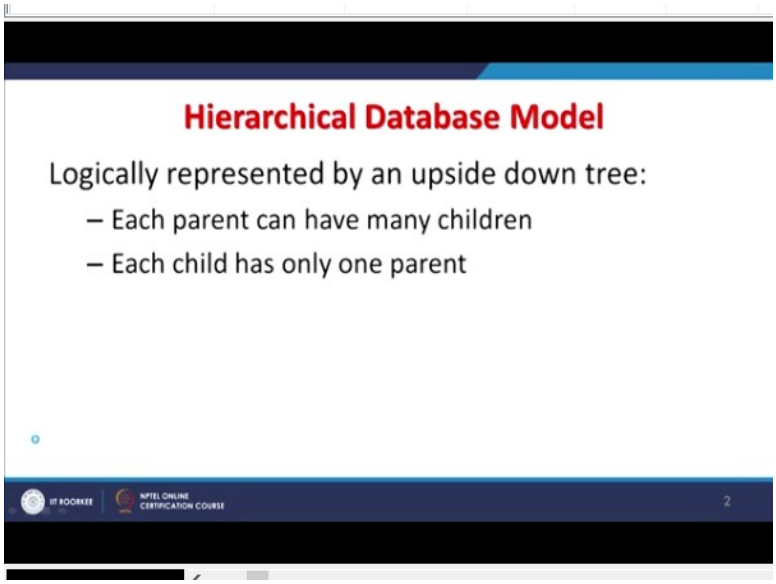
Geographic Information Systems
Prof. A. K. Saraf
Department of Earth Sciences
Indian Institute of Technology-Roorkee

Lecture-31
Spatial Database Systems and Their Types-02

Hello everyone! and welcome to our discussion which we left in the previous lecture that is a spatial database system and their types. So, this is the last part, the second part. And as in the previous lecture, we mentioned different types of database models and hierarchical, relational and network. So, all these we are going to discuss which has got direct relevance in GIS. There might be some other database models but if they are not applicable in GIS are not used then we will not discuss.

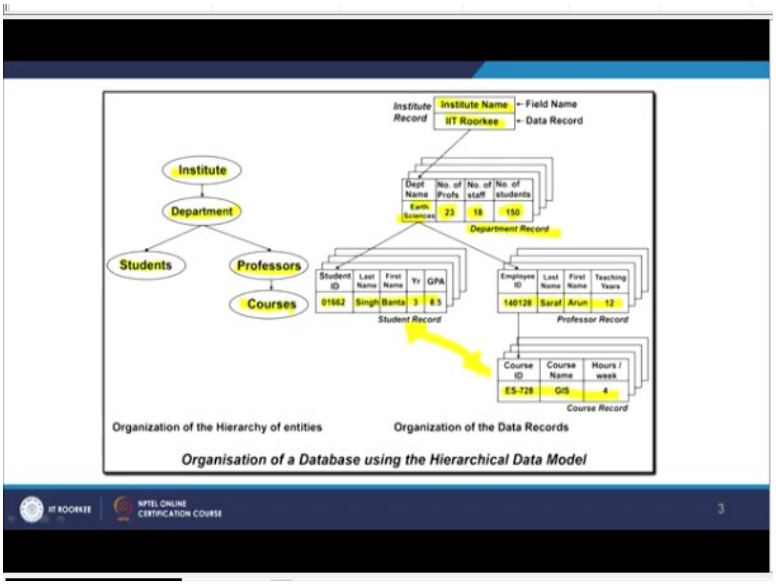
Because this course is not on database management system but this course is on GIS and of course, database management system is a part of GIS. So, only relevant part we are discussing here. In hierarchical database model, basically things or data is logically represented by an upside-down tree and that is why it is called hierarchical that each parent can have many children. So, each you know, say folder can have many files similar like this. So, each parent can have many children and each child has only one parent.

(Refer Slide Time: 01:45)



So, and different files can have only one folder that kind of thing.

(Refer Slide Time: 01:50)



I am taking example from our IIT Roorkee setup and this is represented with like here we are having the main parent is Institute then an institute is having several departments but he or only one department is shown. And then each department will have students of course then faculty, professors and courses. So, if we want to store all that data in our database management system and that is in hierarchical model then you know things are stored like in the file.

And so, institute name and other things then department name, department record is there. Like Earth Sciences, how many professors are there? How many staffs, students and so on? Similarly for each student, we can have their ID, enrollment number and then we can have their records also. Similarly, we can have also implied record as well here and then finally, we can have the links with the courses.

Now, in this model which you might have realized that there is no direct link of the courses and students. And also, there is no possibility of seeing or offering students some elective courses of some other departments which may be interesting for students or useful for students. So, in this kind of model, that kind of information is very difficult to maintain or restore.

Now if we move to different model then obviously, there will be some possibilities. So, first we will see that what are the advantages and then later disadvantages with hierarchical models. So, these advantages as you know that very simple.

(Refer Slide Time: 04:02)

So, conceptually if simplicity is there. Then database security and integrity are also quite good in that sense because people are not accessing very easily the data or cannot access the data. Data independence so data is lying in different locations and it can be modified by individuals. Then efficiency wise okay. It is efficient as well. What are the disadvantages; there are major disadvantage like complex implementation.

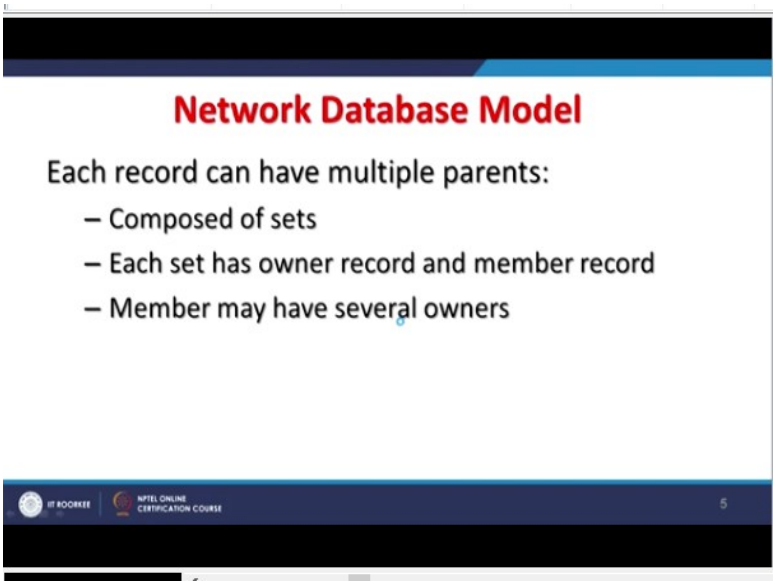
When we try to implement one example, I said that if a student would like to take an elective course or some other department of his interest then he cannot do it in this database management. And difficult to manage and lack of standards because simple things are being kept in separate files or folders and no standards are basically followed. And lack structural independence. Structural independence is not there and therefore, data retrieval sometimes may be very-2 difficult.

And application programming and use complexity. So, whenever we want to query the system or develop an application program for that, there will be complexity or problems. And of course, implementation limitations, there will be limitations that we cannot implement or this is not directly linked say with students and courses and so on. So, the courses are linked through faculty who are teaching those courses but not directly related with.

So, implementation limitations are very obvious in case of hierarchical data model. Now, the second type of data model which was mentioned in the previous lecture is the network database model which is a very good way of representing the data but there for certain

specific purposes. If the purpose of using GIS for network related things, then one should always keep or maintain the data set say in this network database model.

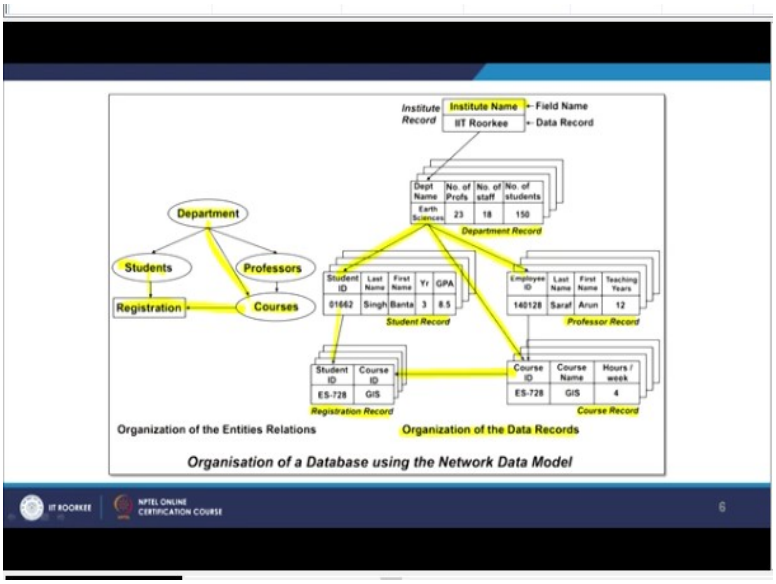
(Refer Slide Time: 06:09)



Now here, each record can have multiple parents. In earlier, we had that each child have only one parent but here it is reverse that each record can have multiple parents and it is basically the network database. When we have been discussing network analysis in the analysis part of GIS in a few lectures before, at that time these points were raised. So, now how they are stored in our network database that is what is being discussed.

So, composed of sets and each set has own owner record and member record. So, lot of information will be required to store in network database then only network analysis can be performed and member may have several owners.

(Refer Slide Time: 07:00)



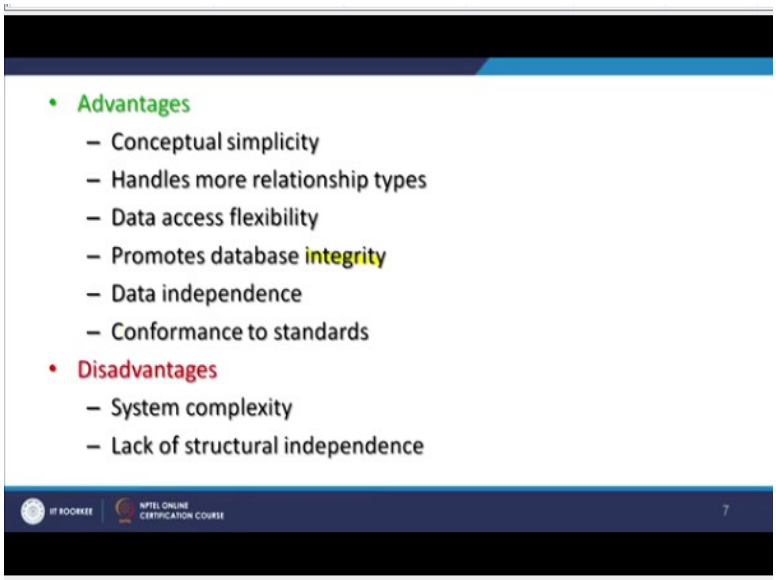
So, in that way compared to hierarchical model that representation or storing the data, managing the data then a network database model is little complex. Now, I am taking the same example but how it will be represented in a network data model that is what we can see. If you recall that there were institute and then department and students and professors, courses. Now through registration of the students or their enrollment number, the courses are now directly connected.

And of course, courses are also connected with professors and courses are also connected with the department. So, student of other department would like to take courses and if database has been managed like this then there should not be any problem. So, in form of like said that inverted tree. So, if we see in the organization in form of the data records, how it will be done, almost same way; you are having institute name then you are having department record, you are having student record, you are having professor record, you are having course record.

This is all common as compared to as it was also in hierarchical data model but what are the additional is the registration record which is being kept here about each course and student's data is linked. So, this linkage will provide some additional information and additional access and network links also. So, here the department is directly linked with the courses, department is also directly linked with the students and department is directly linked with the faculty also.

So, if the database is organized in this network data model, then it would be possible for any student of the institute to take course of any department as an elective without any problem. Faculty can access that information or individual department can also have access. Now, let us see what are the advantages and disadvantages of this type of data model that is network. The advantages are that still it is not very difficult. Relative to hierarchical, definitely it is complex but otherwise conceptually still simple. Handles more relationship types.

(Refer Slide Time: 09:29)



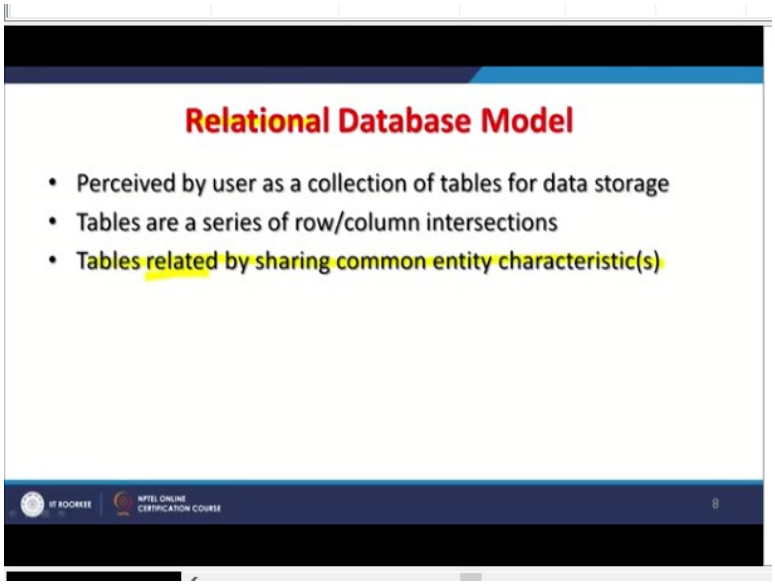
Because there it was very much restricted in case of hierarchical model but here a lot of relationships between courses, teachers, faculty and so on so forth can be handled. And data access flexibility. And student can access the data of the courses or faculty. Faculty can access about the data of courses. Students of other department, faculty of other department can access, that kind of accessibility is there. And promotes database integrity.

So, here we are in a typical database management system and therefore the integrity can be maintained quite good. Data independence of course, it is also there like in hierarchical data model and conformance to standards. Now in case of hierarchical model, it was not possible to follow the standards. Now, here we can follow some standards and what these standards will allow that if we want to integrate our database with some other data sets or databases, it would be easier if we follow the standards.

But if we are having our own standards and not following some real standards then there will be problem and conformance will not be able to achieve. Now, the disadvantage is little less as compared to your hierarchical model. One is the system complexity might be there

because you have to maintain the data. And then connectivity who would have access what. So, that will be there. Lack of structure independence is same disadvantage as in case of hierarchical model.

(Refer Slide Time: 11:16)

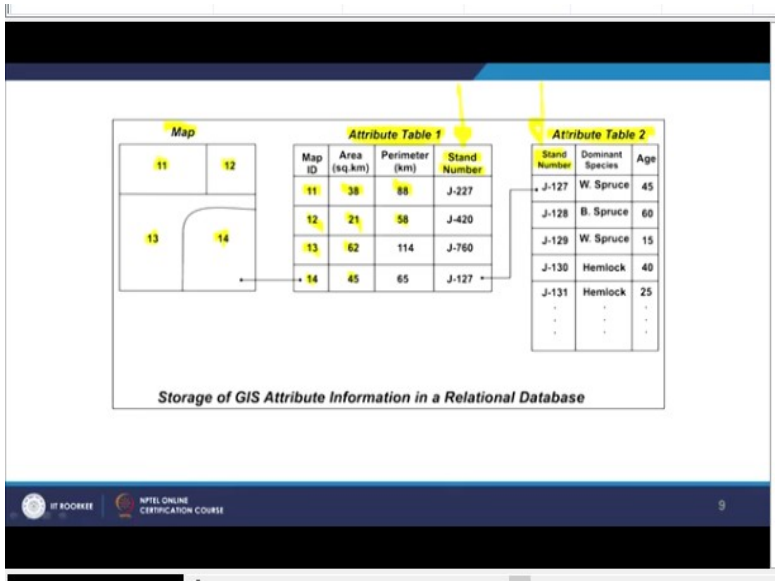


Now, the third type of model is relational database model which is the most commonly used in GIS and very-2 useful compared to other 2 database models which we have discussed so far. So, this is basically perceived by users as a collection of tables for data storage. You know that for each a spatial objects or layer, you are having a table or attribute. If you are having point data, you are having attribute table. If you are having poly-line data, you are having an attribute table. If you are having polygon, same you are having attribute tables.

The information is stored as a relational database but in form of tables at least on the display side. So, tables as you know that the series of rows and column intersections and then tables are related by sharing common entity characteristic. This is very important point and that is why there is a relational thing are. The relational word is basically that is the key here.

That if I am having 2 tables; suppose I am having a data set of say villages of one state and I am having data set of villages of other states or maybe some other similar kind of thing.

(Refer Slide Time: 12:48)



If I want to link with these 2 tables then I have to have a common field which should be common in both the databases or both the tables. If that is there and that common field is having also the common or same identical characteristics, we will also see examples little later. So, if that is there then the relationship can be established. For example, again that I am having soil sample data and for the same area, I am also having land use map.

Now I want to relate this soil data with the land use map that means in both my databases or in both my tables, I should have a common field which should have the common or identical characteristics. If that is there then I can relate these 2 tables very easily. So, this is what it allows. Now let us see the example. Here map is same. Here I am having a map which is having just 4 units polygon map.

Here I am having attribute table 1. And now I get for the same map from some organization or somewhere or within my organization from some section, another table for the same map which is having some other information then what I had in my database or in my table. So, like here the map ID, 4 units are there or 4 polygons so each one has been assigned ID. So, 11, 12, 13, 14 what you are seeing here.

Then against each of these polygons, I am having area information. I am having perimeter information and a unique field which I am having is the standard number; a common entity which I am looking. So, this is my common entity which is a standard number because this stand number is also given in the table 2 in same characteristics or in same format.

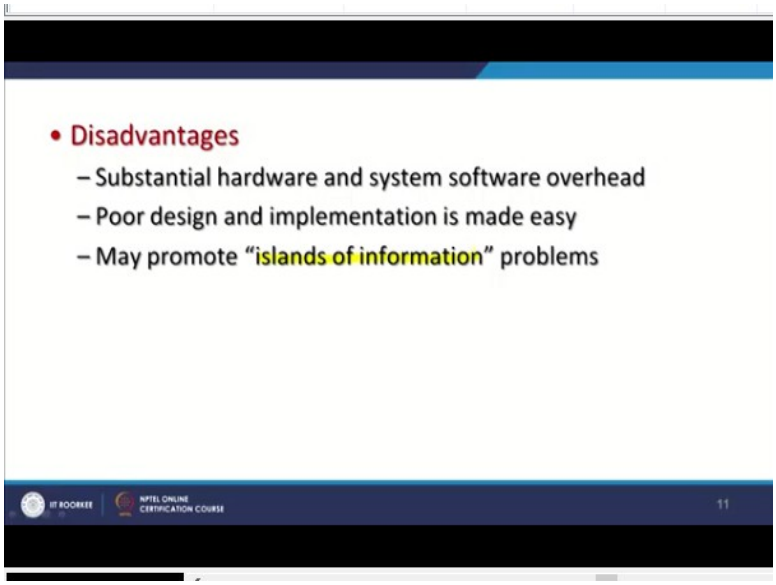
A be would like to talk in like excel point of view. So, that field or that column is having the same characteristics or same format. If it is there then I can correlate or I can relate these 2 tables very easily. Like here, what had been done that unit 14 is the only out of these 4 polygons where I am having J-127 means these stand number are not matching; other 3 stand numbers, top 3 stand numbers are not matching except this.

So, now it is related and I can enrich my table by relating this data also or I can keep as separate but still it is related. Rest is not getting related. There might be a scenario that all rows may get related with the table 2 also of table 1. So, this is how the relationships between 2 tables or 2 databases can be established very easily. So, you need a common field having unique ID. For example, if I give you an example of say, we are having one database of students and I am having that one field which is enrollment of the students.

Now, I am having another database which is about the courses but in that database also I am having one field which is about enrollment number. So, whenever I want to check that which student is taking which courses, I can use that common field of these 2 databases and can relate and develop a new table which will let me know that which student is taking which course.

So, this common entity or common field and characteristics will allow me to relate 2 databases or 2 tables that is what the essence here. That is why it is called relational database and this is what exactly we keep doing in GIS that we relate the data or find out the relationship between 2 databases but the only and main requirement is that both or all databases whichever have to be related should have a one common field with common or identical characteristics then relationships can be established. Now advantages of a relational database structural independence and improved conceptual simplicity.

(Refer Slide Time: 17:31)



So, it is having much better, only one has to understand this common field and rest is ok. And easier database design, implementation, management and use. If database design itself is easy then other part will also be easy. So, implementation, management and uses also easy. And ad hoc query capabilities with SQL. Anytime we want to access any information against any field for any record, we can do it very easily.

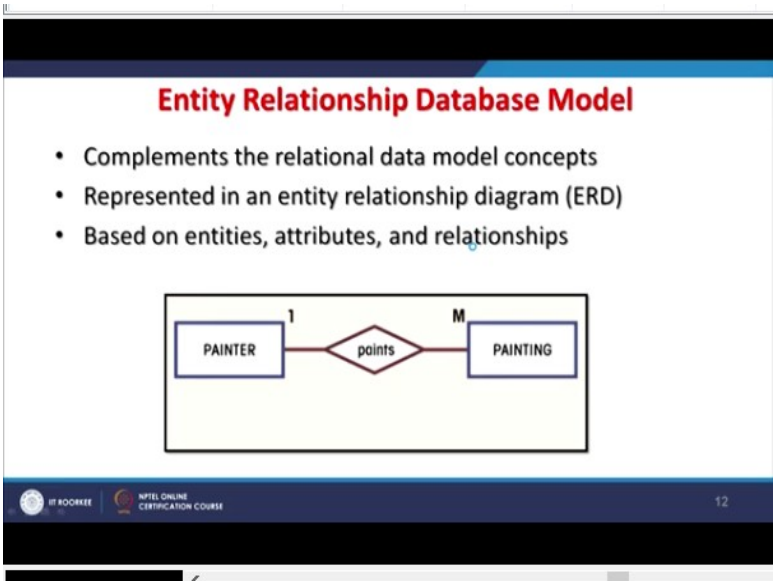
So, ad hoc queries capabilities are very good compared to another database. The most difficult ad hoc queries capability would be network one. Now powerful database management system that is why it is very popular in GIS and have been implemented extensively even in this popular GIS software's. Now disadvantages; among disadvantages, the first one is the substantial hardware and system software overhead because the data is might be kept at different places and anytime you want to relate that is possible.

Poor design and implementation are made easy. And may promote Iceland have information problems; these islands of information problem meaning here is that different databases may have the same information or same field, not one common field but many common fields. And one database has been updated, others have not been updated and therefore when we relate these 2 tables or multiple databases, we may have the problems.

So, these islands of information problem may be there. Now, apart from these 3; hierarchical, network and relational database model, there are some other database models which are also coming up or being implemented into GIS that is entity relationship database model. And

here basically, it complements the relational database concepts and it is represented in an entity relationship diagram that is ERD and it is based on entity, attributes and relationships.

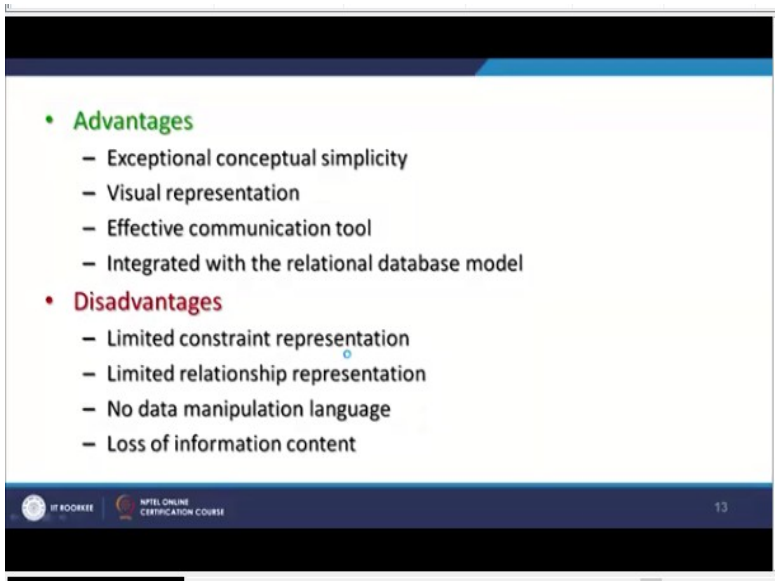
(Refer Slide Time: 19:41)



So, vector entities for example then attributes, non spatial data out tables and then what is the relationship between these 2, spatial and non-spatial? So, like here we are having a like a painter 1 and this is the painting but paints are having a column relationship between painter and painting. So, this kind of ERD; the entity relationship diagram will allow us to complement or the things which are not possible in relational database model might be possible in this entity relationship.

What are the advantages? The advantage among them is the exceptional conceptual simplicity as also in case of relational database.

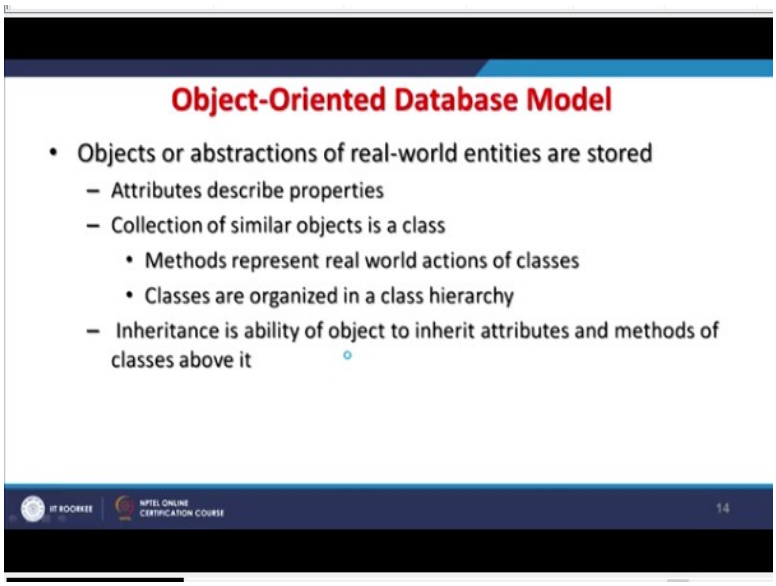
(Refer Slide Time: 20:30)



Visual representations are also possible. Effective communication tool; easy to communicate easy to explain to the people how things have been organized. Integrated with the relational database model, it can be integrated because it is complemented not competing with the relational database or RDBMS. Now, disadvantages; limited constraint representation.

That might be there in some cases. A limited relationship representation, it will not be possible to link or create something like network kind of database. So, simple relationships can be represented but as if you go for little complicated one then the best model would be the network one. And no data manipulation language so if you want to do some modifications or manipulations then it is not possible. And loss of information content sometimes you may encounter this problem if database has been maintained here.

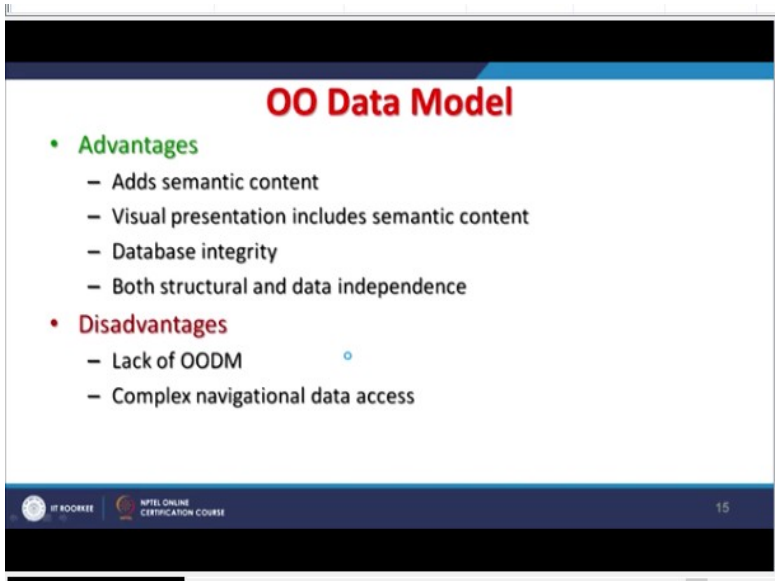
(Refer Slide Time: 21:33)



Now, one more type of database model which we can discuss is object oriented or database model or some object-oriented programming, they people call as OODS. So, it is the same concept that objects are abstraction of real-world entities are stored. This is what we are doing in GIS also. We are representing the real world in abstract form. So, this is what it is being done. Now, attributes describe properties.

So, whether it is a raster or vector, whatever the attributes they are describing what phenomena or what information it is? And it is again collection of similar objects in a class, methods represent real world actions of classes and classes are organized in a class hierarchy. So, in object-oriented database model, this is how things are done. Now, the second point here is inheritance ability of objects to inherit attributes and methods of classes of orbit because this is object oriented so we handle each entity quite distinctly as compared to other database models.

(Refer Slide Time: 22:47)

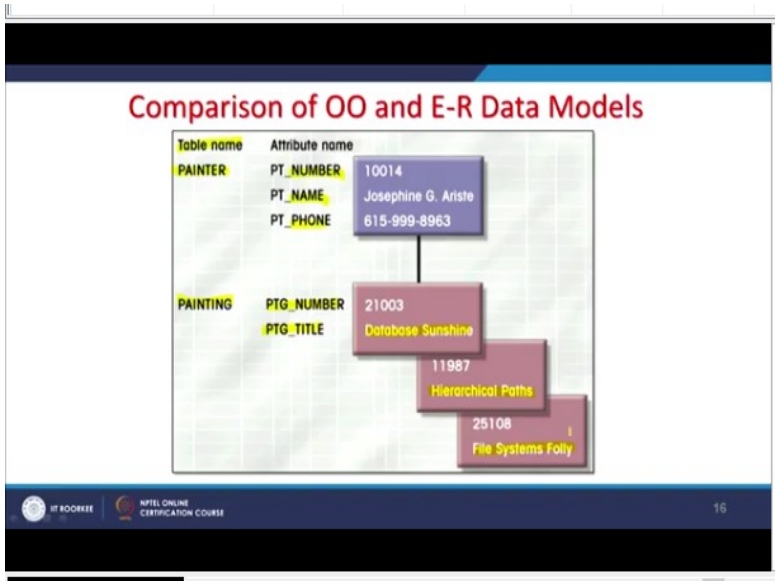


So, this object-oriented database model advantages are that adds semantic content and visual representation includes semantic content. And database integrity is quite high in case of this object-oriented database model. Both structural and data independence is also there which is not common in previously discussed database models. What are the disadvantages? First of lack of object-oriented database management that is not easy very much. Complex navigational data access.

So, if we want to develop a query language then because of this complex navigation that would be a problem or difficult. Steep learning curve; it is not very easy to learn this object-

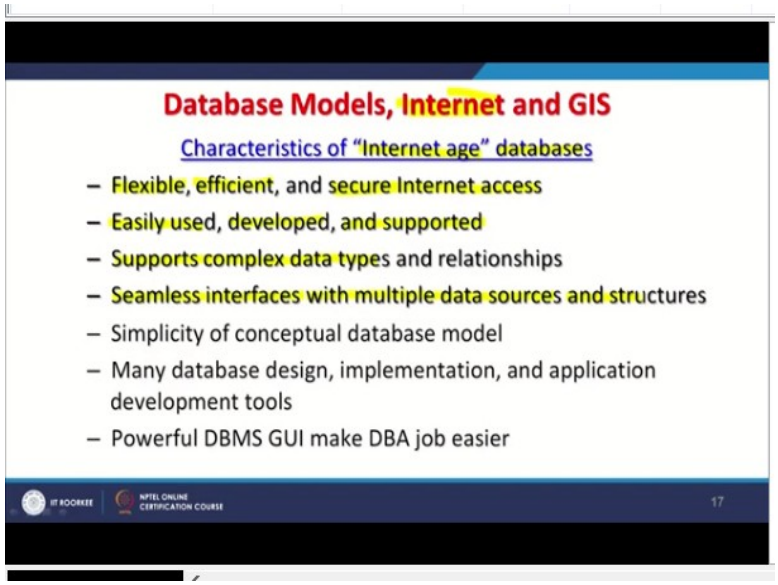
oriented database model. And high system overhead slows transactions because you require a much more efficient computers and therefore, they make add extra cost or if systems are known as having high speed, then that may slow down your transactions or processing. Now if we compare the previously discussed 2 database models; object oriented and this entity relationship data models.

(Refer Slide Time: 24:20)



And then what we find that here like in this table name, what we are having a painter and now painter number, painter name, painter home and again this is a painting number and painting title. Now here we are having database, hierarchical paths and file system folly.

(Refer Slide Time: 24:45)



Whereas in case of other database model, the things might be different. Now, we discuss about what is the real future about this or what is really happening also currently with

databases and GIS and that is what happening through the internet. So, internet here is in between because now we are in the internet age and especially because of this Corona pandemic, people are resorting for many-2 things which are online.

Earlier before Corona pandemic, we used to do a lot of things, lot of operations offline. But things are going online. That means we are now using more internet and handling lot of other things which we are not really either possible or we are not paying attention to that. So, this is a basically internet age and databases that means it is not necessarily the database should reside on my computer. Database can reside on some others computer.

And I should have access and that access should be a secure internet access. And also, that access should be flexible and efficient. Let me give you an example like Google Earth. Many of us have started using on regular basis, same with Google Map. So, if I take example of Google Earth, basically what Google Earth is having 2 main databases, one is in the background which is the digital elevation model of the entire Earth which is coming from SRTM; shuttle radar topographic mission digital elevation model.

Directly on the screen, we do not see but we get the elevation only through that SRTM DEM. And second very big databases which is a satellite image. And these satellite images are of course, your reference. And they are dragged over your model of the earth. So, as we zoom it, we keep going higher and higher spatial resolution. And what basically is that when we install the Google Earth, we are installing just the utility of the Google Earth; a small app of the Google Earth.

But the database is not residing on our computer. Database is being accessed by that app which is Google Earth, installed on our machine through internet and that database which we are accessing is really very-2 huge database. So, lot of technologies goes; data compression is also there, pyramid might be there. And other things are also there. But the main thing is that the database is internet based.

And that is not deciding on my computer but it is residing somewhere else and through Google app, we access that huge data. We can access and we can develop even fly through and see things in 3D or satellite images and so on. So that if we go for internet based

databases then it has to be flexible, it has to be efficient. And of course, it has to be the secure. Otherwise, people will do some mischief with those.

Now, these also should be very easy to use; user friendly. So, Google Earth is one of the examples which I have taken to explain this thing that Google Earth is very easy. Anybody, a child even he knows how to move mouse and other things, immediately he can start using Google Earth. So, user friendly and then accordingly, it has been developed and supported and upgraded on regular basis.

So, if any program has to be there, any database has to be accessed by users through internet. These are the qualifications. Otherwise, it will never take off properly. Now it is also supporting complex data like Google Earth. And not only the digital elevation model, not only the satellite images, lot of other datasets one. And also, you can after doing your own data processing, you can paste your data over the earth also and can check the geo referencing.

So, for geo referencing purposes, we can also imply a Google Earth. And we can collect ground control points from Google Earth. And after doing geo-referencing, we can paste that image on Google Earth and can see whether it is fitting with the Google Earth image database or not. If it is not, that means something is wrong with our geo referencing.

So, very easily one can make check and find out the errors even in geo referencing implying Google Earth. So, this kind of thing is very much only possible if it is available on internet and flexible, efficient and secure. Now, this is seamless interfaces with multiple data sources, satellite images, your digital elevation model maybe the data about the earthquakes, maybe data about the culture, roads etcetera all are there

And then simplicity of conceptual database model; at least at the user end, it is very-2 simple. And many database design implementation and application development tools are involved in the background which are provided through internet the access to these databases. And of course, these are having very powerful database management system. They are having graphic user interface is also very easy.

So that it becomes a completely user friendly and that is why it is one of the very popular tool available on net for seeing satellite images, seeing things in 3D and getting other related information. So, this brings to end of this discussion about the database management system in GIS. Thank you very much.