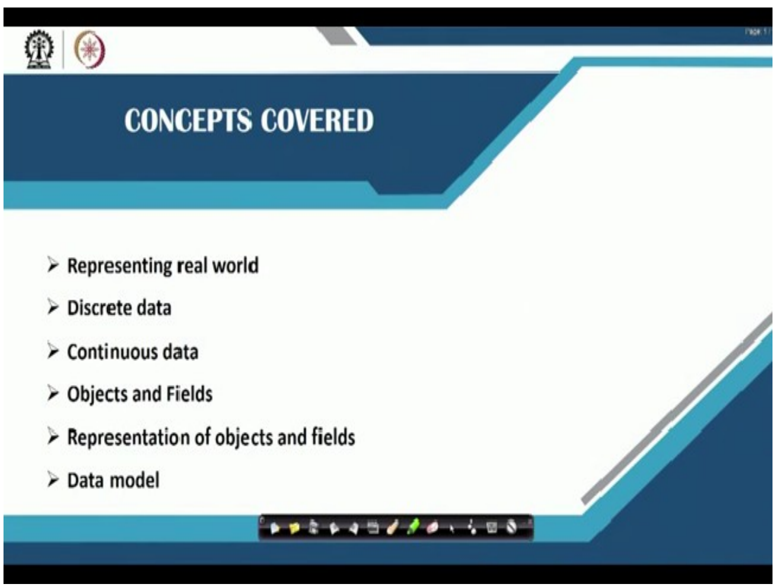
Geographic Information Systems
Prof. Bharath H Aithal
Ranbir and Chitra Gupta School of Infrastructure Design and Management
Indian Institute of Technology-Kharagpur
Module No. #03
Lecture No. #11
Representing the Real World

Hello namaste, I am back here with the next set of lectures on geography information system. So, in specifically in this particular module industry we would look at what is the data, what kind of data that can be handled by GIS, what are the different data models have been speaking about data models all through the last 2 weeks, but today we will understand or may be in the set of this lectures, 4, 5 lectures we will understand what are the different data models with examples.

We will also look at what are different data types that GIS can handle. So these are very 2 important aspect if you have to work with GIS. The first 2 weeks was to understand what GIS really means, but now onwards we will start handling data okay.

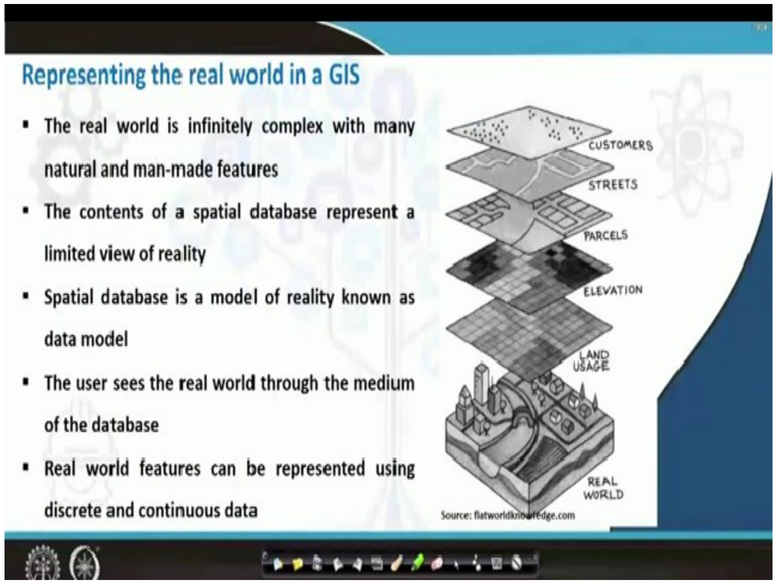
(Refer Slide Time: 01:17)



So in today's lecture so we will be looking at how do we represent a real word, we will also look at what are different datas for example looking at discrete data is one set of data. Other one is a continuous data. So we look at both of these, then we would look at specific aspects called as objects and fields. These are extremely important in terms when you are looking at the data and how you actually represent an object and a field.

So ah, how the representation matters because that is what is actually be converted as a data model and finally a data model, in the next set of slides we look at what are the different data models and the examples with that.

(Refer Slide Time: 02:01)



So now when in the previous week, we did understand what do you mean by a representation of a real world. So when we look at real world, real world is actually a infinitely extremely complex okay. It has both natural resources, it has all changes made by man-made and for a and various features that has been completely change as by the man, but how do we represent it. So we can represent it each of the aspects in the data model as entities.

So now when we are looking at this when once we converted to a data model we should when you have a spatial database it as I said it only represents a limited view of reality because everything is in the form of entity and relationships to the database. So, when we look at the spatial database it is a model of reality known as data model. So whenever you are looking at the data model you are actually spelling out that this is a model of the reality and is representing the real phenomena.

The user sees the real world through the medium of a database, so now database is a heart and muscle of your entire system, without database your data model does not exist neither your real world. So if the real world has to be transformed into a data model then you need to have a database and the ways you can retrieve any information or do any quantitative and qualitative aspect in terms of having information. So real world features can be represented using both discrete and continuous data.

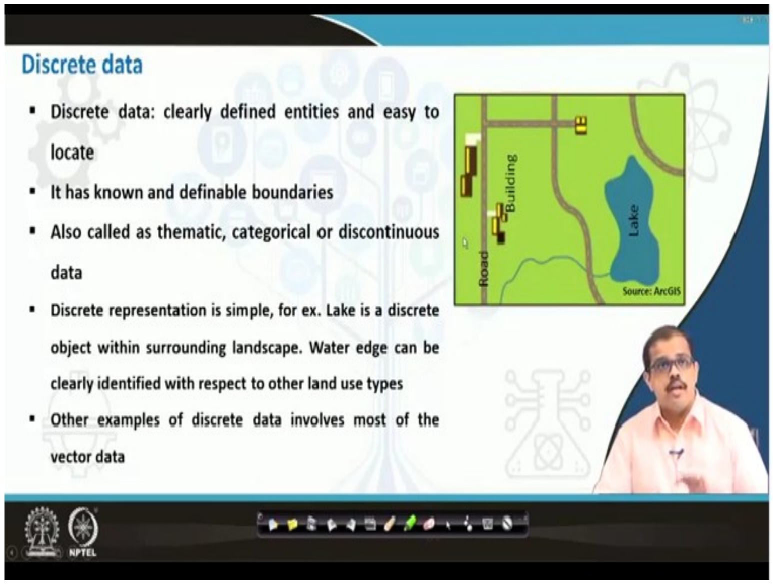
So whenever we consider data there are 2 types of data. One is discrete data, discrete points and another one is a continuous data. So on your right-hand side, I have a set of layers. So this is the real world okay, whatever we have here is a real world. So now when I have a land usage okay. It is in the form of squares or pixels or in terms like last week I spoke about raster okay, set of number of a collected pixels at m cross n, matter x is nothing but your raster image okay.

So each of these pixels we have own specific values based on different intensities of or different values come back to this how it is represented, how it is captured etc. then you have an elevation layer, which is again a raster, you have a parcel layer, which has both a collection of a raster and a vector okay, you have lines, you have set of values, parcels can be only vector parcels can only be raster parcel can be both vector and raster.

So you have this here representing mainly some part of the raster whereas this is a vector, you have this as a vector layer and you have this customer's as again a vector here. So when we look at this each of these points here that is representing the customer layer or a street that is represented here is not a continuous data. It is a point data, number of points make a line segment. So when we look at this it is represented as a discrete data okay.

Whereas this phenomena wherever you see, you have elevation data it is a continuous phenomena, it is not a discrete phenomena, but a continuous phenomena.

(Refer Slide Time: 05:35)

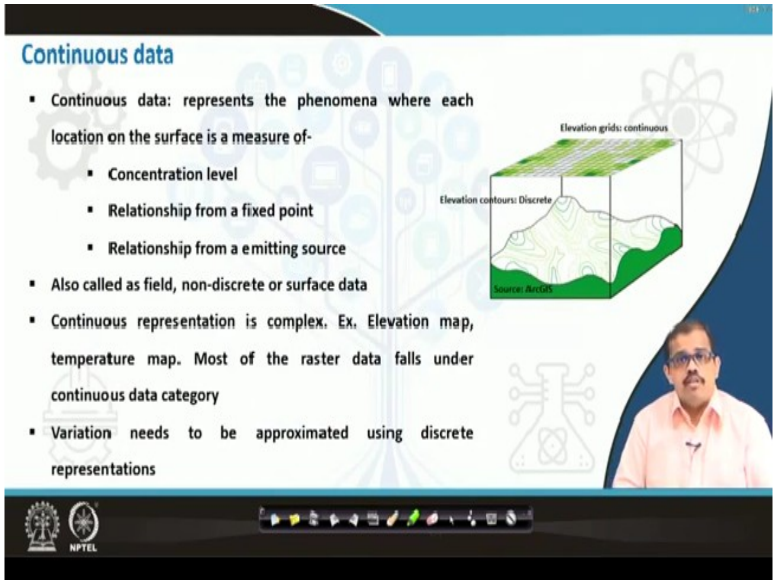


So when I say discrete data it has clearly defined by entities and easy to locate. So, it is extremely easy to locate and it has defined by entities, if every entity has its own value then it is a discrete data okay, it is known as a definable boundaries. For example here if you look at the lake, lake has a boundary okay, it has a specific area, it has may be lake has different characteristics.

So, all of these are discrete datas okay, edge of the lake is it clearly is identified with respect to the other land uses that this concert as a discrete quantity okay. Now if let say I have this layer then I have an elevation layer or any other layer, which has a continues layer that is put on it. So both of these layers represent different quantities, but when you look at this lake it is a discrete data.

But when you look at elevation, which is overlaid on it then it has an completely and continuous data. So that is a continuous data, this is a discrete data. So GIS has tapped efficiency or capability to handle both discrete and continuous data together.

(Refer Slide Time: 06:54)



And when you look at continues data I have given I just spoke about the elevation part of it if this is elevation map you can look at the elevation of the grids continuous grids here. So this is nothing but your continuous data, when I represent a phenomena with each location is surface of on a surface is a measure of may be a concentration level, may be relationships okay, may be relationships on a fixed point or from emitting source.

So, it can be any kind of thing. If you are looking at remote sensing data it can be a relationship between the emitting source because the amount of light that is reflected back and you removed all other interferences and the light that actually reaches the sensor back is the amount of energy that is recorded as a value in the pixel. So that is what is the relationship with the emitting source.

So amount of for example, if you have a satellite here and you have let us say my left hand here is a satellite and my right hand here is your sun okay. Now you have a solar energy it hits the earth's surface and now as it is being captured here before hit the earth surface that is atmosphere, which actually has certain ways of observing certain amount of energy here. So some part of the energy is lost when it hits the earth surface there is some part of them is reflected back.

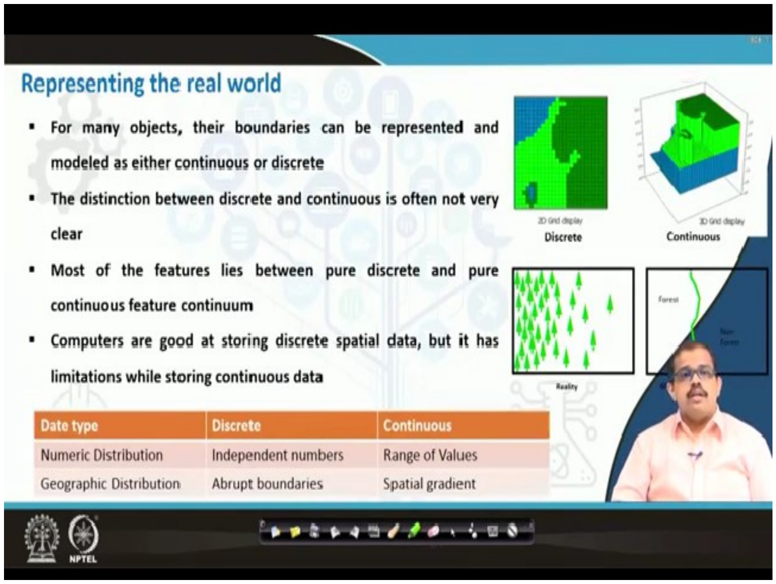
Again you have at now some part of the energy that is removed in the atmosphere and finally, when it reaches the sensor you have very less energy that reached the sensor. So the reflected energy when it reaches the sensor it is the amount of energy that is recorded. So based on the kind of reflection the amount of reflection the you have the DN values that is actually registered on this particular platform okay.

So now when you are actually looking at this it is also saying that you are actually collecting every point on the earth's surface as a source of a pixel. So based on the size that your sensor is able to look at the ground will be able to collect so much of information. If I say my sensor is 30 meter by can sense so much of area and is a 30 meter resolution it means that the details of 30 meter by 30 meter is captured by that sensor okay.

So when you look at this continuous data is these are also called as field non-discrete or surface data. There are different types of services data which I will also speak to you probably in the next lecture. So when continuous representation is normally complex, so you have to look for example when you look at elevation map temperature maps, etc. is a quite complex in terms of representation and interpretation also. Most of these are raster data may fall in the continuous data category.

And most of it every all of it basically falls under there are continuous data categories and it is different it is not so easy to differentiate each categories in a very smooth way as we do in a discrete data. So variation needs to be approximated using discrete representation certain times. So what you do is you convert the continuous data into a discrete data in order to show the variations, that is extremely important when you are actually handing the continuous data okay.

(Refer Slide Time: 10:32)



So when we are trying to represent a real world as I said real world it first of all of complex world to represent and moreover it is dynamic. So, to capture this dynamic complex real-world, you need to have for data model which can handle every kind of data. So when we are looking at the boundaries that can be represented model as discrete or continuous values we have to look at what kind and how it has to be captured.

So the first thing that you do is you look at the distinction between what is a discrete and what is a continuous values that are there on your data set or in your study area. For example in this image when I am seeing this particular image this is a discrete image okay, so you can easily find out, you can make this particular raster data discrete easily because you can understand what is the variations in this particular raster data.

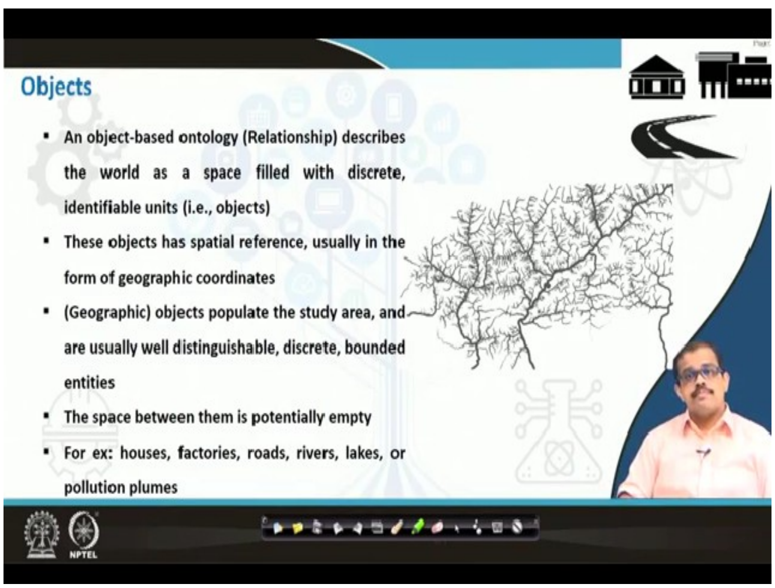
Whereas when you look at the continuous data, it is difficult to find out the variations. This is another example of discrete data, if you look at this is again a discrete data. So most of the features in a continuous or a discrete data have the certain values associated with it and when you look at the real world it does not have anything that is separately discrete or separately continuous. So you should look at something that is in between a discrete and continuous.

So when you are looking at the real world, you will not be given anything that is completely discrete or completely continuous. So your data or the database should be able to handle or a data model should be able to handle both discrete data and a continuous data in an effective way. And when you are looking at the spatial data storing of a spatial data in your computer especially in the form of a discrete is much easier.

But when there are your continuous data, your computer has certain limitations by which we will also see how it is quite limited. So handling and storing computers have a lot of limitations in looking at continuous data and when we look at data type, so when you have a numeric description if it is a discrete this are independent numbers and when you look at continuous it is a range of values okay.

So when you have a geographic distribution it is abrupt boundaries and when you look at continuous values it is a spatial gradient you have a gradient of values okay.

(Refer Slide Time: 13:26)



Now when you look at this discrete or a continuous data the first thing that we have to look is objects. Objects are extremely important, it is equivalent to your entity relationships okay. But when you look at an object object-based relationship or object-based ontology it describes the world of space that is filled with discrete identifiable units. For example here, if you look at this image, this is a discrete data.

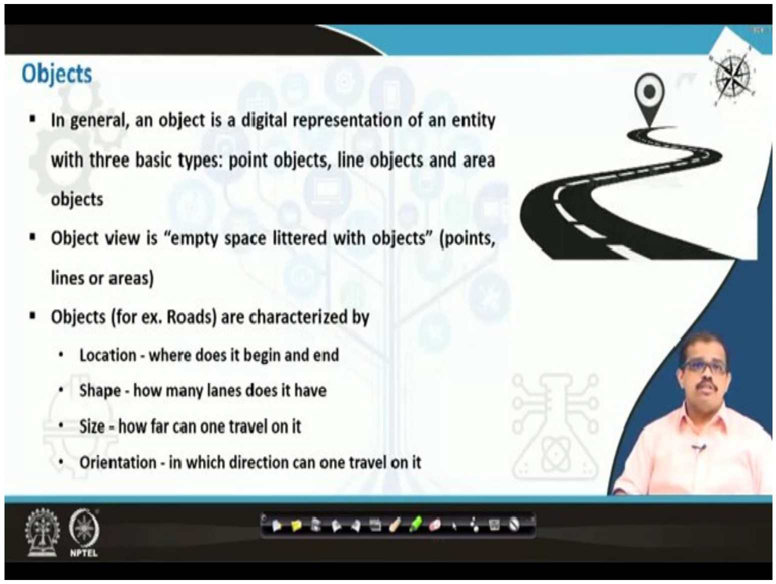
These 3 are 3 different discrete data, so these are identifiable units that are present there. These objects, each of these objects have has a spatial reference, usually in the form of geography coordinates, geographic objects. Now we have objects, so, there are 3 different objects in that image, but in case you add geographic value to it, it becomes a geographic coordinate and this geographic objects populate the entire study area.

And most importantly is that when you look at geographic coordinates these are quite distinguishable, so you can distinguish each and every object, for example here in this image, you can distinguish this is an house this is a single storey house, whereas we can see it as industry that is next to the house right. So when you look at it these 2 are discrete quantities okay, and are distinguishable discrete and have a bounded entities.

So when you look at this, it is a bounded value the space between them is potentially empty. That is what we mean by discrete continuous has continuous values if there is a house at struck then you have another house at stuck, any of the industry so that is a continuous value whereas discrete you have a space in between them. So for example, we have houses, factories, roads, reverse, lakes of pollution plumes.

So continuous efforts there then it becomes a continues data you can distinguish the boundaries but if it is discrete, you have a bounded value, you have a space, then you have another may be another object, then it becomes quite discrete okay.

(Refer Slide Time: 15:49)



So when you are looking at this object is a digital representation of entity, you can call object as an entity, but in the digital representation becomes an object, but in a real representation it becomes an entity okay, when you look at the object it can be defined in 3 different ways, you have point line, point objects, line objects and area objects. So object view when say that it is an object view is an empty space littered with objects.

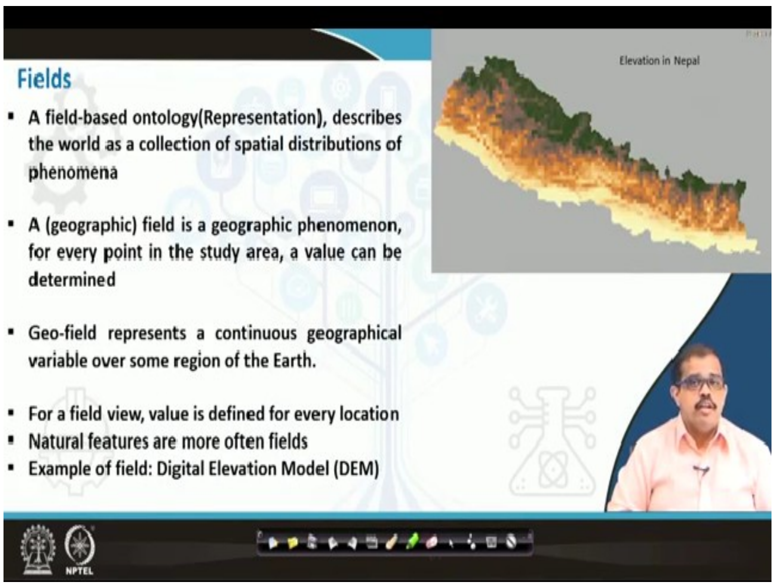
So, if you considered this if you have this entire slide, if you do not have anything written here it is an object in a database okay. So now I put in some data on it, some object on it, it will be a point object, it may be a line objects. Line object is representing this road, point object is representing this location here as may be an industry okay. So each of the entire board is nothing but an empty space okay.

But it is a collection of point objects, line objects and area objects, it may be collection of all of these, but even one of these. So that depends on what kind of object it is, now objects for example, if we consider it as roads are characterized by may be location where does it begin and where does it end because road is characterized by line segment. Line segment is made up of 2 points. So you need to have a starting point, you need to have an ending point.

Then shape how many lanes does it have, it is a 2 lane road, 3 lane road, 4 lane road or a 6 lane road, so depending on how many lanes that it has, then size how far one can travel on it. So number of line segments that connect how far it travels, then you have orientation in which direction one can travel. So that direction is extremely important in case you are trying to find out what is the route structure in the entire database.

If you are trying to give the distance we are trying to give a route, so you need to give the orientation otherwise you will not be able to tell the user that this is the route that you will take to go into this place okay.

(Refer Slide Time: 18:05)



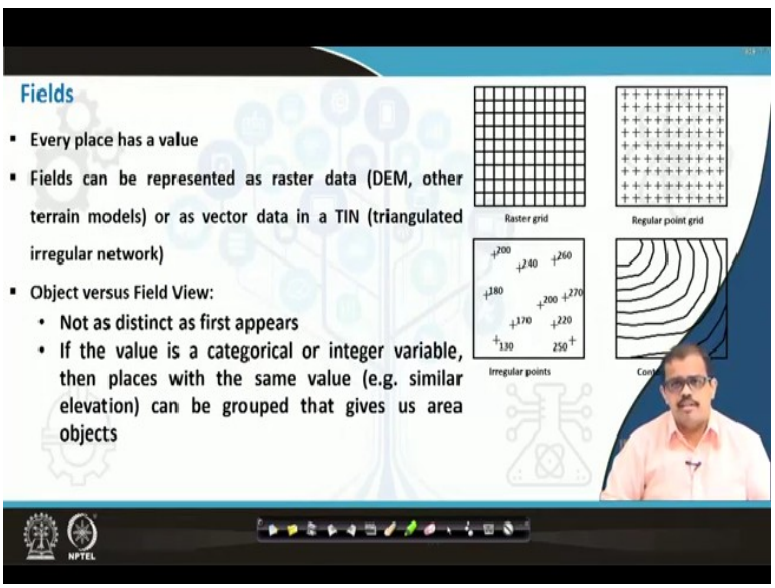
So similarly we have fields or field when I say a field based ontology or field based representation describes the world as a collection of spatial distributions of phenomena or it is a continuous data. For example on the right-hand side elevation detail that I have represented is a continuous data. So when you say it is a geographic field it is a geographic phenomena that is for every point in a study area a value can be determine or geographically there is a value associated as far as earth reference is concerned for that particular phenomenon.

Then geo-field represents a continuous graphical variable or some regions of the earth or if it can have **a** I mean when you look at it is not only just one value can be many value also so but it represents a continuous geographical variable. Please understand this, it is a geographical variable. So, it has a geographical quantity associated with it, but these quantities are variable based on earth's reference okay.

You have to always look at the earth reference when you are actually representing any GIS data. So how what is earth value, how it is represented as a data model. So looking at that it is a geographical quantity also which means it has certain ways of represent, it can tell you where you are exactly, but it will also tell you what you are okay. What is of value, what is this I mean region that is dependent on it.

So all these values are given as a field in a continuous data, when you look at the field of view, it is a value is defined for every location natural features are more often fields, when you are representing the very good example is whatever I have represented here is a digital elevation model, which is actually represent in the form of field or your entire land use data can be represented as a field.

(Refer Slide Time: 20:08)

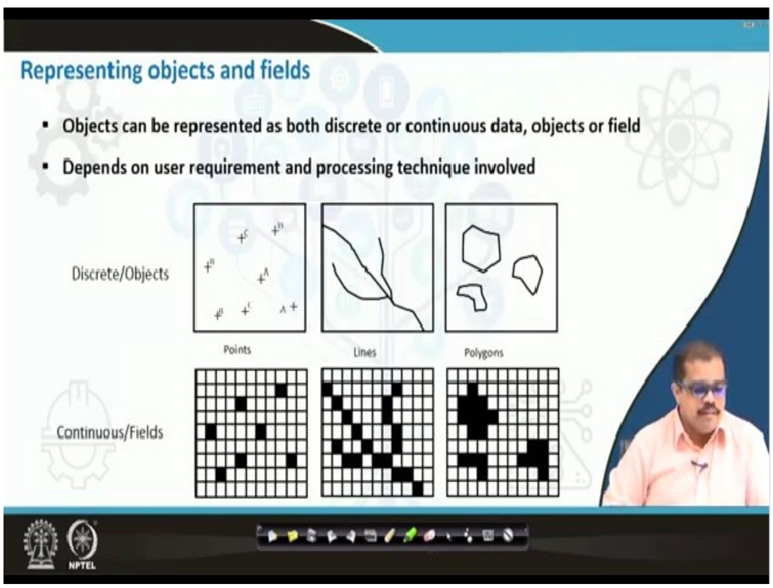


Land use data can also be represented as again it can even represented as an object, in case it is converted as a line point on a polygon. So if it is all land-use features are converted as polygons then it becomes a polyline objects or polygon objects sorry, so that is how we use objects and fields. Now, when you look at fields every place has a value. Please remember this is extremely important.

Every place has a value, fields can be represented as a raster data for example DEM and other terrain models we look at this, it can also be represented as a vector data in form of 10 that is triangulated irregular network okay. Object versus when you look at objects verses field view if it is the first one is as distinct as possible whereas the field is not quite distinct it is a continuous value.

If it is a categorical value or integer variable then places with the same value can be grouped in an object, for example if there are number of polygons with the same value it can be easily grouped okay, whereas if the entire elevation data is same it is not possible to group because it is a continuous phenomena. So that is why the object can be easily grouped whereas the field of cannot be grouped much easily.

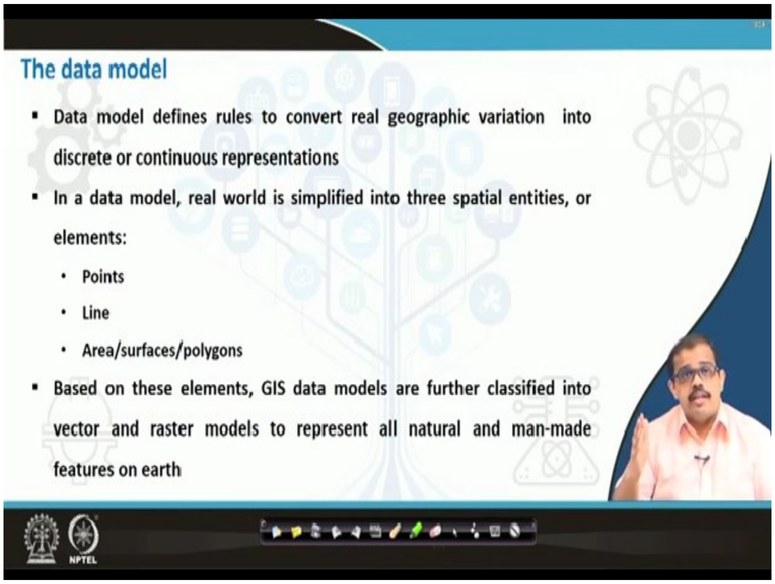
(Refer Slide Time: 21:43)



And when you look at representation of an objects and a field, if it is a discrete value and is called as an objects, then it is point lines and a polygon okay, which can be represented as simple as this number of points okay 2 points from a line okay number of lines number of line, number of line segments that are connected together forms with the closing of value forms a polygon okay.

When it is a continuous representation or a spring about fields, then it as number of squares that are together in a form for matrix okay, it has m number of rows and n number of columns, then it is represented as a continuous of fields or continuous value okay.

(Refer Slide Time: 22:30)

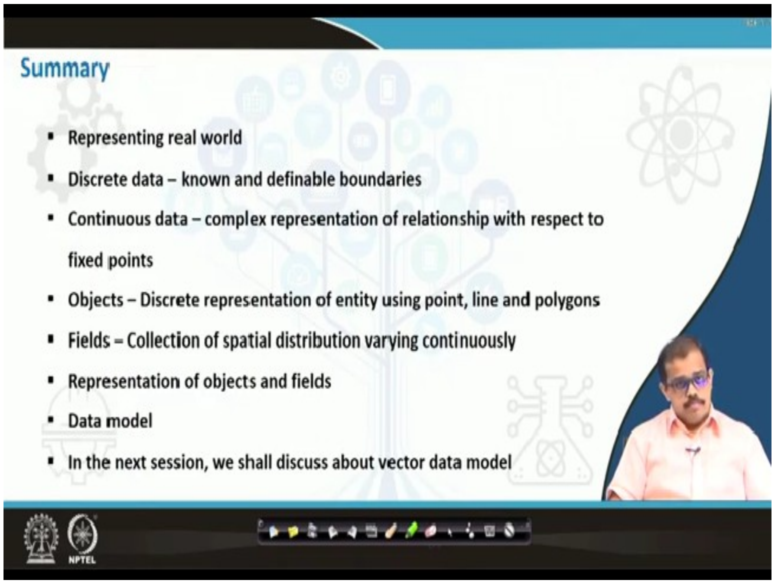


So now we have been speaking about data model, so now what do you really need mean by a data model. Data model defines the rules to convert real geographic variations into discrete or continuous representations which means that you are converting the real world in to either a discrete value or continuous value, without which a data model does not exist. In a data model it real world to simplify into 3 spatial entities it is this point, line and polygons okay.

When I say polygons some people can even I mean put it in the word such as area or surfaces, it can be areas, surfaces and polygons. So point line polygons are the things that this data model signifies. Now, based these elements GIS data model can be classified as vector data model and raster data model. So in a vector data model as I sad is a collection of point, line and a polygon. Raster is nothing but a grided cells, which is actually representing different quantities of based on it may be its reflections may based on its various at representations automatic features.

So, to represent all the natural and man-made features on the earth, so just to signify that what you mean by a data model, it is nothing but rules inverted quotes rules to convert real geographic variation. So please keep this in mind, it is a real geographic variations okay, each of these real quantities are converted as discrete or continuous representations. So you are converting the real value to discrete or a continuous representations. So this is nothing but, what is called as the data model.

(Refer Slide Time: 24:37)



So when we come back in the probably in the next session we look at different types of data model. For example, as I said, it is a vector data and raster data model, will also look at what are different vector data's that are vector data model that is there okay, there are different ways of representation of a vector data. So we look at those types of model, there is also raster data models, different ways of representation of a raster data model.

We look at least one of them how do how does that particular data model represents the entire data space. So with this summarizing today's class we revise what how do we represent a real world. Real world is extremely dynamic and is complex, so this has to be represented in the form of a data model. So normally when we are creating a data model everything is in a form of entity and because of which it may real word.

But essentially not the entire photo scope of the real world okay, so when you look at a discrete data, when you look at the real world it can be divided into discrete data and a continuous data, you can extract it, so when I say discrete data it has a defined boundary whereas continuous data is a complex representation of relationship with respect to a fixed point okay.

So when we are looking at objects and fields, objects are discrete representation of entity using a point, line and a polygon. Whereas when I say fields it is a collection of spatial distribution varying continuously. So you have a huge amount of spatial distribution values. So this values are connected to a geographic coordinates and these vary across space and it has been captured by various measures into the data set.

So that is what is called as a fields, we looked at how do we represent objects and fields. So objects as point, line and a polygon fields the same point line and polygon in a field representation is represented as number of pixels okay, if it is a point it may be a single or many pixels surrounded whereas a line is a collection of pixels okay, in a particular way. And if it is a polygon it is a huge area or it is area earned the collection of pixels okay.

So, we also looked the data model so with me what you mean by a data model. These are tools that are representing the real world which convert the real world phenomena into a database and helps us enquiring. In the next session we would we shall discuss about the vector data model first and different types of vector data model that can be represented. So, till then thank you very much, have a nice day.