Geographic Information Systems
Prof. Bharath Aithal
Ranbir and Chitra Gupta School of Infrastructure Design and Management
Indian Institute of Technology – Kharagpur

Module No # 01
Lecture No # 04
Introduction (Continued)

(Refer Slide Time: 00:28)

CONCEPTS COVERED

➢ Spatial data acquisition

➢ Spatial data source and generation

➢ GIS attribute data

➢ Data display and investigation

Hello Namaste today we would look at the lecture 4 of the module 1 where we would be looking at concepts of how acquire a data. This is very important part I would be introducing certain data sources where you can get data so we have sufficient data when we look at Indian subcontinent every data is available for the users only thing is how we access it is the thing we have to understand.

So the first thing that I would speak about is how we acquire this data then the second thing is looking at spatial data source and generation are there certain data sources which can be used or can we generate this data's from different sources is the thing that we would look at in the next part of lecture. The third part would be GIS as an attribute data so how attribute data is used as a GIS capability is what we will understand and finally we would look at how will display the data how the integration of the data will be happening and with this look at the entire capability of GIS.

So with this lectures I am trying to show you how the data generation happens to the analysis part. So this is just a introduction part of how GIS work but if you can understand how the spatial data is acquired to and the data display and integration it would be extremely helpful in understanding this as a basic part.

(Refer Slide Time: 01:56)

So when we look at data acquisition is normally is the first step of whoever starts looking at GIS whether it is GIS whether it is any other part of research first is that everyone as to be understand is how do get data okay. Many research scholars who look at the first thing they give extremely give research questions but they would not know the sources of data. The first thing that everyone has to look at is that first look at the source of the data then is what your proposal comes.

The first so when we look at Indian context since 2010 the most of the government agencies at different levels as setup data for public sharing. So most of these data's available and this geographically referenced most states most of the government agencies have put the data in the open domain. Now spatial data acquisition there are different methods of acquisition for example some of the spatial data can be acquired using interviews, the land surviving and the numerical methods.

There are huge ways of looking at different models and methods in order to do the ground based acquisition then laboratory measurements of samples for example you go to the field collect

samples come back to the laboratory analyze those samples then you have measurement by in-situ sensors. This is what is the thing of today so you have sensors that are available at very less cost and even it is designed by many of us who can understand how the senses works.

So the measurement by in-situ sensors are one of those aspects that people have taken out in great way understanding the ground ways acquisition. So you do not need any human resource but you have a sensor that is there on the ground which gives all the time all the data that is required for any kind of analysis.

(Refer Slide Time: 04:03)
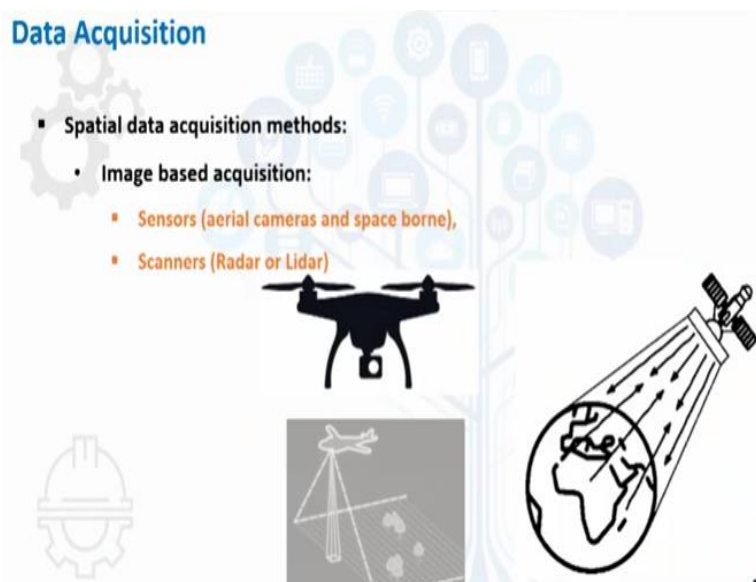
So when we look spatial data acquisition so from the ground to the spatial data we can look at sensors it may be aerial camera's it can be space borne sensors it may be satellites it can be scanners something like Radar and Lidar and this is what is revolutionizing the way that is acquired. Now the data because of new regulations government of India is now pushing into so that the data the geographic data can be collected using the UAV's or the drones so that has made it more easily you know easily available data for the users.

(Refer Slide Time: 04:48)

Data Source and Generation (1/6)

- Survey of India portal
- Free topographic maps of scale 1:50000
- http://soinakshe.uk.gov.in/

So that is how the data acquisition starts and it can be the data can be acquired from various portals. For example if someone wants to get survey of India topo sheets now government has made it the all most of the toposheets a digital. So anyone who wants to access this toposheets can go into survey of India portal which is mentioned which is soinakshe.uk.gov.in so if you can go into this particular website you will be welcomed by this particular prompt.

So once you are welcomed by this prompt you have to just log in where you have to provide your UID or Aadhar number. If the aadhaar number is added here and you have to enter the captcha code. So once you enter the captcha code you get an OTP on your registered mobile number which the Aadhaar so that OTP is entered here you will be able to access the map that you would need only how to access this maps what kind of maps is required what is the scale of this maps etc., that will understand in that alter stage.

But now it is one of the data sources if you want to get certain data of a particular place that can be got one of the sources is nothing but survey of India portal it is extremely good lot of information is put on the website. So if someone is interested please look at it I would suggest everyone looking at it.

(Refer Slide Time: 06:15)

Data Source and Generation (2/6)

- United States Geological Survey – Earth explorer portal
- Free satellite data of resolution up to 10m (Landsat, Sentinel, IRS-LISS etc.)
- http://earthexplorer.usgs.gov

The next data source if you are if someone is looking at remotely sense data there are huge number of free resources available but very well-known medium resolutions satellite is nothing but Landsat. So this Landsat data can be easily downloaded using earth explorer. At the explorer also gives you Indian satellites data now certain data of LISS3, LISS 4 senses are actually given by the USGS.
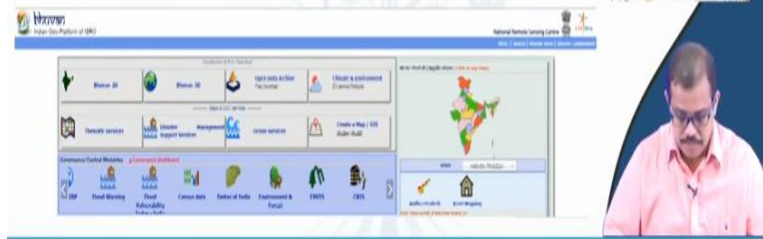
So this is the website there you can download some of the satellite data just go and give the address name and place and what date to what date so it gives you the geographical information when you give the name of the place and once you give what date to what date will ask you what are the datasets you need. Whether you need a remotely sensed data whether you need certain attributes data etc., all every data set is listed here.

So once you click on this submit you can find out there are large number of datasets available from almost from 1980's to till today till date. And you can also see what is the quality of the data if you are okay with the quality of the data then it can be easily downloaded in free of cost only thing is you have create your own account into this portal and then you can you are free to download.

(Refer Slide Time: 07:39)

The third and the most referred aspect is Bhuvan it is India's own website that actually showcases how geospatial data the geospatial information has progress over past 2, 3 decades in Indian context. So when we look at Bhuvan is a portal which also gives you a lot of data for example you have a Bhuvan 2D data you have a Bhuvan 3D data is something like your Google earth then you have open data archive this has a huge sets of data that is available in Indian context.

Then you have a climate and environmental data this are (()) (08:27) derived product then you have thematic services so you can with certain analytics you can even look at the different thematic structures. For example if I am trying to look at what is the literacy data across India if you go here this thematic services will give you the entire literacy rate where the literacy is high where the literacy is low where the literacy is reaching at good levels.

So looking at that one can understand which state as to be constant I mean have to be given more of concentration in order to develop it so that is what with thematic structures then we will also look at thematic services when we are looking at the software also. Then disaster management support system this is what I was speaking about Uttarakand previous class then you have ocean services and most important is we can create a map in using this particular portal it means to say you can download probably the boundaries of particular region you can create your own maps you can get subsets of that data.

So of these can be done so and second this that also information services which means in terms of flood warning, flood vulnerability senses data is made available Bhuvan which is the new thing that I have seen the way it has been represented the Delta's of India the environment and forest in India. So every information from ministry is being put into the single portal so these are various things that find it on Bhuvan and most importantly is that you get a remote sensing satellite data of Indian remote sensing satellite IRS-LISS 3 and also if you want to understand the elevation of region you get CartoDEM.

So this resolution is about 23.5 meters so you can download this data from 2008 or 9 till 2019 the data is available free of cost so anyone who wants to work on remote sensing data and GIS data there is a one stop portal where most of the things of Indian subcontinent can be easily found. So you also have state remote sensing portals every state as its own remote sensing centers which are otherwise called remote sensing centers you can go those portals you can download a lot of information's that is available many of those have an interactive website when the data is available.

(Refer Slide Time: 11:12)
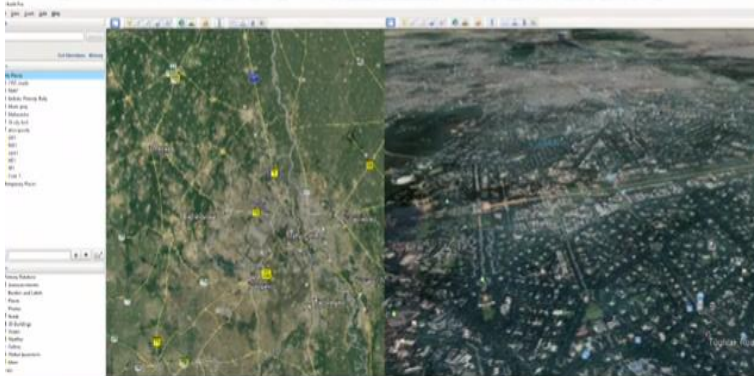
Then the biggest servers of the data generation for any of the analysis is OSM the open state maps. So one thing is you have free data in terms of point, line and polygon so if you go this particular website you can find out what open state maps are they are extremely interactive will give you huge amount of information many of your maps, many of your data outputs are based on this open street maps.

So another big data source is Google earth so I hope many of you either used seen this maybe in a various ways it may be in your movies it may be in other ratios you see Google earth as one of the major aspects in order to pointing out the geographic place. So it has you can create free vector data I will say what is the vector data. So as of now we can create free point lines and polygons point is represent particular maybe a building line is to represent a road segment may be polygon represent a area. So you can create it by yourself you can look at 3D building they look at terrain visualization and more so all of this capabilities are inbuilt with Google earth.

(Refer Slide Time: 12:35)

And very recent development that actually speaking the research community by real storm is Google earth engine extremely good this has it is own data sets is already available the entire course of various algorithms are already available you can use python to code it and this helps this is a platform which gives you without even having the system it can give you the entire output (()) (13:10)etc., can be done in a much easier way without having a capable system with you.

So Google earth engine is extremely capable and has huge amount of data that has been stored. So analysis data combine is what is GIS so you have everything you have data you have analysis capability and information can be put out for example there is one of the information that was put out by Google earth engine where it is providing a climate model from a Google earth engine one of the models the data that is a dynamic nature it is exhibiting how that climate is changing over a period of time.

(Refer Slide Time: 13:50)

So this is about data generations or acquisition so very important part of data acquisition where people miss is meta data the first and foremost important thing as to be seen when you acquiring a data is a metadata. Meta data provides answers to for example Who? Where? What? Now How? When? And Why? So when I say who it is about who as actually created the database who manages the data.

For example if there is an updation every 5 years who is the person who is doing the data because we should know the legitimate quality of that data if the quality of the data is bad then downloading of that is no use. So that tells so that metadata gives you information of who is doing it. So if it really a good source may be a government source or may be any other source which is extremely good so it may be useful

Where it is means where does the particular study area is for example if you downloaded satellite data it tells your where that location is giving you a coordinate it can I access the data is it source data so how what is the kind of source from where it has been acquired? So it is not that is passed on from once source to other what kind to data it is what is the data content what are the data what is actually it is providing as a data.

Then what are the source to generate that data is been used whether it as collected over GPS whether it is collected over sensors what day time night time or what part of the time is it morning, evening, afternoon? So all of this (()) (15:34) it is all of these will be put in to the meta data then you have how the data is created what are the procedures that has been used to create this data then is the period of the content I mean what is the period of the content and when it has been created.

Then what is the data created are there any missing value? So all of these will be in the metadata so whenever anyone is looking at any data source which is not generated by you the first thing is look at a meta data. Meta data is data about data so if you not understood what is a metadata so if you are seen the meta data then there is every chance of your information going wrong when you are extracting information from a data.

(Refer Slide Time: 16:24)

Data Acquisition – Metadata

- Core components of metadata involves:
    1. Metadata Record Information
    2. Identification Information
    3. Constraints Information
    4. Data Quality Information
    5. Maintenance Information
    6. Spatial Information
    7. Reference System Information
    8. Content Information
    9. Symbology Information

So when we look at metadata the very core component is you have metadata record information you have identification in information you have constraints. So this is very important parts so constraints will give you what is actually the thing that information that can be extracted from your data. Data quality information so how good is the data if it is have the cloud if it does not have the clouds if it has certain errors if it does not have errors is has scan lines it does not have scan lines.

So all of these information will be put into it you have maintenance information the very important as per the spatial information where it is located. Then the reference system what kind of reference system because when you are analyzing data everything should be in the same reference system this is very extremely very important then for example if you have one data in one reference system and other data in other reference system so when you overlay both of these one will sit may be in America and other one will sit in may be in somewhere in India.

So if this both are not at the same reference system so this are digitally divided so it has to be in the same reference system. Then you have content information that is important in terms of understanding what is a content. Then the final thing is symbology what symbology is used in that data because every data as its own way of representation.

(Refer Slide Time: 17:55)

Data Acquisition - Metadata

- When there is sharing of data publicly metadata acquisition is very important
- Metadata stores all information about layer/image/data
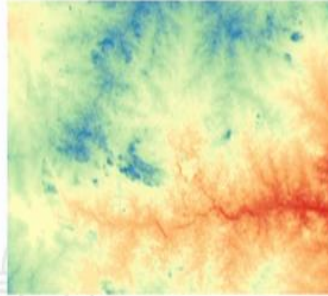
Driver	
GDAL provider	
Gtiff/GeoTIFF	
Dataset Description	
F:\Slope\elevation.tif	
LAYER_TYPE=thematic	
Dimensions	
X: 7200 Y: 7200 Bands: 1	
Origin	
608849,1.32835e+06	
Pixel Size	
30,30	
Data Type	
Int16 - Sixteen bit signed integer	
Pyramid overviews	
Layer Spatial Reference System	
proj=utm +zone=43 +datum=WGS84 +units=m	
Layer Extent (layer original source projection)	
608849.16,1105562.93,828944.28,1328348.47	

And when there is a sharing of publicly the metadata acquisition is extremely important especially the public data whatever you are downloading. Metadata stores almost all information about layer, data or image for example on your right hand side I have given a list of metadata for example here GeoTIFF is a format that is provided by GDAL. So this is the driver that is actually giving you there is a format that is provided by GDAL. So this is the driver that is actually giving you this is the format of that this image then the dataset description it is an elevation so it is elevation.tis so this is where it is stored okay.

What is the dimension of data it is 7200 column rows by 7220 columns and it is 1 band data so origin is this then you have a pixel size each pixel is 30 by 30 meter data type is integer 16 bit data 16 bit integer type so a similarly you have a projection is utm zone is 43 datum is WGS84 and units is meters. So this is what has to be similar for all the layers that you are trying to analyze. Then what is the layer extent so it gives you the original layer all and this all the coordinate starts in particular section system. So this is about the metadata.
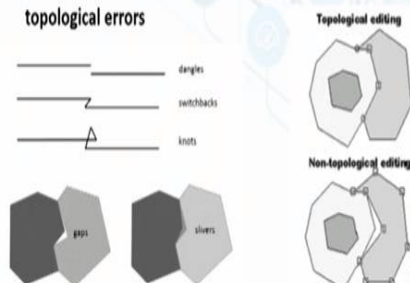
(Refer Slide Time: 19:24)

So as I did say data can also be digitizing orthophotos are extremely capable ways of looking at data. So these are created from satellite images or maybe even converted from GPS data, survey data with specific coordinate information so orthophotos also a good source of deriving a data layers.

(Refer Slide Time: 19:49)

Then data acquisition may involve both new and old data in terms that these both data have to be unless for analysis of a certain GIS based operation must necessary match these information's. For example the geographic information the coordinate information etc., Additionally both existing and use spatial data must be edited if they contain digitizing errors this is the first step in analyzing any spatial data.
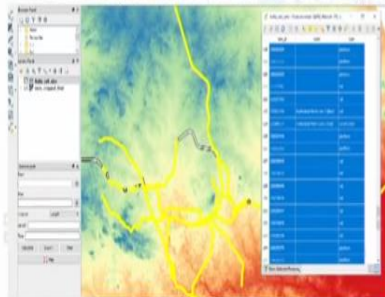
If you have any errors topological or a digitizing errors has to be removed so we will also look at some of the errors for example when you look at topological errors you have something like dangle you miss this particular line. So you have a switch way when you actually drawing a line so somewhere you have a switch back somewhere you have an error when you are actually creating because you are actually doing on using a mouse or may be digitizer.

So such things have to be removed or you have something like a silver which has a over shadowed polygon or you have something like a gap which is under shadowed polygon then you have a gap between a polygons. So something like this as be to be corrected so I have shown here how the topological editing is done will go to in details of it in a coming lecture this is non-topological editing.

(Refer Slide Time: 21:16)

So the next set is attribute data in GIS as a database management system that can handle attribute data but these attribute date are most likely stored in the relational databases. And when we look at this tables can be prepared maintained edited separately all of this tables can be edited separately and only thing is that you have linkages. So you can look at mathematical you can look at statistical and database related operations that can be performed on all of these tables in any of those attribute data in a GIS.

(Refer Slide Time: 21:51)

Data Display

- As an representative factor for all mapmaking plots are the final interface to GIS
- Data display is an integral part of GIS because symbol, size, color, shape, and pattern affect the user's perception and the map's overall meaning
- It involves:
  - Selecting the symbol to represent features
  - Symbolizing a layer based on attribute values
  - Using a classification method to represent numeric values
- Maps have certain formal methods of representation.

These would be dealt in 7th/8th week of the course

So the final part is looking at GIS as a display so this is you have to look at represented the factor for any of the map making parts or final interfaces of GIS. So how do you represent a particular map that is essentially important we will look at every part of a map also we will understand how maps are done what are the different parts of a map how maps are organized etc., It involves selecting the symbols to represent features symbolizing the layer based attribute values using different classification methods. It can be any kind of classification methods we would look at it when we are actually looking at how different map features are there in may be in seventh or eighth week of this course.

(Refer Slide Time: 22:35)

Summary

- Data acquisition – Ground based and image based techniques
- Data source and generation – SOI Nakshe, USGS Earth Explorer, Bhuvan, OSM, Google Earth and Google Earth Engine
- Metadata and characteristics
- Attribute data and relational databases
- Data display and mapping

So in order to summarize we have data acquisition the first thing we look ground waste and image waste techniques. So I have you good number of links where the data is available I would suggest everyone to look at each and every link because this has good data source. If you become familiar with the data sources it would be very easy for you to understand different aspects as you move a head in this course.

So the first is to look at some data so once you have this course can only be learnt the data and more practice so understanding data and having more practice is extremely important is to be more successful. So when we will look at data source and generation we looked at SOI Nakshe which is really good as far as the governmental data set is concerned it gives you everything based on maps.

Then you have USGS earth explorer which is well-known land set data set provider along with various other data even I did inform you about the Indian satellites the data that is available then you have Bhuvan. Bhuvan extreme capability of how we look at data how different datasets are there what kind of data is available and you can utilize Bhuvan itself as a platform in generating a data.

Then we looked at open street maps and it is features and Google earth and Google earth engine so if time permits I would request everyone to look at Google earth engine as a and it can be extremely innovative and it can help you out in your carrier in a better way. So look at have you looked at what is a meta data what are the characteristic of a meta data then we did we even look at how attributes what is the attribute data and how they are linked it need not be just in form of a columns in a single sheet but this data can be over a multiple data sheets and it can be linked over by different aspects.

Then you have a data display and mapping system so this data display and mapping system is where the last part of where you add in what are the map characteristics is required what are the different the labels are required different map place makers are needed in order to define that has a final map. So with this we come to end of today's lecture so we will look at more aspect of how the data is can be used in the next class so thank you till then.