Geographic Information Systems
Prof. Bharath H Aithal
Ranbir and Chitra Gupta School of Infrastruture Design and Management
Indian Institute of Technology-Kharagpur

Lecture - 62
National Spatial Data Infrastructure (NSDI)

Hello, namaste. Welcome back to the course on Geographic Information System. As I said in the previous class we would discuss about the National Spatial Data Infrastructure or well known as an NSDI. So most of you have would have heard about this government is aggressively looking at how we standardize data that is actually generated from the Indian subcontinent.

So in this today's class, we would look at what is the vision basically vision of the NSDI.

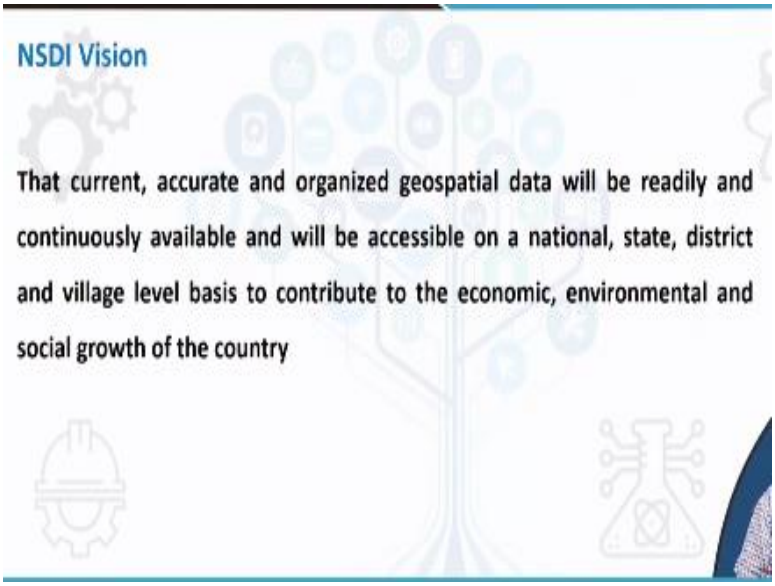
(Refer Slide Time: 00:52)

Then we will look at how the National Data Sharing Policy has been created. There is already existing National Data Sharing and Accessibility Policy. How and how does that affect as a researcher and academic we look at. Then we have a need and objective of and the scope of this policy which is as of now very broad. So let us look at that aspect in today's class.

Then we will look at the benefits of national policy, why do we need it basically? So that is very important in terms of understanding the policy by default. Then you have

type of data access. What kind of data can be accessed and how it can be accessed? Then what is there a pricing or a legal framework that has to be worked out? That is also we can look at as we progress.

(Refer Slide Time: 01:37)



So when we are looking at NSDI, so NSDI vision is as simple as to make the data accessible, okay. The first thing is create current accurate, organized geospatial data that can be readily and continuously available, and will be accessible on a national, state, district, and village level basis. So each of this data, whether it is national data, whether it is state data, whether it is district data or a village level data, they have different way of representation for sure.

But this data has to be organized in a particular way, which the organization should be understood by the users and most importantly, this organization and the data representation should be accurate and more importantly should be current. So you so whatever the data that was normally available, previously where before the NSDI came into for maybe 10 years or 12 years old data, which is quite unorganized in terms of the representation.

So if it has to be organized, then it needs to be clearly distinctly represented. So that way of representation is the vision of NSDI. So why it is actually the vision this NSDI vision is created is to be basically contribute to economic, environmental and social growth of the country. So that is why and most of the NSDI has its stake on the OGC vision.

So when you look at the outlook of NSDI there, it has a lot of things that has been very influenced by the OGC standards that are already present.

(Refer Slide Time: 03:22)

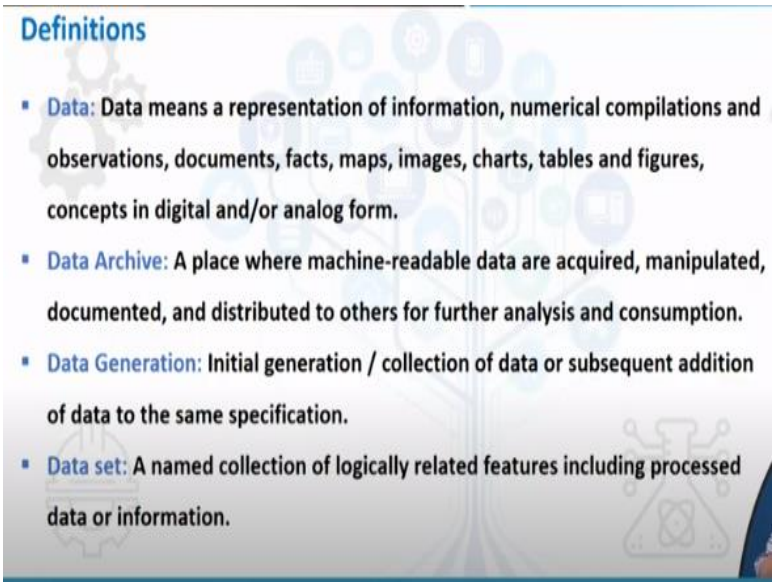
And when you look at the national data sharing and accessibility policy that came up in 2012, it is basically designed to apply and shareable non sensitive. So it is very specific, it is a non-sensitive data. When I say sensitive data these are certain data which cannot be shared in the public. So non-sensitive data available either in digital or analog forms. And basically, it clearly states these are generated by the public funds.

When I say public funds any academic institution, any research institution, any of the governmental organization which are in the ambit of receiving public funds for their regular work. So they are supposed to generate the data all the data either digitally or analogaly and this has to be designed in a specific way as prescribed by the National Data Sharing and Accessibility Policy.

So whether it is funded by ministries, departments, subordinate officers, organizations or agencies of government of India, this policy has to be taken into consideration. And when you look at NDSAP policy that is the data sharing or accessibility policy is designed to promote data sharing and enable access of the Government of India owned data. So whenever you are creating data from public funds, it is basically not any personal data.

It is the data that is owned by public and owned by the Government of India. So that to have an access of that data, okay for National Planning and Development. It can be national, regional or even the state level planning for any kind of planning and developmental activities. So this policy is very clear cut in terms of defining how the data is created, how where the data source is from, what is the ambit of that particular data and why this particular data will be used and implemented.

(Refer Slide Time: 05:24)

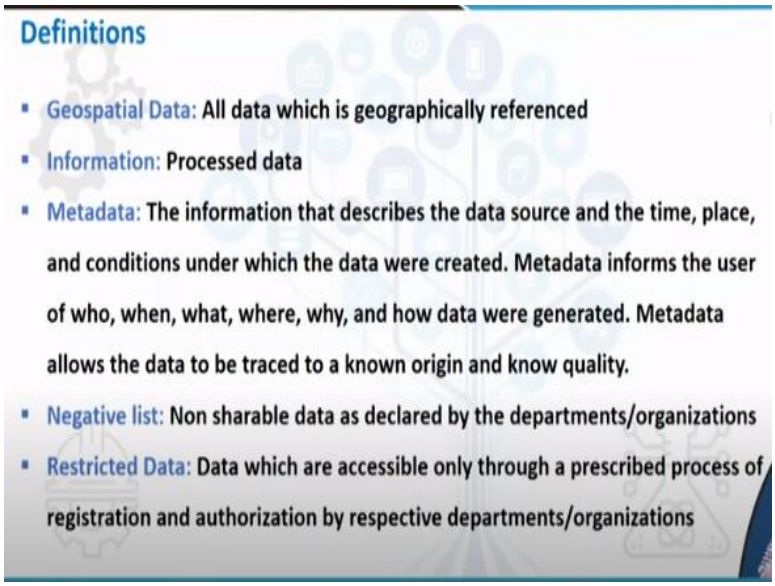


So when we are looking at this policy, there are certain definitions that we have to look at. The first definition is about data itself. Data when I say data means any representation in the of information, numerical I mean compilations and observations, documents, facts, maps, images, charts, tables and figures, concepts in digital or in analog form. So all of these are data, anything that gives you information in any form is data, okay.

So that is comes under the definition of data. And when I say data archive, so this is a place where machine readable data are acquired, manipulated, documented, distributed to others for any further analysis or in terms of representation or consumption. So this is where it is called as a data archive. So all of these are very defined in a clear cut way so that a user or a data generator understands how the data has to be stored and transmitted.

So when I say data generation, initial generation and collection of data or subsequent addition of data to a specific, to the same specification. That is what is called data generation. Then you have data set. It is a named collection of logically related features, including process data or information. That is called a data set basically, okay.

(Refer Slide Time: 06:55)



So when you look at geospatial data, so when I say geospatial data all data that is geographically referenced, okay. So anything that has a reference to the earth's surface and anything that can be easily located is nothing but a geographical data is also called as geospatial data. So this information, whatever the information that we get from the data, the processing of this information or the processed data is nothing but information.

Then you have metadata. I spoke about metadata in my previous lecture. So it actually defines it in the same way. It is the information that describes the data source. So I am very specific here. So it is data source, time, place. It is extremely important because when you are looking at temporal changes, you need to account for time and place and conditions on which the data were created.

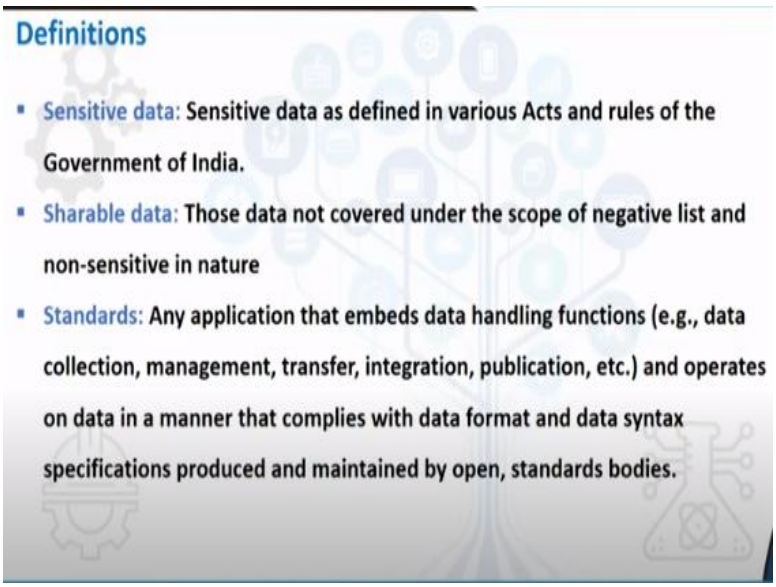
So you cannot have a few, let us say five sets of data, three sets of data taken on a summer day and three sets of data taken on a rainy day to measure the temperature and say that this is a temperature variations across five different days. So that is not possible. So it has to be in a specific time, place and conditions has to be very clearly

mentioned. Metadata informs also the user of who, when, what, where, why and how the data is actually created.

So and metadata also allows the user to be traced to the known version and known quality. So this also is the same OGC standard that comes out here and NSDI. And when you look at negative list, non-shareable data is declared by the departments and organization beforehand. There are I mean, most of the data is shareable, but data on a very finer scale is kind of non-shareable and it is already declared by departments.

And when you say restricted data, these are the data that is which are accessible only through the prescribed process of registration, authorization by the respective departments or organization and comes under the ambit of any of the data policy that each of the organizations have, which may be shared, may not be shared. It is dependent on the organization which is actually creating the data.

(Refer Slide Time: 09:12)



So there are also set of sensitive data. When I say sensitive data these are defined in the various acts and rules of government of India which cannot be shared in the public. So these sensitive data are also defined very clearly in terms of the national spatial data sharing policy. Then you have a sharing data which means, those data are not covered under a scope of negative list are and non-sensitive in nature.

So these are shareable data. So you can share this data with anyone but look at the standardization okay. Then you have standards, as in OGC standards, you have this as

an any application that embeds this data handling function should have certain way of standardization that is already represented by any of the open standards bodies. It may be OGC itself. Many of them have follow OGC standards itself.

So should have the data should have a particular syntax specification produced and maintained by open standard bodies is what it defines.

(Refer Slide Time: 10:21)

Then you have a need for policy. That is evidence based policy of socio-economic development, okay. Then it also says that it has to be shared and utilize the data generated by the government departments itself. Then efficient sharing of data among data owners and inter and intra governmental agencies is very easy in terms of when you have such policy. That is why the policy was proposed.

(Refer Slide Time: 10:48)

Then, when you look at the objective of this policy, it is to facilitate the access to the Government of India own shareable data and information in both human and machine readable formats.

So there has to be an access of human readable format, there has to be an access of machine readable format through the network all over the country in proactive and a periodical updatable manner, within the framework of various related policies, acts and rules of the Government of India, thereby permitting a wider accessibility and use of public data and information. So that is what is the objective of this policy.

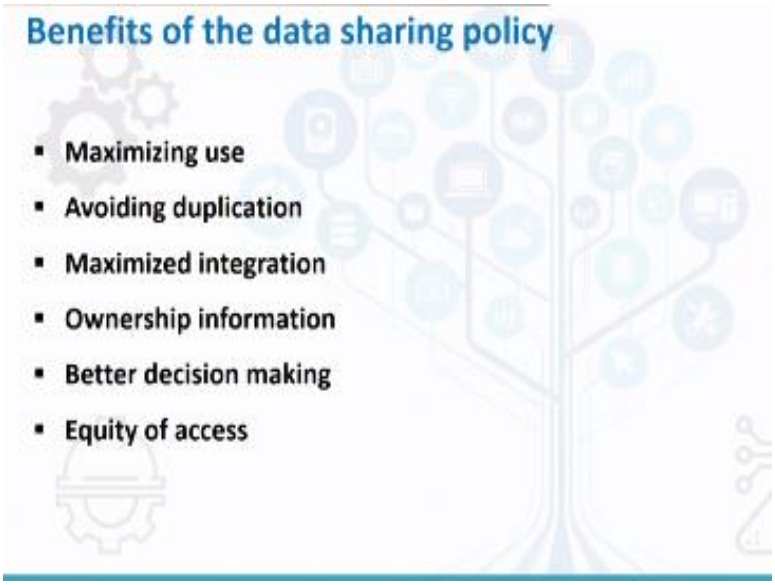
(Refer Slide Time: 11:25)

And when we look at the scope of the policy, the national data sharing and accessibility policy will apply to all every information that is created, generated,

collected, archived using public funds provided by the Government of India directly or through an authorized agency okay or state government or state government agencies, by various ministries, departments, organization agencies, autonomous bodies.

So it all of these are under the scope of this particular policy. Then certain data that I did mention that other than those data all data has to be shared under this particular policy.

(Refer Slide Time: 12:10)



And when you look at the benefits of the data sharing policy, the it stresses on maximizing the use, how do you help in maximizing the use of this data and avoid duplication. So I have created a land use record. The same if the record is not erroneous is quite has a good quality data. So the same thing need not be repeated by someone else to create the same data.

So that is the main thing of why this data sharing policy came into force. So this avoids complete duplication. Then maximized integration of data okay. Then ownership information that will be created along with data. Because once the data is out of your box, it may be manipulated the way it is required to. So ownership embedding ownership information entails that owner can represent whatever the basic data that he or she has created and helps in defining how the data has been created.

And most importantly is for decision making. So that is the policy is towards developing most important decision making much easier. Then equity of access. Equity of access is never shared or never there in the spatial data. So that this particular policy is towards taking towards the equity of access. So that is why you need such policies, in terms of even with whatever the data sharing we have.

(Refer Slide Time: 13:34)

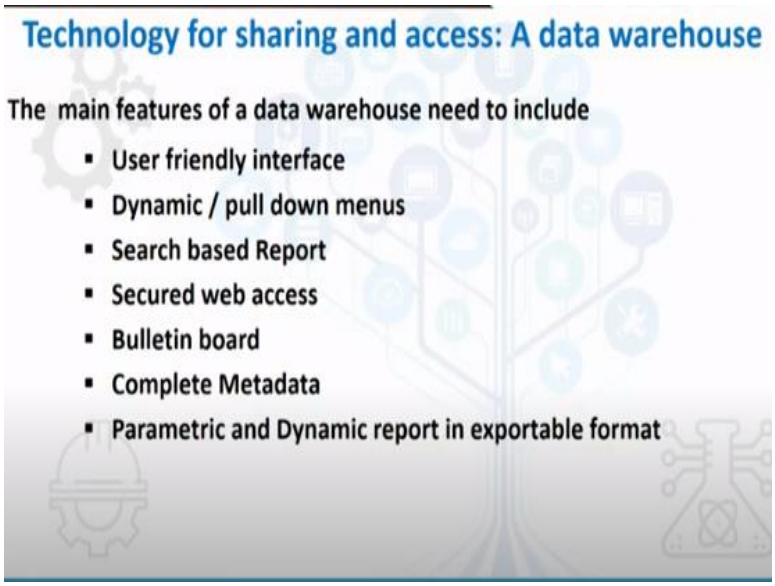
And when we look at the policy it also defines very clearly that the type of access that is there, for example, open access. When I say open access, the access to data generated by public funding should be easy, timely and user friendly and web based without any process of registration and authorization. So any data that is generated by public funding should be available to the public without any authorization or registration.

So this has to be now it has to be web based. Before this the first when the beta policy was proposed, the web based was there, but now it defines very clearly it has to be web based okay. Then registered access that is the data sets which are accessible only through a prescribed process of registration authorization by respective departments organization will be available to the recognized institutions, organization, public users only through a defined process.

So we cannot cut across that process. Then, there is a restricted access. So data declared as restricted by government of India by its policies, by its activities. This will not be accessible through any of the authorization or to any organization other than

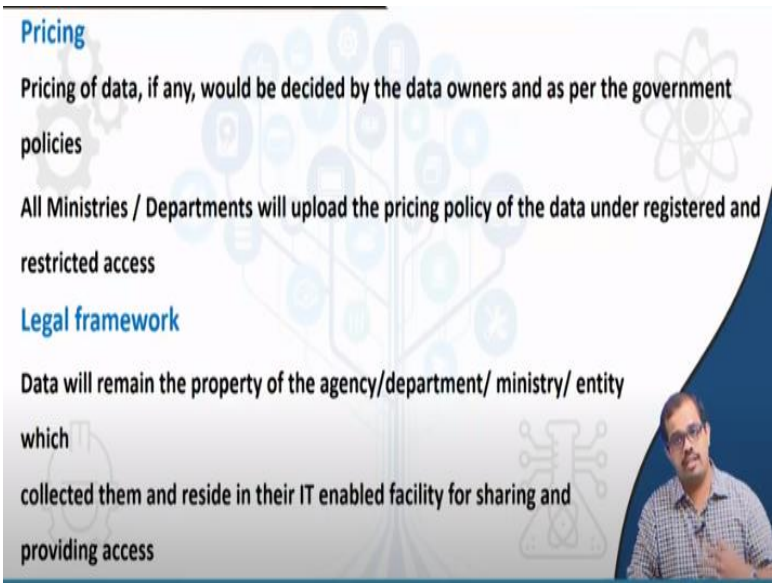
for the decision policies of Government of India or the particular organization which has generated that particular data.

(Refer Slide Time: 15:12)



So that is about the access. Then we have technology for sharing and access. So you should, which is actually a data warehouse. So when you look at the main features of data warehouse, it needs to include a user friendly interface, dynamic pull down menus, search based reports, secured web access, bulletin board, complete metadata, parametric and dynamic report in exportable format. So this is how the sharing or accessing of the data usually happens.

(Refer Slide Time: 15:44)



And when you look at the last part is very important is pricing, okay. Whenever you say pricing, pricing of the data, if any would be decided by the data owners as per the

government policies only. So we any anyone cannot have their own pricing for the data that is generated from the public funds. This has to be in concern with the Government of India policies, the sharing policy and the pricing policy.

And all ministries and departments will upload the pricing policy of data under registered and restricted access. So every ministry has its own policy and how the data is shared, whether it is registered data or it is restricted data or a public funded data. So based on these three categories or public funded or open data based these three categories the ministries or departments will normally give out the data.

And when you look at the legal framework, whenever you look at the legal framework, data is always a property of an agency, department, or a ministry or an entity which collected them and reside in their IT enabled facility. Which means to say that I have collected the data, if it is in my IT enabled facility for sharing and providing access, it is my property the data.

Only the access to that data is given by me or the sharing of the data is given by me. So I am the original owner, but it can be shared or used by other access based on different restrictions that we have, okay? If it is a restricted data, it cannot be shared and used. Only the user who has got the permission will use it, okay.

(Refer Slide Time: 17:24)

So this is about the thing. And when you look at the policy, it is actually to develop an institutional repository, data centers on a domain and a national level, which means to

say, whether it is state whether it is national, or a village, so it is to develop a particular institutional repository, so that it will be easy for us to look at any kind of data sharing policy.

Or it may be looking at the decision making or developing a spatial data infrastructure in order to derive additional capability that can be done through this policy. And when you look at National Data Sharing Policy and Access Policy, it is basically to promote the technology based on the culture of data management, as well as data sharing. Before this policy, there was no culture of data sharing at all.

All data, even with the public funds were thought to be the personal data. So now and with this data sharing policy, any data that is generated by public funders should be represented and should be put out to the public for them to analyze and interpret. So that is about this policy. This is what I wanted to speak about the NSDI and the National Data Sharing Policy.

(Refer Slide Time: 18:41)

So we have looked at all of these, whether it is a need, objective, and scope of the policy. So we defined very clearly what are the different types of data that we have considered, how the data that will be considered, what are the data that is available, and what is the pricing. So as I said pricing cannot be defined by us, but defined by the Government of India's price book.

And finally, the legal framework is we are the legal I mean owners of that particular data. Whoever has generated the data is the legal owner of that particular data. So this is about the entire NSDI and the National Data Sharing Policy. It is yet in a very broad representation. It is actually growing and now you can see the newer versions coming up in some days. So this is about the entire course.

So and probably, this Mr. Chandan would speak about geo server, which is as I spoke about the web based service. Mr. Chandan would give you a detailed overview of how a web based service can we built, how the data is being used, and how do you represent it as an as a web service. So if that is an useful output in terms of whenever you are representing, standardizing the data and putting it on the web, this is very essential to understand.

So maybe that is the last part of this particular course. So I would wish all the best for everyone for their exams and for their career. Thank you very much.