Digital And The Everyday: From Codes To Cloud
Prof. V Sridhar
Department of Multidisciplinary
International Institute of Information Technology, Bangalore

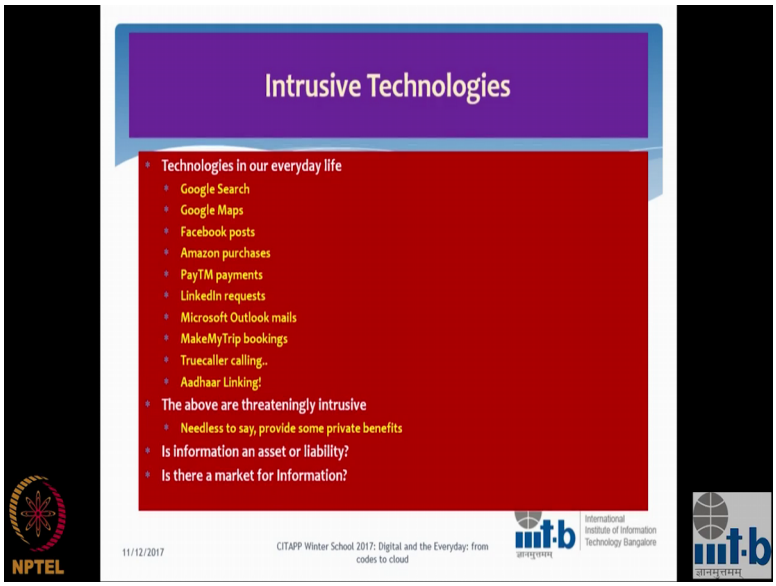
Lecture – 06
Data Protection & Privacy Regulation in the Digital Era Part 01

I am Sridhar so, I basically work on telecom regulation policy, but as I have told you before. So, I have started looking at data protection and privacy issues as it is closely related to telecom. So, this particular session you will try to give you some background about data protection privacy; you know how to understand the issues surrounding privacy.

How does it affect you know for example, the privacy of individuals, firms, organizations and what are the regulatory mechanisms that should be in place and we will also take some examples from the existing acts and policies and have what they cover I mean that is basically the succinct of a you know one and half hour lecture. So, if you have any questions please stop me at any point of time.

So, the context is that suddenly data protection and privacy has become a big issue right.

(Refer Slide Time: 01:04)



I mean we are known for giving private information at ease with who neighbors and whenever we go in the train, we are very happy to divulge about our private information.

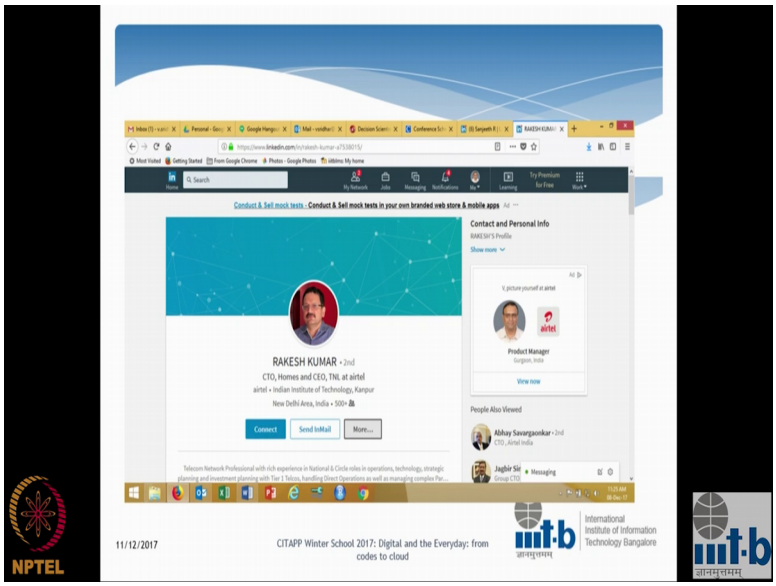
But suddenly it has become a big issue right and supreme court is looking at it and this is the most important national law school of India University is here is here right.

So, why is it becoming a big important issue so, we need to understand that in everyday life we use technologies right. I mean that is the whole theme of this particular winter school and so, we deal with Facebook, we deal with linked in. We use mobile we use Wi-Fi, hotspots for connecting to the internet, we do Google search right and we buy things at Amazon.

We book cabs through Ola cab, all these information suddenly is digital and therefore, it is possible to collect all these information and then use it for whatever purpose it is intended for or non intended for right. And therefore, suddenly after the digitization has come in place the privacy and data protection has become a very really very important issue because its surrounds our everyday life.

So, I am being surveilled right so, everywhere you are being surveilled and that information can be used for any purposes. And therefore, it is becoming very important in today's context, and these technologies or intrinsically intrusion in nature right. Even if you do not know it is being intruded upon your information is being intruded upon at any point of time right so, it is very intrusive in nature.

(Refer Slide Time: 02:52)



So, for example I was just looking for a speaker for my mobile even India even which I normally conduct in the first week of January and I was looking for Rakesh at the CEO. I mean who is in charge of networks homes at Airtel and suddenly linked in gives me a job right so, you can take this particular job in Airtel. So, you can see this profile I have been matched with the job in Airtel right.

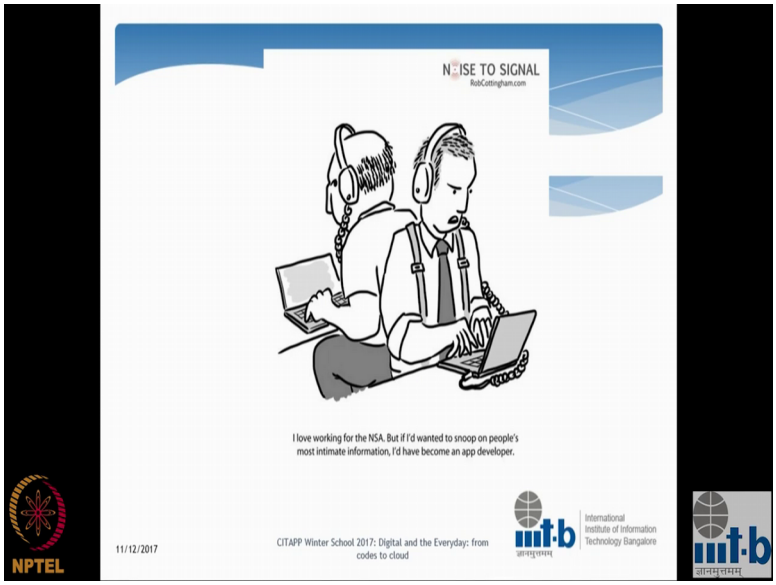
Student: Yes.

So, because of my linked profile and the search that I did this information is being used for projecting something which I may or may not be interested in right so, this is how information is being collated organized.

Student: Analyzed.

And analyzed and then it is posted for your use consumption so, it is very intrusive in nature.

(Refer Slide Time: 03:38)



So, take a look at this cartoon right so, for example, if ever you want to snoop on someone's life you better be an app developer. If you are an app developer you can get all the information that you want on your app right. Where this guy has gone? What this what you know pages that is visited? Which places he has gone? Right, all this information is available to you; if you just develop a mobile app and put it on that particular person's mobile phone right. So, it is very easy not only that the information is

become digital, but it is become easier for us to snoop right its easier for us to collect information.

Student: Yes.

So, hence the context right, but the important thing is that we still have tremendous confusion about privacy. Nobody knows exactly what privacy means if you ask. So, for example, sir what is privacy mean to you? What is privacy meant to you?

Student: Privacy depends on (Refer Time: 04:35) subject or definition.

Correct.

Student: It is something.

That is the most important point yeah.

Student: Something is not problem for others other people only the thing is which data I would like to save with others.

Ha [FL] so.

Student: It is it is showing the data.

Ok

Student: The data which I do not want to share.

Ha.

Student: if it is being shared it means it is got some problem.

So, what you are saying is I will share some information the information. The information that I do not want to share if it is shared, then it is a violation of privacy. What about you? What is the definition of privacy that you can give?

Student: It is durability to safeguard information my interest things that pertain, but there are for personal nature my ability to save current for right now.

So, if you are not able to safe guard the information that is of interest to you then it means it is a violation of privacy that is what you are saying.

Student: That is of a personal nature yes.

That's of personal nature if you are not able to protect it, but some intrudes on it then it is a violation of privacy. So, what about you? What is the definition of privacy that you can give?

Student: I would say if someone's using my information without my knowledge.

Somebody is using without your knowledge.

Student: Yeah.

You are not so, it is very similar to, but is it different from what he is saying.

Student: No.

Shushravya, what is your definition of privacy?

Student: So, some of that do not should not be used without our knowledge same (Refer Time: 05:51).

Without your knowledge, so, similar to what about you sir?

Student: Same.

Same

Student: Without my.

So, we are congruent.

Student: Without my intention anything information.

Ha so, without your consent if the information is shared then it is a violation of privacy.

Student: Or even collecting of my information from third party.

Good point so, collecting of information from third party without your knowledge or with your knowledge.

Student: Without.

Without your knowledge then it is a violation of privacy is there any other definition of privacy that you can give.

Student: Following the cracking my absence.

Without your knowledge or with your knowledge.

Student: Without our knowledge.

Private information or public information.

Student: (Refer Time: 06:21).

For example you come to you have come to here and you might have posted in your linked in if somebody follows that is that violation of privacy.

Student: Yes.

Is it violation of privacy?

Student: Yeah, I do not think so.

Ha what about you.

Student: It is same right if my privacy is been intruded. The data.

What is that that is intrusion means?

Student: Like my.

So, for example, we are you know NPTEL is recording this is this violation of privacy.

Student: No, they have taken my consent.

Here they have taken so, if they as long as they have taken your consent then it is not a violation of privacy.

Student: Yes.

Ok

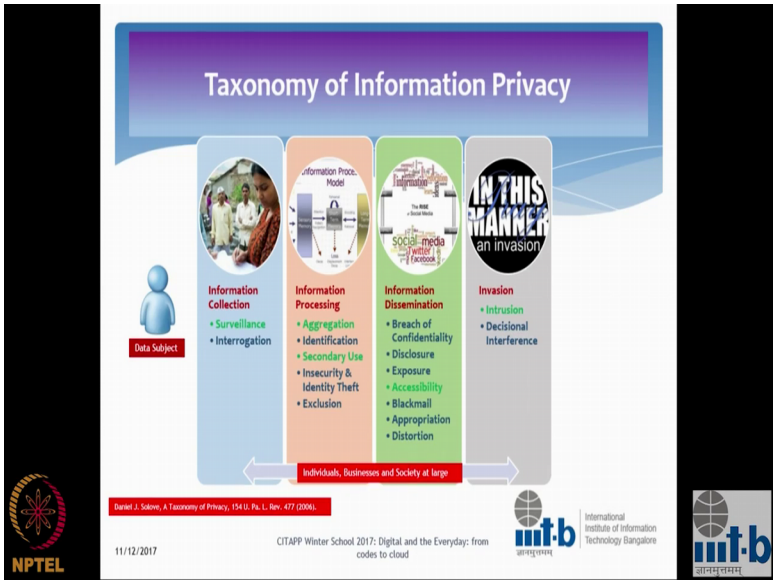
Student: And then it has to be known why or yeah and how it is used why it is used where it is used.

So suppose for example, NPTEL takes this information use it for purposes other education is it violation of privacy.

Student: Yes obvious right (Refer Time: 07:11).

Possible so, I think it is I mean the, I will just skip this one.

(Refer Slide Time: 07:21)



So, it is very important to understand a taxonomy of privacy, otherwise we will just go into loops and then we will not be able to figure out, you know how to punish a particular violation of privacy. For example, if it is if the information is shared without my knowledge. Is it a greater crime? Compared to for example, intruding it to my personal domain, breaking my security and taking that information anyway right. So, there are various degrees of violation of privacy and only if you understand the taxonomy of the privacy. That we will be able to have the clear the understanding right and so, this particular I like this framework it has been formulated in 2007 by this guy

called (Refer Time: 07:54 it is been I mean I have given the references you can look at it you must have (Refer Time: 07:58) review.

So, what this basically does is I will give some examples later on so, information gets collected, processed, disseminated and so, on. At each stage there can be violation of privacy right and it can take different dimensions you know depending upon which part of the cycle that we are in right. So, for example, in the case of information collection it can be surveillance. You know for example, there can be a surveillance camera, which can be mounted in your private home or in a public place, which are retail space, which may be monitoring you all the time and then which may be collecting information right.

Now, we need to see whether it is a public space, private space, it is a intra, it is a covert or overt you know all these dimensions we have to look at it before deciding whether there is any violation or not. Suppose for example, CCTV surveillances is here. All retail stores now have that right. Now, they have publicly posted it right, now it is overt they have already told you this CCTV camera is there visible for you and is a public.

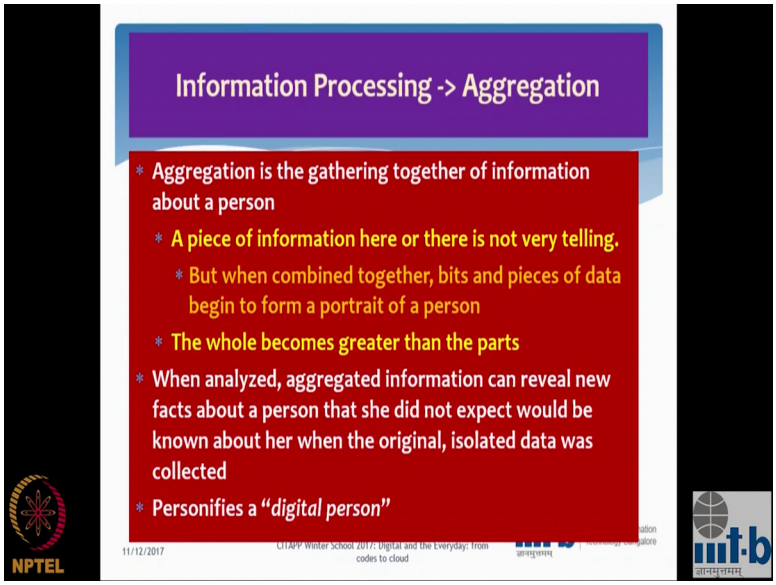
It is a retail store all people come here right and so, this is very different invasion of privacy compared to for example, I mean see you know that it is a overt CCTV camera which is which is surveilling you in a public place right. Will you do something? Which you won't do? Normally will you do that? Will you do so, the something that you have done in your private unsurveilled place, in a public surveilled place? You will not do that because there is someone who is monitoring you know so, you will careful right.

So, we need to look at these dimensions you know before deciding whether there is invasion of privacy and if there is an information privacy, what degree that it is right. So, you have data subject and you go through all these processes and then you know individuals go through this process, business goes through the progress, government goes through this processes and there can be violation at different levels.

So, this is exactly what we are trying to sort of figure out right I mean it is very in one half hours it is very difficult to go through each and everything in detail, but I will just go through some examples and then see how for example, it fits in to the taxonomy. And then it is very easy for us to figure out. Whether it is a violation, how serious is the violation. What is the punishment that we need to give? Is there any act which governs

this particular violation and so, on and so on right. So, the first important aspect which we normally do not associate with for example, privacy is aggregation.

(Refer Slide Time: 10:26)



What is aggregation? See if you tell the Aadhaar number, right Aadhaar number is 12 digit Aadhaar number. Is it a random number? Is it a random number or not?

Student: No.

Or a once you say the number you can say that oh it is professor sridhar (Refer Time: 10:42) and all can you say that.

Student: No.

You cannot say that know which are random number right. So, just I will give you my number nobody will be able to associate anything with that particular Aadhaar number. It is a random randomly generated number, but if you associate that number with name, with address, with mobile number right. With your Google password, Facebook password or Facebook you know login ID and so on and so on. Then what will you get? Will you be able to say that it is professor shridhar (Refer Time: 11:14) ? You will be able to say and this exactly ah.

Student: And what is purchasing on Amazon.

So, you will find, but so, many things about me as long as I divulge you certain other information, which can be collated with my Aadhaar ID right. This process is known as aggregation right so, for example, your Aadhaar number by itself does not cause any privacy problem. If somebody steals my Aadhaar number what can they do with that, they cannot do anything with that right. However, if you aggregate it with other data then it becomes every powerful right, then you I can profile that particular person.

So, aggregation is a very serious offense. In fact, we do not know right at this point of time you did not even say that aggregation is a problem right. So, aggregation is a very serious problem and you can use aggregation in many ways right so, for example, Google does that right, Google knows your personal ID and it knows your browsing pattern and therefore, it is able to collect all these things and then come up with a profile. Oh this guy is interested in telecom therefore, let me throw some advertisements relating to telecom linked in did that right so, this aggregation is it violation of privacy.

Student: Yes.

Is it violation of privacy? How serious is it we do not know the answer right at this point of time you do not have a clear cut answer right. But aggregation is a biggest problem in information technology and it is easy for us for technology companies to aggregate data right and to put some kind of using algorithms that you discussed in the last session, is possible to profile a particular person right. So, for example, this is a classic case right.

(Refer Slide Time: 12:56)

There is something called as Facebook memes which came around 2 years back. They give you memes right 10 concerts they will give you right and you rate them whether you like it or not like it right. And finally, there will be a question which says that, you know tell me the last consort that you went that you really liked ok. Then you will tell the name. They will collect this information right and then they have your Facebook I you know login ID and so, on and so on.

They will be able to using this they were able to profile for example, whether a person who has answered this 10 meme concert question is a male or a female, black or white. Which country he had come from? India or Pakistan or U.S. or Russia right after that they will be able to show advertisements which are relating to the profile of the particular person. Simple 10 questions they aggregate it right and then make a profile out of it and using that profile they will be able to target. Why companies do that? Companies might do it for its own personal purposes right.

They will be able to you know sell advertisements, they will be able to rain a revenue and things like that, but is it a violation of privacy. Is this violation of privacy? 10 meme concert question, they did not ask you anything right, they asked you to login through Facebook right and they had this 10 questions that you answered and there are able to push advertisement, which are they are able to they have been able to you know successfully profile you right. Then your gender, your ethnic background whether you are a Carnatic, classical music lover or film songs you like film songs right. You like Hindi film songs or Tamil films or a Kannada films. They are able to do all those things using that 10 question that you have answered right. Is this violation of privacy?

Student: Yes.

In what degree in a scale of 1 to 10 how would you rate.

Student: 1

This as a serious violation 1 to 10, 1 being low 10 being high.

Student: 1.

1.

Student: 10.

1.

Student: 7 because they are aggregating without my knowledge and also they probably monetizing it right yeah.

Ya, that is the business model right you are able to do Google search free only because of advertisements.

Student: Yeah.

Otherwise each search will cost you like 10 rupees. We will not be able to do Google search know. Google search is free why is it free because they have aggregated information and use advertisement for their monetization model right. Now, you can always ask this question right. Did they ask me before aggregating?

Student: No.

You might have put a tick mark and somewhere in Google, Facebook.

Student: (Refer Time: 15:37).

Long time back.

Student: Long time.

Right that I agree to all the conditions.

Student: (Refer Time: 15:40) most of the terms and conditions they say that you might change the terms and conditions anytime they want.

Anytime they want.

Student: So, even when you sign in.

Student: Like that particular clause may not been there they can insert it later as well.

Right now, so these are so, we are evolving. At any point of timed new business models are emerging around aggregated data right and we know that there is a violation right.

How serious is a violation depending upon this particular context right. So, for example, so this is one of the one of the important problems in n privacy right aggregation.

So, for example, why are people worried about Aadhaar linking suppose for example, I link my Aadhaar card which is a random number right. To for example, my pan card to for example, my mobile number. Is it possible? That using that one identity called Aadhaar number they will be able to collect all these different information to which I have linked and create a profile and say that oh this guy is not paying income tax right or this guy is fooling income tax department or this guy is you know he went to for example, Meghalaya in October of 2017, 23rd of October.

Will they be able to do that? Possible if your Aadhaar number is linked to mobile number, mobile number knows where you are and therefore, they will be able to track you right. Now, the question is it good or bad right. Is it good or bad? Is this aggregation of information good or bad? See everything as good the positive side and the negative side right then there has to be some trade off so, is aggregation good?

Student: 10.

Suppose for example, in the 10 meme concert question I get an advertisement and I get a coupon right which gives me 250 rupees off on my desk to next ticket.

Student: (Refer Time: 17:33).

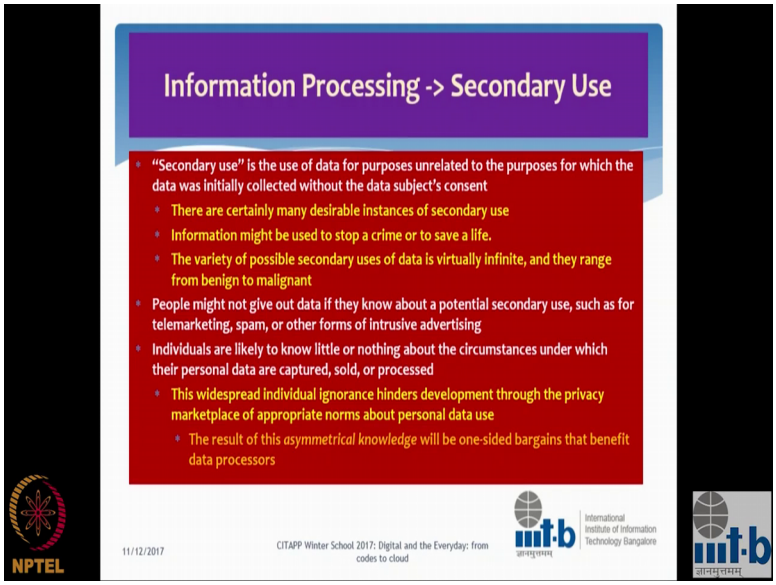
Right to a concert it is good for me.

Student: (Refer Time: 17:36).

Right otherwise I would not have found out. Now, I have got it right. What about the bad thing? It is quite possible that I might be profile and I might you know can be used for some other purposes right. So, there are there are tradeoffs right and so, we will discuss both information goods as a market later on. So, this gives us some framework you know for example, aggregate information. If it is taken without my consent it is not a very serious violation of privacy, but it is there you know between 1 and 5 something like that I can put a number to it right.

The second important aspect of information privacy is secondary use. This is a gross violation I would say what is secondary use.

(Refer Slide Time: 18:18)



You collect information for one purpose right it I use it for entirely different purpose without even you know getting a consent form you right. So, it is possible that you might link our Aadhaar card to your bank account right. The primary intention being that you know everyone cash transaction should be tracked and you know they should be a accountable and things like that.

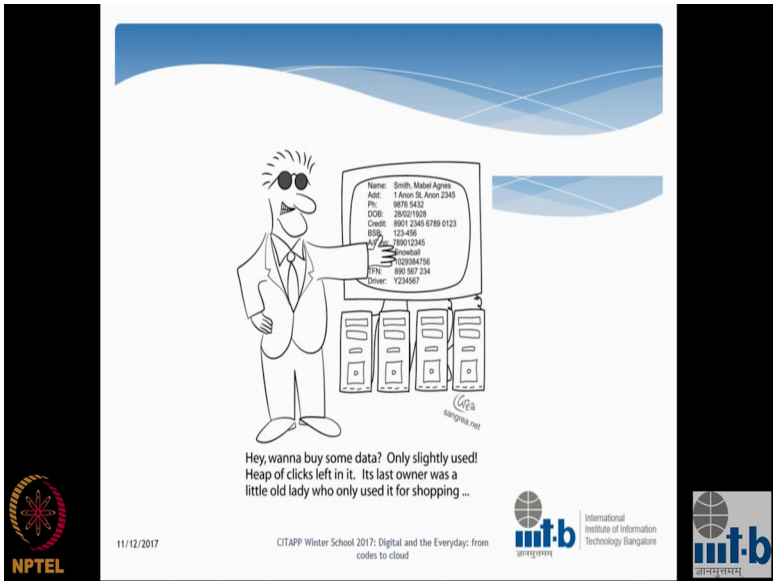
So, we should not have a you know bank accounts without a specific personality associated with it so, there has to be a linkage and so on. But if it is used for some other purposes then it can called as a secondary use and secondary use is a serious violation right because you are not taking the individuals consent. There is absolute information asymmetry between me and Google right. I mean knows my search pattern, but what Google is going to do with it is absolutely unknown to me right there is absolute information asymmetry.

So, that information asymmetry is being used by the data capture person right. It can be easily monetized right it can be leveraged, it can be intruded upon. I do not have any way of knowing because I do not know exactly how the data collector is going to use that information for right and therefore, secondary use is a biggest concern. So, in scale of 1 to 10, how would you rate secondary use as a violation of privacy?

Student: 10.

10 so, this cartoon gives it alright.

(Refer Slide Time: 20:08)



Last owner of this particular information set was used only for shopping, but you can use it for something else. So, the secondary information can be traded the primary data collector. For example, Google may not using the data, but if it sells the data to someone else say for example, insurance company. That insurance company might be able to use the information that Google has provided in order to profile me, sick person, healthy person, rich poor whatever it is right.

So, the secondary use not only exist between the data object and the immediate data collector the primary data collector, but it can actually propagate right. That is one of the important aspects of secondary use and it is can cost serious violation because you really do not know who is using it because that chain of secondary data collectors is unknown to you right and there is absolutely information asymmetry between the data object and the data collector. So, this is a serious problem and people take notice of it.

(Refer Slide Time: 21:20)

You know for example, WhatsApp. We all use WhatsApp and WhatsApp was bought by Facebook last year and now Facebook is part of WhatsApp right. What is it that (Refer Time: 21:32) WhatsApp collects, that Facebook does not have? WhatsApp collects some information, Facebook collects some information. What is it that WhatsApp has that Facebook does not have?

Student: Mobile number (Refer Time: 21:45).

Number.

Student: Contact.

Your number Facebook does not need number, but WhatsApp requires your mobile number right and therefore, WhatsApp knows using your mobile identity your GPS location, the places where you have visited, your conservation that you have posted. All those things are known to WhatsApp whereas, not known to Facebook.

Student: And also the friends contact.

Friends contacts everything right friends friends phone number everything it knows right. Facebook does not know does not have any information about mobile numbers right, telephone numbers is not known to Facebook. It is only goes by Facebook login ID. However, Facebook has bought WhatsApp and therefore, you do not know exactly

what is happening. With the information that is collected by WhatsApp right. So, it is quite possible that WhatsApp can generously hand over everything to Facebook and Facebook with Facebook information can profile you much much better because it knows much more granular information about you right.

So, it was taken a note off by Supreme Court right so, supreme court is still case is in progress. They said that you know what is the deal between WhatsApp and Facebook right can WhatsApp share all the information to Facebook right because WhatsApp. When you are signing up for WhatsApp I am not signing upon for Facebook. It is quite possible that WhatsApp is owned by Facebook right. But when I am doing WhatsApp I am only contracting with WhatsApp as a service provider right and so, for example, WhatsApp has created this peer to peer private conversation.

So, if I make a conversation to you on WhatsApp it is guaranteed that it is encrypted right using peer to peer without any central key and therefore, it is not possible even for WhatsApp to crack it right. That is the guaranteed encryption protocol that WhatsApp gives for communication between peers or in groups right, but I am not saying that WhatsApp can share the information to Facebook. So, that Facebook can profile me based upon my telephone number right and this is secondary use this exactly secondary use right.

So, you collect information for your primary purpose, but it is shared with someone else is the secondary service provider right so, supreme court has take a note of it and the case is in progress. WhatsApp has been asked to explain, what is the information that it is sharing with Facebook right? So, we are conscious of it right on the other hand the U.S. under the chairmanship of Ajit Pai who is the (Refer Time: 24:20) communication commission chairman, has reverse track in the recent ruling what they have said. Is that you know internet companies like Facebook, Google they all collect information, they all use it for whatever purpose they want right. The telecommunication companies like Vodafone, Airtel right.

Student: Jio.

They are all bounded by a regulation they were not able to collect information that they wanted for their own use let us free that one also right. So, Ajit Pai has said that telecom companies can actually collect information. They can sell it to the secondary or third you

know tertiary service providers without getting the users consent it is a major overruling of the previous administrations protection bills right.

So, countries or taking sides now is secondary use good or bad suppose. For example, I you know we all know that Google that is the secondary use beautiful right, that is one of the reasons why Google search is free. The reason why is search is free is that because of your organic search potent you get the advertisements which are targeted to you right they know what you want and therefore, they are so, that information is passed onto the advertiser. Advertiser posts the advertisement and you click through it Google earns revenue and you also get coupon. So, secondary use has some benefit looks like right. But is it to have secondary use Google does not, does not explicitly asked you before posting the advertisement, oh do you want this advisement to be posted. Does it ask you anytime? It does not.

Student: But we can give feedback with that on a each and every act.

Opt out right all the policies that we have are all opt in I mean this is basically sorry these are all opt out policies by default you are opted in.

Student: Yeah.

By default you agree to that terms and conditions, you do not even read it because it is so, complicated it is not user friendly right. So, it we can only opt out for example, you know that there is something called as do not call registry (Refer Time: 26:30) keep on getting SMS. You know your mobile number has been leaked by the Telco's to all the service providers around the world. They keep on sending you SMS, insurance agency call right opt out of policy right by default I am opted in. If I want to opt out then I have to put my name in the do not call registry most of the time it does not work, but if it works.

Student: (Refer Time: 26:51).

Then you will not get that SMS from that particular company right. So, the question is should we have opt in as a policy right so, before secondary use I have to explicitly opt in. Otherwise I should not be by default opted in and that is a policy question right that is a policy question. So, as of now most of the secondary use survives because we have in

most of the cases opt out policy, unless you explicitly opt out by default you are actually consenting to secondary use of that information. Is that correct? So, how do you rate it now in a scale of 1 to 10?

Student: 15.

On a 10.

Student: 10.

10.

Student: 6 7.

7.

Student: And if it is if it is not 1900 it is a multiples with refreshing if you are linking Aadhaar and then associating with some kind of crimes and it can be friend or something except this laws I mean except this kind of cases I think for everything you see this is a open market. So, why that, why my data should not be?

Used.

Student: Yeah.

So, you so suppose a so, there is a claws suppose say Google says that you know if you divulge this information I will give you coupon worth 250 rupees. Then I obtain to get this advertisement get the 250 rupees coupon then it is then it is not violation of privacy know.

Student: Yes sir it is consent.

It is not a, it is by consent.

Student: Yeah.

So, as long as there is consent there is no violation of privacy.

Student: Correct, but they can. But they can use it.

Which one.

Student: They can use it for secondary.

Ha the important thing is consent you give based upon what this is the purpose behind which the secondary use user is going to use it. But as we have discussed before there is a absolute information asymmetry between you and the service provider right you I mean there service provider might say something, but you can use it for some other thing you know it is very difficult to find out whether my consent is for this purpose or for that purpose its very ambiguous right so, but anyway. So, one of the ways by which we can surmount this problem is using opt in instead of opt out that is possible scenario right. So, this is a case of secondary use and this is why important.