

Course Name: Cyber Security and Privacy
Professor Name: Prof Saji K Mathew
Department Name: Department of Management Studies
Institute Name: Indian Institute Of Technology Madras, Chennai
Week: 12
Lecture: 36

Okay and now we come to the important discussion which is about the need for privacy and the need for unique identity. So both are actually important and therefore there is a need for regulation and as I said government initiated this well in time and a bill was available by 2017 and they did excellent job I would say, in drafting the first data protection bill, personal data protection bill of India. And the terms of reference, if you read what the government asked the Justice Sri Krishna Commission to do, you know it is a very difficult order, to unlock the data economy while keeping data of citizens secure and protected. You see what comes first, to unlock the data economy. So for digital India being digital is important, do not ask to stop being digital or as we discussed in some of the cases, go back to the manual controls and do not give internet connectivity or protect the privacy of people and then nothing more. That is not what the government did.

So you should not actually deter or destroy business models or businesses that run successfully on data and create a lot of value for citizens. And so by the time this, this TOR was given, Facebook is already in India, Uber is in India. What does Uber work on? Is it on anything, any assets? Just data, purely data. So data based business actually has become huge success across the world.

So Google is right there and they are making car. So that is a new economy and do not destroy it but at the same time take care of citizens' privacy. That is the order. So government recognizes, recognize the transformative potential of the digital economy to improve lives of Indian citizens. At the same time upholding the privacy of individuals was also important and therefore there comes this regulatory, initiative for regulation.

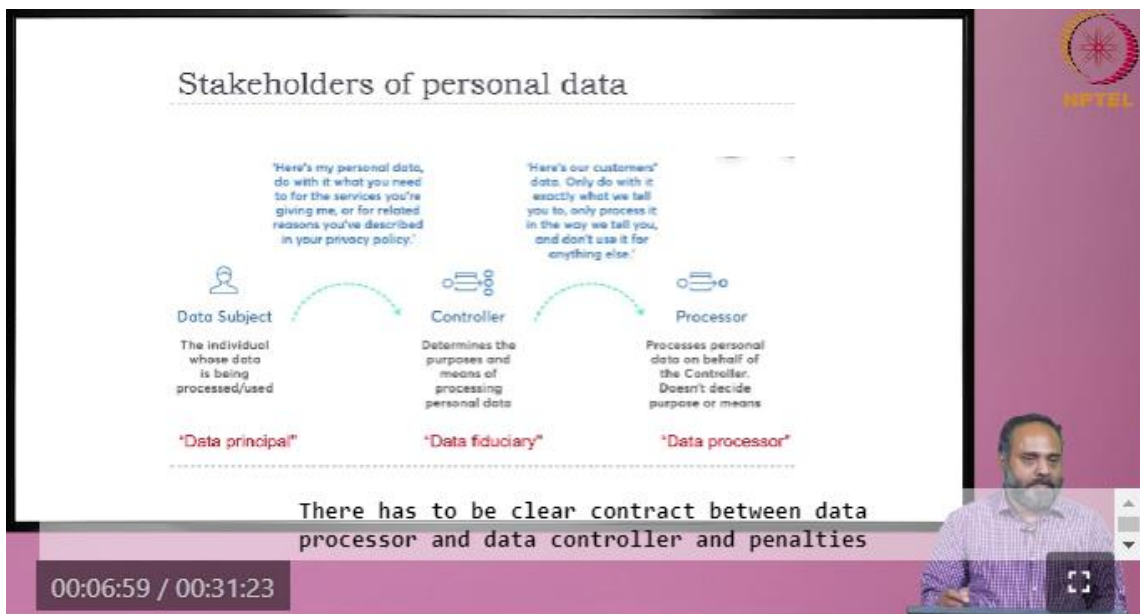
Government recognize the need. And of course, this PDP was presented by the Union Cabinet in the parliament in December 2019. And but, of course there was opposition because what happened was that the PDP as drafted by Justice Srikrishna Commission was not presented as it is. Government made amendments to it. The amendment was basically to give more power to the government or government's right to access personal data.

And particularly there is a clause which actually provided this, you can call it undue powers but government sees that certain powers they need to access, to have access to

personal data. So this amendment made by the government was not acceptable and therefore it did not conclude and government finally dropped this in the last year and then they said they will go for an alternate bill. So if you look at the PDP which is the base document, the current DPDP is building on it or it is making it more concise and it is making it, I would say more inclusive. We will come to that. So you can draw the parallel between the GDPR and the PDP by looking at this sketch.

So there are three actors or three major stakeholders in data. One is the data subject. So in PDP and DPDP the data subject is termed as data principle. Data principle is you and me, the individuals, persons whose data is collected. So that is a data principle and data controller is called data fiduciary in the PDP language or in our bill, the data controller.

Data controller could consist of multiple entities but essentially the best example is Aadhaar. Aadhaar is a data controller. What are the functions of the data controller? It will collect data, it will store data and it will also take decisions about the data. For example, with whom that data will be shared or what kind of analysis or processing will be done on the data, who will do that processing, would this be by the controller itself or by an external entity. So there are lot of decisions that actually is vested with the controller.



So therefore GDPR calls that entity a controller and in our language, it is a data fiduciary, the one who actually is the guardian, fiduciary in the sense who is a sort of someone who can be trusted with data. And the third entity is the data processor. So same in GDPR as well as in PDP. So GDPR, sorry a data processor only processes data that is shared by the data controller. So if you outsource analytics, so the processing agency is the analytics

company.

But analytics company is not the data controller, the data is collected and decided on by a data controller. And when we discussed GDPR, we saw that since data passes hands and the data principle loses control on with whose custody my data is and who is analyzing my data. The GDPR made data controllers responsible for their contracts with the data processors. There has to be clear contract between data processor and data controller and penalties would apply to data processors as well in the GDPR. Before GDPR that was not the case, the data processors are not liable.

In certain company, in certain contracts there may be liability clause, in certain contracts it may not be there. But GDPR made this domain aware that there has to be clear contracts and financial implications. Well, I will move on, this is a bit political. When government actually brought the PDP to the parliament, Justice Sri Krishna said this is not the draft I drafted. This will result in an Orwellian state, George Orwell, we referred before.

So, because government actually amended it and more sort of, so brought more controls on it. So I am not politicizing this, so it is about the government, it is not about a party because government needs, feels that it needs more control for governance. We saw the history of privacy in India, so it cut across political parties. So now we are actually discussing a bill called the DPDP bill. Data sorry, Digital Personal Data Protection, PDP with the digital added.

Personal Data Protection, it has changed to Digital Personal Data Protection bill, DPDP in short. So this is under discussion and it is open for public comments. You can just go DPDP bill, you will actually get this in public. It is a document that is available in public domain. And one interesting point here is this DPDP 2022 has 22 clauses and experts notes.

So lot of things that I am going to present now is based on expert views. I am not a expert in law, I am a teacher of information systems. So interpreting these bills also require legal expertise. So I am not getting into that domain, so I am making some general observations and also referring to opinion pieces that came in business newspapers. So that is my reference here because there is no journal article or there is no very highly sort of unbiased opinion available.

So it depends on which newspaper you read also, you get two sides of it, so that also exists. So the one point is the, out of the 22 clauses, the central government has been provided with rule making power in around 14 clauses. What does this imply? Is this okay? Government can actually make rules about 14 clauses in the regulation. So a

regulatory body is supposed to be independent, free from the influence of government. RBI is supposed to be an independent institution and it should take decisions for that particular domain not dictated by, even by government, that is the point.

So you see what is called conflict of interest. I am not suggesting that any government would just let data go without a control on it because then it will actually lose ability to do its function of knowing people or investigating about people etc. So government has a legitimate need but there is also a conflict of interest here. That government itself is the fiduciary or the data controller in several cases. Look at Aadhaar and today you have GST, of course GST does not come under personal data but government is in possession of huge amount of public data and then government also becomes a decision maker or government can decide on the regulation.

So it is like, I make the law and I execute it, so both are not correct. An independent judiciary, now we say the judiciary should be independent of the legislative and the executive, that is to make this, the estates independent. So that there is no conflict of interest. So there is a arguably a conflict of interest situation here but that is not a final opinion, it is an opinion by an expert.

I am just quoting it. So if you make a comparison which I tried based on expert opinions available in public who are lawyers of course. So this is sort of tertiary data. So I am actually compiling what is said by somebody else. So the first point is my observation, PDP is very detailed, there are 14 chapters, it is publicly available and 56 pages and the DPDP is 6 chapters and 24 pages. So it is made very concise, it is half in size, very very concise.

A quick comparison

- › Detailed PDP (14 chapters, 56 pages) vs concise DPDP (6 chapters, 24 pages)
- › Addresses only personally identifiable data; non-personal data not addressed
- › The right to be forgotten now under the right to erasure
 - › Right to be forgotten covers consent for data sharing
- › Penalties much higher in DPDP, cap at ₹500 crore; penalty applies to data principals also
- › Dropped three tier classification of data in PDP (personal data, critical personal data and sensitive personal data)
 - › In PDP the last two must be stored in India

So the first point is my observation, PDP is very detailed, there are 14 chapters,

00:12:16 / 00:31:23

And DPDP addresses only personally identifiable data, it does not deal with all kinds of data where PDP was trying to regulate data as a whole. So the purview is slightly different and there are some minor issues like the Right to be Forgotten was a clause in PDP but now it is brought under the Right to Erase, it is called the Right to Erasure, but both may not be the same, that is what the legal opinion. Right to be Forgotten also would apply to data sharing among entities but the right to be, try to erase data may not ensure that, this is one minor issue. And penalties are much higher in DPDP, the new bill has higher penalties for breaches. It is capped at 500 crore, that is the maximum that would be chargeable for data breaches.

And interestingly in DPDP penalty applies to data principals also, that is if you actually make a wrong claim, you may actually end up in paying penalty, which was not the case with PDP. So you can see there is a bit more influence of the other stakeholders, not just the individual's perspective, it also actually looks at the regulation from the corporations point of view or the business entities point of view. So they do not want this irritation of complaints all the time coming. So this will actually probably reduce the number of complaints. But the last point is very critical point, so we will close with that.

In PDP there was a three-tier classification of data, the PDP classified data as personal data, critical personal data and sensitive personal data. So let us go from the other end, sensitive personal data or SPD means data that is very sensitive, it could be your financial data, it could be your, data about your what you call, personal preferences like your matrimonial data. So a lot of data that is very much linked to your individual space and it

is very also sensitive. Passwords, that is SPD, health data, that health data, financial data, passwords, all these are highly sensitive and PDP brought this under a category called SPD, sensitive personal data. Then it had the second classification which is critical personal data and it is not defined in PDP but it gave the rights to the government to decide what could be critical personal data for the country and government did not define it.

But if government wants to tell organizations, well this data, this category of individuals data should reside within the country and you know, other data can be passed or stored elsewhere etc. Government had that control possible. The rest which was not critical or sensitive is personal data. But in DPDP there is no such three-tier classification of personal data. It only says personally identifiable data.


So this classification is done with. But in PDP, so that is where actually the politics or the debate actually comes in. In PDP the last two categories, critical personal data and sensitive personal data, PDP recommended it should be stored within the country. Cross-border transfer of SPD and CPD was not allowed. And that is called data localization. So you must have heard about this issue of data localization.

Data localization is proposed in GDPR. You can see cross-border transfer of data is expressly permitted with only a few countries. And for rest of the countries, it has to be based on contract. So therefore this particular aspect of data transfer when it was considered in the regulation in India, they made data localization a clause. But data localization also known as data residency, data sovereignty in some literature basically requires that citizens data, citizens that belong to a country, their data should be stored in data records or archival within the country.

It should not go outside of the country. And you can see that after PDP, government started arm wrestling with WhatsApp, Facebook, Google, payment banks, all of them. Citizens data should be within the country. Now, is it good or bad? What do you think about? Data localization is a highly debated topic. So on one side government which is or a regulation which is trying to protect personal data wants data to be stored within the country. But what is wrong with it? That is the, that is data sovereignty.

DPDP amends that. But what is the problem here? So one issue is that the data actually has value. So we can analyze the data to get some, identify business patterns and so forth. So if in case there is no data localization, multinational companies could take this data to their home countries, analyze the data and develop products for the Indian market from abroad. So potential source of jobs in our country will be lost. Like you can, they can build customized solutions, say customized apps on the Play Store from the US using data which was acquired from India.

So you want data trade or data to be taken out of country for analysis. That is what you are saying. For analytics, analytics. An analytics could happen in India and the results of analytics could be used to develop apps and other digital services within this country. So that value is being transferred, being taken out of this country through, without data localization.




Data localization

- ▶ Data localization means data about the citizens of a country to be collected, processed, and/or stored inside the country, often before being transferred internationally. Also known as data residency/data sovereignty.
- ▶ How does this affect different stakeholders?

That is what I think.

00:19:58 / 00:31:23



That is what I think. But don't, foreign corporations actually outsource analytics to India. We are actually a place where we have the competence to do data analytics in the country. So in one sense the country is saying that, well do analytics within India. It is actually a boost to the Indian analytics as well.

Don't transfer the data outside of the country. But you are talking about scenarios which we do not know where data analysis in other country would bring more insights. Okay. Any other reason? Mostly it is related to technical trade. When data is stored in other countries, if any, as per law, if the data is stored in US data center, if the law forces them to the company, for example, Azure or anyone, to share the data, it is their liability to share it. So if the data is stored inside India, the government has protection over it and it is not liable to share with any other countries, until it is the concern of India.

Okay. So essentially you are saying if Indian citizens data go to a country X. So it comes under the jurisdiction of that country X. And if data sharing is permitted within that country, like many countries, if a data trade is free, so this data which is residing in their jurisdiction can be shared because it is no more within the jurisdiction of India. So

basically the PDP or any data regulation to be effective, data has to be within the country. Once it goes out, it is in another location, the same laws cannot apply.

So that is a valid point. So data localization is essentially trying to ensure citizens data are protected within the law of the country. Agreed. But this is like skewed, this is over protecting individuals. Why corporations should worry? So there are multiple stakeholders, there are data subject is one but then there are data controllers, they are collecting data for some purpose. What is their concern? One is potentially if they are analyzing data elsewhere in another country, so they have to actually change the contract.

So there could be complications, so that is one. But is there any other major issue? This actually if you look at DPDP, this clause has changed, data localization is diluted. So post PDP government actually was trying to have complete control and suggested data localization but now it is not. DPDP does not have a strong clause for data localization. That is actually by corporate power. But what could be other concerns that corporations have in data localization? Actually this may be right or wrong but I remember reading few years back that security is a major concern where if the data warehouses, like multiple, some MNCs data warehouse when their data centers present in India are not that secure and whereas there are some specialized data sources in some islands or something where the data can be secured more safer or something but they belong to some multinational organizations.

So they are willing to come forward but if in case we are, like the Indian government is in, to go to, like you know if they are to go in for contract with those organizations they will be obviously losing some control over the data. So I remember reading that it was, you know the trade off between security and control. Security and control. So what are you arguing? Suppose data goes outside of the country.

We will have, the data might be more secure. It might be, I have read somewhere, I am not too sure of the context but yeah. It might be more secure because of the data center. Depending on which country. Not which country but organization I remember. Some specific organization had its data center in some island and they were about to, some talks are going to exchange data with that particular organization but again control was some issue which they were discussing about.

Okay. So let me write down. So plus side is government's control, the minus side is corporations losing potential value. Are there any other pluses and minuses you can think of? What is the, what is a major loss because corporations actually arm twisted the government to change the law.

You see. There could be other potential reason. Okay. So let me give you clues. What do you think about the data center economics? One potential negative is like it will increase the costs of operations. So earlier they used to have clouds where the actual warehouse is not actually within the home country. It will be in some other country. Now they have to build dedicated data centers within each country.

Yeah. Yeah. That is a major issue because all corporations or large corporations like the Google you search for where are the data centers of Google located. They have actually rationalized that across the world and they have, particularly Google has built their data centers close to hydroelectric projects basically for saving costs, saving on energy. So this is something they have already invested in. So data localization will require these companies to build data centers within India. So on the plus side, a data center business, data center and analytics business will prosper within the country, the plus side.

But cost of data center or cost of data storage would definitely go up. They do not have the scale and they enjoy scale advantages when they actually store data within certain premises, within certain areas. And that is something they have rationalized. The cost goes up for them, if they have to build domestic data centers.

For India, this becomes a new business. It becomes new prospects for the IT industry because data center business comes to India. So government has that rationale but not corporations. And you can see the corporate interest actually, because if corporations become non competitive then they would not like to do business or it is not possible. So therefore you can see this particular balance. It was actually skewed towards the government or privacy but now with the DPDP, government is also permitting cross-border transfer of data.

The sensitive personal data, critical personal data, this kind of classification was also dropped because if that exists, the argument becomes strong but now there is only personally identifiable data and data localization is not now very strict in the DPDP, thanks to corporate interest. But we should not see this as a, you know as simply negative. It does not make economic sense for corporations. But if you talk to consultants, they will also say that well, is data center a huge business potential for India? Many disagree because data centers do not create many jobs but they use up a lot of domestic resources. For example, if Amazon, actually there was a negotiation with Amazon in Hyderabad to set up a data center.

So Amazon's conditions were that they want, say 100 acres of land and they want full security given by the government around that particular infrastructure and they want no taxes. So it should be all favourable for them to establish a data center here free and they

may employ, say 50 to 100 people from India. So this is one insight I got. So it shows that data center as such, IT services is different but just having a storage facility does not really create a large number of jobs but it actually take away our resources. So probably these are reasons why government also relax this sort of strict condition on data localization.

Okay. That is about Indian regulation. So it is on debate, so you can continue to watch this. Any questions? Yeah, so beware of newspaper articles if you have to balance between news, expert opinions because they can only look at certain, some overly would be critical of government or very much in favour of government but there are pluses and minuses and government has to balance these different forces. Corporate interests are also important, you know, you cannot just neglect that but at the same time, privacy is also important. So let us see where it reaches and that is it for today. So next class we will discuss the economics of privacy because we touched on that today, There are different players or stakeholders in data and therefore we cannot neglect any stakeholder from this.

And we also say that data is the new oil. If oil becomes very expensive then you cannot do business with data. Manufacturing business will not be competitive, oil is very expensive and so is many corporations. If the privacy laws becomes very tight and data becomes very expensive then those businesses cannot survive or be profitable. So we have to look at that side, data is the oil. So it has to be protected but at the same time should not become very expensive to do business with. Okay, we will meet in the next class.