FINISHED COPY

APrIGF 2017
“ENSURING AN INCLUSIVE AND SUSTAINABLE DEVELOPMENT IN ASIA PACIFIC:
A REGIONAL AGENDA FOR INTERNET GOVERNANCE”
BANGKOK, THAILAND
29 JULY 2017
WS83
UPGRADE REQUIRED: OBFUSCATION AND CRYPTOGRAPHIC STANDARDS FOR DATA PROTECTION LAW IN THE AGE OF LINKED AND BIG DATA
1:30 P.M.

Services provided by:
Caption First, Inc.
P.O. Box 3066
Monument, CO 80132
1-877-825-5234
+001-719-481-9835
Www.captionfirst.com

***

This text is being provided in a rough draft format. Communication Access Realtime Translation (CART) is provided in order to facilitate communication accessibility and may not be a totally verbatim record of the proceedings.

***

>> ARTHIT SURIYAWONGKUL: We will start soon. We are still waiting for some of our speakers to come. Thanks.

She's coming. She's in the parking lot. It's always a problem with speakers from Bangkok. They have a problem of car park. We have the privilege of not having that problem coming from abroad.

Okay. I think we can start. So maybe a little bit of the reordering, so we can start. So maybe I will just give a background first of this panel. This panel comes from the discussion during several APrIGF in the past, including in Taipei that have a lot of lawyers and they're like, yeah, we should come up with something that will project, right, the security, the privacy of Internet users and consumers in general, about consumer rights, about political rights, civil rights, but then there's a question of, oh, there are a lot of bridges that the law is not updated enough to cover those bridges.

It is not actually a bridge according to the law, or there are some laws that lie accorded. Or these are the things you should do, one, two, three, in order to protect the data of the citizen. But those criteria, it's no longer relevant. Some of the things that work in the past, like ten years ago, the technological measures is no longer work now today, because the assumption, right, that used to pass the law ten years ago is no longer true. This assumption is no longer, true, right?

For example, when we are talking about the length of the encryption key, right, it used to be suggested, like, okay, it should be this X long, right, because we think that the computing power back then, ten years ago, if you are going to craft this, it's going to take, like, 200 years time, which is in terms of personal data, it's not relevant because the average life span of a human is less than 200 years.

So given that this encryption key, it's possible to protect someone for 200 years, it's good enough, right? But then later we found out, oh, with the advancement of algorithms, advancement in computing power, together with all the techniques it, shortened the time, 200 years into, say, 10 or 20 years, which in this case is well still a long time, but it's shorter than the human life span.

So in terms of the protection of the person's data, it's no longer good enough. So this is, I think, the key question, why we have the panel today, talking about -- when going to, like, code something, in terms of standards, as well as the law itself, as well how we are going to think about this, to maybe the regulation, actually future-proof.

So to -- I think probably we can start with Juggapong first. Maybe I will let each of you introduce yourself first, and then back to Juggapong to give us technical background of how it is possible in terms of, like, the technical qualities, right. Maybe introduce yourself first.

>> JUGGAPONG NATWICHAI: Juggapong Natwichai from Chiang Mai University from the north of Thailand. My background is computer science and I'm working on the data privacy in terms of the technical part of that, not policy, but I can chat with you my thoughts about that as well.

>> My major job is the development of the data and artificial intelligence, technology, and in this year, we also work with governments, also industry to set up the data privacy framework, try to address the requirement that when government or enterprise try to release data to the third party, then there is some

identification of the data. Make sure the data cannot be reversed linked to the personal identities, and we follow the international standard mechanism and technology, and we will brief you on our framework later.

>> ARTHIT SURIYAWONGKUL: No Mr. Juggapong, please.

>> JUGGAPONG NATWICHAI: Okay. Actually, the plan is that -- to have Dr. Thitirat to go first with the overview, but I will show you my background work first. It's technical. If you think about the topic of skilled data or prevention release of the data or sometimes we call it data masking, right, to do that.

Okay. I show you this is the outline -- I mean, the guideline from Korea, this one. And I think -- anybody from Korea?

Korea?

Okay. Because I -- I look at this one as an example. This is the guideline to the identification of personal data, right? In that guideline, it's really -- I think it's very well designed. They are composed of four steps before we can release the data or we can disclose the data for the public for other uses.

The idea may be very -- not really safe, if you think about customer point of view or citizen point of view. It means that after some organization trying to collect our data by any means, using some of their services, right, you -- for example, one of the data that now being used in Thailand, quite a lot, is the data from the cell phones.

Let's say you all have, like cell phones your hands. Usually you be tracked by the provider. Of course it's not like 5 meters or 10 meters, but at least, you can be located, like, in cell site, and we can track somehow to which location that you can go.

That type of data should not be released, right, in terms of regulation and the viewpoint of laws, however, there could be some services that could be improved, I mean, for human life to use this kind of data. We compute a lot, and trying to put more resistance to the cell system, and maybe some of the partners of the telco can put some of the merchandise or to make people's life a bit better somehow.

Occasionally, when you are trying to get that kind of data, there could somebody regulation involved, right? However -- however, the data to be released has to be somehow limit firstly, the ID, right, the identification part, for example. Right? Name. Maybe citizen ID, or maybe part of, like, date of birth or something like that, right? So this kind of framework in Korea, that I just found is composed of four steps. I think it's really, really common.

Okay, the first one, they are trying to do the preliminary review with this data to be released is contained at the ID or contains personal information. Some not, but some may have something related to this kind of data.

And in the second step, they are going to do the de-identification. Of course, the name will be removed. The ID will be removed and something specific for people would be removed.

And then in the first step, they are going to have a bit of further de-identification. In the Korean guidelines, they are even at the eliminating. This is a standard algorithm for being identified again. The fourth step, of course, they can release this kind of data. That's the general view of the process.

However, if you are familiar with this picture. Anyone in the room seen this picture before?

Right. Very interesting. This has been used a lot to show that the data, the releases are not that safe anymore.

This is from Sweeney, a professor in some -- yeah, yeah.

(off microphone comment).

At that time, 2002, right?

The theory is like this, suppose you go to some hospital, right? And, of course, all the data can be recorded, right. On the left-hand side, they have everything, visit, diagnose, procedure, medication and total charge, right? That's very, very private data that should be hidden or maybe it should be recorded at only the hospital, right?

But this kind of data, really, really, commonly requested from maybe researchers and maybe even medical doctors or maybe someone else would like to use this data for, I mean, maybe public good, right?

And then they say, okay, to release this kind of data without any ID could be safe. However -- however, it's no longer the case, right? Because we are in the world that everybody trying to collect the data.

People may have data out of the left-hand side. In this case, Sweeney showed that she can obtain what she lists for democracy election, right, and she can overlap and cross-examine the data to get the data. She's trying to get the information. From that data, suppose that the interviewer knows that this guy admitted to the hospital at that time, and maybe we would like to know what is the result of that diagnosis result, right?

So we know that he's from the area.

Okay. The area is Massachusetts, right? And there were only six people that had the same date of birth as this guy, right? And from six people we can narrow it down to only three because of the gender, right? And then from three, only one that matched the ZIP code. So it means that it's really easy to get one data which seems to be safe, no ID and then to do the cross-examine with other kind of data outside and to get to know the sensitive data inside.

Are.

>> ARTHIT SURIYAWONGKUL: You have different databases.

>> JUGGAPONG NATWICHAI: Right, we can mix it. So to remove from this article, it seems like only to remove the ID maybe is not enough, right? So that's why people invent the thing called k-anonymity at that time. I mean, almost 20 years ago, right?

And, for example, from this part, right, we can remove the name. This is really over simplified. We remove the name and then we are left only sex and age, right? So we can somehow do data masking on the age to make it a range, not exact age, trying to obscure the data from being discovered again. That's the idea for the k-anonymity.

In this case, the k is a number of the people that can be mixed together without any way to reidentify again this one. K is two, because two people can be somehow put together and you cannot reidentify. That's the idea, right?

And, of course, there are -- even though we are trying to do the data protection like this, it seems to be safe, but, actually, there are more problems to this kind of situation. For example -- okay.

(Off microphone comment).

Okay. Let's think about it. I suppose that we are trying to look at this data first, right? And if you look at it, it seems to be okay, right? Because if we know someone is male at age somewhere at this range, right, we don't know about that disease. I mean, we can hide one from two, right? It seems to be safe.

For example, if you do multiple release of the data, we update the data. We put more here, right? And we are trying to -- to do the data masking again to make the data something like this, right?

If you look at it, we have male and 63 and male 80. There will be somewhere like this, right? And you mix that disease into something like this. It seems to be safe, right? You seem to be safe because we can hide someone in a group of four, however, if you put more data like this, in the release of the data, we put more data, right, and we do data masking again, it seems to be okay.

But if you they about it, to do multiple releases of the data, from all the time, I mean, just at the data, we can still cross-examine to reidentify the data again. For example, if you look at that. Suppose we know that he was admitted to the hospital at this time, the second release of the data.

Okay, he will be somewhere around here, but we don't know which one of the four that he can be. Right? But however, if you look at this two data set, right, we will know somehow that he maybe don't have heart disease. He maybe don't have flu. So he can be somewhere around here.

>> ARTHIT SURIYAWONGKUL: So you compare different --

>> JUGGAPONG NATWICHAI: Right. Right. Right. We can still see.

We are not sure which disease that he can have, right?

But if you look at the third list of the data again, where could Jack be, right? If you compare, you can compare again and we can know which disease that he can have.

So this is only the issue that can be happen. I mean, because we always collect the data, however, it's still somehow trying to protect the data but from -- if the time goes right, and we can see that maybe -- maybe if someone is trying to attack, trying to reidentify, it can be possible somehow if we are not aware of that.

Or I show you another case and I pass to the next person like this. This is what we call the trajectory data. Suppose -- suppose that if you are trying to compute which clinic in the hospital, in needs upgrade, I mean, in terms of more doctors, right, maybe we are trying to improve the route after the patient gets first diagnosis until they are discharged from the hospital.

Maybe they start from this A clinic and D clinic and then diagnosis. It seems to be safe. But what if -- what if we are having an attacker knows about the trajectory of that patient. I know that this guy going to a hospital this morning and I try to follow him, right? And we know that, okay, he entered into this room and go into another room. If you are trying to get this kind of data, we can somehow have the diagnosis result, if this kind of data is released.

But I'm not talking that we shouldn't use the data, to collect data and use data, it somehow can be a good thing for people, but to protect the data is something else that we have to do.

Okay. So I pass to --

>> ARTHIT SURIYAWONGKUL: Thanks a lot. So this is the technical background of how things can go wrong sometimes, right?

We are back to the big picture, a little bit. We have now Thitirat, she's a lecturer at the faculty of law at Thammasat University.

Would you like to join us? Not now, right?

She -- she will give us an overview of the international framework and probably you are going to talk a little bit of the European Union data law, right? So I hand it over to Thitirat.

>> THITIRAT THIPSAMRITKUL: Okay. Good afternoon. Thank you for inviting me to this event.

So, yes, I think we have seen the -- from the research, the architecture side, how data can be de-identified and de-anonymity, and make it related to an identifiable person in reality.

So let me start with these chart, these pictures. I think you may be familiar with this one. This is one of a very famous diagram from Professor LeAnne Litvat. Usually we think about law, but in the professor tried to convince us that not only law. Actually, there are many other rules, many other norms, many other factors that actually control the behavior of individuals.

So law creates artificial constraints to enforce or to control individuals in the society, but also we have the social norm where we create cultural constraints. For example, in Thailand, you cannot pass by the elders without greeting them, something like that.

>> ARTHIT SURIYAWONGKUL: Hello.

>> THITIRAT THIPSAMRITKUL: Yeah. Hello. And also the market, which create economic incentives. And now what the law today focused about is actually about the architecture. So the architecture or the technology control physical or structural -- creates physical or structural constraints, and that comes to the next one, the next -- oh, it's fine.

Okay. The next diagram, that actually the law today is not trying to control directly the individual behavior, but also control the other factors that will let individual behave and the way that state thinks, it fits or it's suitable or desirable. Okay?

So we see that law does not only regulate us correctly, but it regulates many other things to make us follow the path that the law aimed to achieve, okay?

So let's see some examples, that is quite -- I think this is an example, an easy one that will relate to our daily life. For example, the technique that the law used in general, it implored general goals for everyone. For example, in traffic -- in traffic law, we have the law enforcing the goal of safe driving and we have discretion. We the driver, we have discretion how to drive safely. So that can be subject to individual interpretation of the safe driving. And then we also have the law that sets some more specific rules or instructive rules of desirable behaviors. For example, if you want to drive safely in certain location, for example, on the highway, you should not exceed the speed of 80 kilometers per hour or something like that. So this is very specific. So you have very little discretion to interpret this kind of law, right?

Then the third one, the law that tried to change the structure -- so I mean by saying law here, I mean the state power, trying to change the structure or the architecture of that environment. So that will control directly the behavior directly, such as the state might build a washboard road so we slow down the speed of the car.

Even though the law is saying you should not exceed the speed of 80 kilometers per hour, still some violate such law, right? But if you set these -- some architecture, change the architecture of the road so the car must physically follow that instruction. But in many situations, the state cannot change the architecture directly, because the architecture or the technology are not in hand of state. So they are in control by private company and many other institutions.

So here the legal technique today, it's trying to -- the law is trying to apply private company to import certain structure, such as enforcing driving control function in car engine. So if you want to change a lane, you must put the blinker on. Right now the law allows some company to do that, but in the future, the law might enforce every car company to do that. Okay?

So under the -- the data protection law, what we have, if we try to look at legal techniques, that have been used in data protection law in many jurisdictions around the world, firstly, of course, enforcing the general rules that every personal data should be protected, right?

But there we still have some discretion, some room for interpretation. What is personal data? How does the personal data should be protected? That is still under the -- the discretion of each actors or players in the environment.

Secondly, so data protection law also sets some prohibitive, instructive specific rules, such as the per person principle, or many other principles that, for example, the data must be accurate. The data should have the right to access this data. So these are the rules that we have in almost every jurisdictions. But still, we have different interpretation. Each rule in different society.

So what about changing architecture? Can the data protection law change the architecture of data protection or the process of data today? I could not find any example of the law that trying change the architecture itself. So if you can think of one, please let me know, because I -- I believe that actually the structure of the -- the technology of using data today, they are no longer -- or they have never been in the hands of state control. So it is in the private sector and also in the hand of engineer or scientists, rather than the hand of state from the beginning, that what I believe, but if you have other example, please let me know.

Okay. So then data protection law comes to the reason and the most used way to control individual's behavior in data protection environment. So it obliges private company and other public institutions, public institutions as well, the rules and the guidelines of technology that will be used to protect data such as encryption and the anonymity. You want to ask something?

Okay. Here. I want toed to make some point here before we look to -- sorry. Okay. Before we look at further specific rules. So when we talk about data protection, we talk about the right to privacy, but I think one thing that we need to make it clear here is that the right to the protection of personal data, or even the right to privacy itself is not an absolute right. So it must be considered in relation to its function. So how we use the data, how it can be useful for many sector,M., education, or many others in society, and it must be balanced against other fundamental rights. It's not only the data protection, it's not only right to privacy. It can be the freedom of expression. It can be the freedom of religions or other things as well that might be in conflict with the data protection.

And also, the other principle that we need to take into account is the principle of proportionality. So the law cannot protect personal data without considering other important factors and also cannot do it in the way that -- in proportionate when you compare with the benefits of the individual -- the protected individual. So this is the -- an excerpt from GDPR, of the EU, recital 4, which reminds us to think about data in the way that actually it is not absolute.

So it's come to actually what was said in the beginning. So we need to find a way --

>> ARTHIT SURIYAWONGKUL: You were not here they beginning!

>> THITIRAT THIPSAMRITKUL: I think I know.

(Laughter).

So we need to find a way how to maximize the data protection and -- but at the same time, how to minimize the murder of relevant actors, okay? So we have legal issues. We have some legal issues here. The first of all, what is personal data? So of course if it is not a set of data or information is not personal data, it is -- it falls out of the score of many data protection law, including the EU's GDPR as well. So if a set of data is not personal data, it must be treated as the other -- the data protection principles will not apply.

And also, so we look at -- sorry, I would like you to look at each of -- categorize a group of data which actually it can be overlap as well. Okay?

First of all, encrypted data, the data is already encrypted by some method. I think -- okay. So this is definitely personal data. So encryption is one of the -- is an appropriate safeguards to protect personal data, but -- but encryption, it doesn't make that data become nonpersonal data, which means even encrypted, still the data must be treated under the principle of data protection. Okay?

But by the method of encryption, that set of data can become anonymous or synonymous data. So this will create some gray area within the application of law. Okay?

Next one anonymous data. In many law, I -- I didn't find any law that says anonymous data is personal data, but in GDPR and also in many others, anonymous data is not considered as personal data, however, the non-identifiable data -- so here we need to be careful with the use of word here. Non-identifiable data with possible of becoming identifiable data later if it is associated or combined with other information, such as the one that we see in Juggapong's presentation. That is very good that the appearances it might look like anonymous data, you cannot identify the person, but with the other existing information, we can identify someone from that data.

So here, we -- it is very difficult. We have the difficult problem here is that some data how can we know -- what kind of related information access is out there in the world? So many --

>> ARTHIT SURIYAWONGKUL: Now and the future as well.

>> THITIRAT THIPSAMRITKUL: Right. Now and the future as well. Each data control her, they have different capacity depending on the technology, depending on the personal -- the human resource they have. So some can deanonymize, so we have the absolute approach, which say, if there is, even in theory, some way to reidentify that data, that data is definitely personal data under the protection of data protection law. But in the relevant approach, we have other approaches as well. In relevant approach, they consider more the burden or the practical reality of technology at that time, or even the potentials or the capacity of data controller or data processor, if you consider those context, actually, those who support the relative approach, who say that we need to look at each data controller capacity, and environment of technology at that time, to identify if it can be anonymous, it can stay as anonymous data or not.

So this is the big, gray area and the big debate among many experts, lawyers and others there.

The dynamic IP address that this went to last year, that the advocate general ECJ, they have different opinion there. So we have to follow the court opinions but little the advocate general opinion's was also considered an important one. So the court took the position of absolute approach, that if it can be even in theory combined with the other data and reidentify, that data is definitely personal data.

But the court, at the end of the judgment, that even if we have to wise call application of personal data, still it doesn't mean that every personal data cannot be processed without consent. There are many exceptions such as many other legitimate interests for detection and other legitimate regions that can be exception of processing data without consent. So we need to separate two issues. So the first one, what is the personal data? And the second one is how we deal with that personal data. Is there any exception?

So it must be treated separately, okay?

So the next one, the pseudonymised, it is still personal data. If you can make sure that no one can identify from that statistic data. So its not, but with some exceptions.

For example, for some are particular disease, if there is only one person in that country who has that disease, definitely you need to -- you need to delete that data from the set of the -- the whole set of data, and this one is established as rule in Japanese law already. So we know that the Japanese law has changed last year and become effective this year, that they tried to set some standards to use the big data without consent, but with the many means of anonymization, and the pseudonymised. So these are the character numbers or numbers, anything that identify a person, you must delete all of these data from the set of big data. Okay?

So should I stop here? Right?

>> ARTHIT SURIYAWONGKUL: We will come back later.

>> THITIRAT THIPSAMRITKUL: Okay.

>> ARTHIT SURIYAWONGKUL: But I think -- we talk a lot about this de-identification, re-identification, right, an example from like, different parts of the world, right? And also, like, you are talking about the balance. A lot of people, oh, yes, actually yes, data, big data can be used for innovation, for better, like, public policy, right? I think maybe on the panel, we have like a little more concrete -- partly technical and partly policy as well, from Taiwan. I think I will introduce you. Please.

>> Actually in Taiwan, we enforce the personal data information. The personal data protection laws since 2012. And under such kind of legal framework, that every -- everyone that -- everyone intentionally or unintentionally collects personal data they collect to third party, they have to face criminal charge. And in the other way, they have to pay some -- they have to pay some fine. And for every one record, they have to pay $520.

So after we enforce such kind of legal framework, that many government officer, they -- they don't want to release to industry because we know that in the era of big data, data by itself has -- data can be released to third party, reduced, for research and then we can find very -- at the same time, we have to consider how to release data without violating the personal data protection. And so as the speaker just mentioned, there's some gray area. What is a personal data? What is not personal data?

Actually, in current legal framework, in Taiwan, there's no formal definition of what is personal data and what is not personal data? This only says that the data can be -- it's personal data. But what kind of data, in we make data anonymous, but in some case, it can be -- I can identify later. So we need some standard to make the definition for the personal data -- the personal information.

So starting from three years ago, the government starts to set up some standard. Actually, we already have the global standards and ISO29, 100 and the ISO29-191. So based on this ISO standard in Taiwan, we have the -- the CNS stands for Chinese National Standard, CNS29100 and CNS29191. In this standard, we identify some process, if an organization wants to release data to third party, what kind of process they should follow.

So first of all, they should define their risk, and what kind of data they release and what kind of attack they have to face, and after that, they have to describe what kind of mechanism they use to de-identify the data. So it's sometimes a process, a standard for process, not for technical but for process.

So in Taiwan, we set up the recordation standard that follows the national ISO standards. And in the other way, we have some control process. It's also referenced to ISO29 and 101 process. It's more a detailed discretion of how to do that and we also have some kind of guideline, some kind of framework and guideline and to further government and for the company to follow.

And after that, we have two system. The first -- the first system is that we have the accredited organization. We set up the process as a business requirement of a company, and we help the company to define their data and define their process and we help the company to -- to go through the whole process and in the other -- in the other way, there some accreditation organization, they will check in you are process compliant with the national standard, and the international standard.

And after they confirm that your process was complying with the national standard, then they will issue the certificate to the organization. So that the system will -- we already set up in Taiwan.

And under such kind of system, what kind of technology we use currently, we use several different technologies to do the data, the identification. So it's just mix the average of the data and the other one is what Mr. Juggapong just mentioned, we need to make data more -- in some -- in some way, making data be not to easy to be traced back and we also have some other mechanism. This is also very traditional technology, statistical disclosure control. It was developed by the academy. It is also a technology more than 40 years.

And we also have another technology called differential privacy. Under such kind of technology, we can, add noise to the data.

>> ARTHIT SURIYAWONGKUL: Can you explain that, differential privacy. What is that?

>> For differential privacy, we just try to add noise to the data and try to make the data more --

>> ARTHIT SURIYAWONGKUL: Obfuscation?

>> Yeah. Yeah. We send the data. But for differential privacy we can add the noise to some other data, some record to the data.

>> ARTHIT SURIYAWONGKUL: To make it more difficult to reidentify a person?

>> Yeah. Yeah. Yeah.

>> ARTHIT SURIYAWONGKUL: Okay.

>> And we also develop some cryptology methodology, but also cryptography cannot contribute to the identification too much, because most state of technology is called homocryptography, it's very difficult and it's very computational. So actually, we have to see if such technologies can be deployed to the physical deployment of the data process.

So right now, actually, according to our standard, right now, actually, this the most visible and the most political method to do the de-identification.

And so this is the current status under such kind of standard and such kind of technology. Right now we have two ministry already adopt such kind of standard and the technology. The first one is the Ministry of Finance. They use -- they adopt the national standard and also the k-anonymous technology to -- to deal with the text income statistic. And after they use the -- they adopt the standard, they can adjust and provide the data to third party for research. And this the most important one, the Ministry of Health and Welfare. They have the health data and the social affair data. So right now they have adopted a national standard and also the k-anonymous technology to their social welfare statistics.

And right now, we are working on the national health insurance data because this is -- it's a very big database. In Taiwan, we have national health insurance for every people. So actually, it has the health record of every citizen. Actually, database, a lot of research, they want to do such data to do research. So the government doesn't want to release the data to the third party. So right now we are helping the ministry of health to adapt the standard and k-anonymous technology in such a way so they can convince the civil society that the data released third party is obscure.

Just as speaker has said, we can't make sure 100% that the data cannot be released back to the original data but we can say we do our best to -- we adopt the standard, the national standard and we adopt the state of that technology, and we do our duty that to make the data can be reversed back and we think we can release this data.

And we suppose the data. We try to September our government data marketplace. And everyone can contribute data and the academic of industry can just apply for the research, and once the project is approved, they can go to the marketplace and they can just collect the government data to do their project.

And right now, we are also with telecommunication and financial holding company, they also have a lot of data and they -- there's also -- they also want to provide the data to third party. So we also -- we are now -- now we are working with telecommunication and financial holding company to adopt the national standard and also the technology, such that they can just release their data to third party.

So that's a brief overview of the current status of data privacy in Taiwan. Thank you.

>> ARTHIT SURIYAWONGKUL: Thanks a lot. So I think, like, something in common, that we hear from this panel, it's like -- at least like -- take some of these public policy decisions. There's a lot of example from the medical, right, or the public health. Without the data, the government cannot determine whether to use their limited resources, like to build a new hospital or where they are needed the post, right?

But then again, there's the question of the protection of personal data so projects like data masking that Juggapong talked about. Well, in order decide about where to build a new hospital, you don't actually have to know the exact age of a person in that province. All you need to know is probably the range, right, of the age, of people in the province.

So by these techniques, the data will be useful for making public policy decision but we will not be able to identify a particular person, right, or in the case of, like, talking about the differential privacy, right, as you mentioned, it's like -- so putting a lot of noise, right, but in a manner, like, those new records that are added to the database, we will still, in total, right -- it still reflects the statement statistics or attributes, right?

So if the database has like 60% female, and 40% of male, the noise should be the same percentage as well and you add this noise into the database to make it more difficult to re-identify people. That's an idea. I would like to take something from the floor and probably we have, like, more comments from our colleagues. Can you -- would you like to -- would you like to?

Anybody with questions or comments? Yes, please.

>> PARTICIPANT: I'm a lawyer from Taiwan, and I found that the data protection practice since 2012, when the Taiwan personal data protection act was amended, significantly amended at that time. So I want to share some experience for Taiwan. For example, what Dr. Wong said about criminal liability, it was criticized a lot, because if you negligently disclose personal data to someone else, should you be held criminally liable? So there has been a huge debate in Taiwan, and it has been amended in March of 2016. Now you will only be held criminally liable if you intentional breach data or if you want to make some profit or gain some illegal benefit then you will be held criminally liable.

And I want to share one recent case. The Supreme Court make a decision in January this year, it was about hour health insurance record. It's the famous health insurance system, and the system protects all of our health and also it collects all the data of Taiwanese. And the government actually opens the data to all the research institution, for them to study medical issues or ill prove health, and -- but before the government released the data, the data was encrypted.

So the government, they think that they are not releasing personal data because the data was encrypted, and the research institution does not know who the individual are, but human rights society, I think there's one member attending this forum. They think that's not right. So eight individuals file a lawsuit against the government to prohibit the release of data to the research institute.

So the Supreme Court made the final decision on that issue. There are two -- there are two major results that I can share. First of all, regarding whether encrypted data should be identified as personal data. The case is that the data was transferred between two different government agencies, and before the data was transferred, the data was encrypted. And supposedly the second one, the second government agency shouldn't know the identity of the individuals, but the problem is there are common staff among the agency and that common staff holds of key of the encryption.

So the Supreme Court says if they are smarter, then the receiver of the data wouldn't have been seemed receiving any personal data because the data has been encrypted. So that's what the judge said.

And from the court decision, we can tell if the data has been successfully encrypted, and the receiver is able to unlock the data, then the judge will think that the data is not personal data and it can be used and forwarded. That's one implication from the judgment.

And another one is, I think what the other Taiwanese would be criticizing our government because under our current government protection, they have a wide range to use personal data. It's opinion its statutory data, they can collect and use personal data.

The judge think that for the Ministry of Welfare and Health, they are protecting the benefit of the whole society. The judge actually encouraged the flow of data. They want more research -- researcher to research the data and obtain benefit result for all human being.

Actually, the judge said if possible we can -- we show a lot of data to be released and to be further used. So that's something I can share from Taiwan.

And I think our law with this personal data, one is you can identify an individual directly. That's the easy one. And another set is the difficult one. That is if you can use data indirectly, identify anyone, digitize personal data. There's no way how to determine that and it's actually changing for different people. The data collector and the data receiver, because they have deepened the ability to cross track the different database.

So the sense that the data may be personal data to one group of people but maybe not for another group of people. So that's another thing that is affecting the interpretation of the Taiwan personal data act. So I think the second one is what this forum wants to find out, how to determine what kind of data is -- can be used to indirectly identify a person and what kind of data can not. And I think that's something that technology can help us. Thank you.

>> ARTHIT SURIYAWONGKUL: Sure. Thanks.

>> THITIRAT THIPSAMRITKUL: Can I ask for some information? This is very helpful for me as well as a lawyer. So when you said the criminal liability, you said it must be intentional and with the purpose of gaining profits. So they are the two criteria of establishing criminal liability or they are separate?

>> PARTICIPANT: Actually, you need to -- that's a whole set. You have the intention to gain improper benefits.

>> THITIRAT THIPSAMRITKUL: Okay.

>> PARTICIPANT: And it's more than just welfare. You need to have certain intention to do something.

>> THITIRAT THIPSAMRITKUL: Okay. So if someone intend to release data to breach data but without any benefits but without any commercial benefits, like, just to bully someone, would that be counted as criminal liability under this law?

>> PARTICIPANT: There's another one, damage another person's rights.

>> THITIRAT THIPSAMRITKUL: Okay.

>> PARTICIPANT: So that would be the second category.

(Off microphone comments).

>> THITIRAT THIPSAMRITKUL: Okay. So it can also be under the civil law, the usual categories as well?

>> ARTHIT SURIYAWONGKUL: Yes, please.

>> PARTICIPANT: Yeah. I'm from the Philippines. First, a clarification on the Taiwan case, because on the presentation, I'm very impressed because it seems that across government as an enterprise, there is an appreciation and all agencies are Marching in step that the data has to be protected because unfortunately, that's probably the exception in our country, in the Philippines. It's a new law, aside from lack of awareness of all of these things.

There is also different -- you know, different interests at play. So the national privacy commission, we have one of their officers there. That's one. Ways to get people from different agencies to March in step. How were you able to generate that general strategic that we should go in step. That's a comment and a question. That's not an easy task.

I guess we feel that everything is negative in the Philippines. That the health sector is a bit advanced in our country. The health committee was ahead of the national and a possible model that we could join to the discussion is the idea of sectoral privacy codes, where -- because there are just simply different domains that not all are familiar with. So in the health and medical domain, there are specific sensitive information that only health professionals will understand as such.

So there have to be developed -- we cannot assign that all of these things in different sectors to the national privacy commission.

So we feel that a multistakeholder approach is better, where it's multi-stakeholders, the sectors develop their own rules and standards.

Within the standards, there's also multistakeholder representation in what should eventually be a privacy board, a health privacy board, national health privacy board, they are advisory expert group which hopefully we want to convert it to a national health privacy board.

So I want to throw that in the context of government doesn't always march to the same drum. Especially in such a -- such an important issue like this.

>> ARTHIT SURIYAWONGKUL: Sure. Can we go to second questions and then back to the panel.

>> THITIRAT THIPSAMRITKUL: I would like to add one thing. Some trend in the -- in the law -- in the data protection law is what we see before, there were more sectoral law, but right now, many countries change to general protection. So I think that should answer you as well.

>> ARTHIT SURIYAWONGKUL: Okay. Please.

>> PARTICIPANT: Yes, hello, I'm from Hong Kong and from Open Data Hong Kong. I want to talk to the panelists, how do you see this relationship between data privacy and data innovation? Because that is -- I mean, if you are in start-up circles and everywhere, they are very data hungry companies, most of them, and they always assume the data they use is their asset and their IP, in a way. And so I wonder also in this whole context of smart city and so on, there as a lot of data intensive discussions. So I would like your idea of the data mark place of the government. You have open data but you also have the privacy very strong. And then I thought about this idea of architecture and privacy by design.

So could we have more innovation in this field? You said you have no example for data privacy by architecture but if I think about the car drivers, the most effective way is actually to have these bumps in the road. That's the only way we really prevent them. I was thinking we need to find something drastic for the industry, for the designers and for the software developers. So what can we think about?

Thanks.

>> JUGGAPONG NATWICHAI: Okay. I answer -- maybe I can't answer your question. I think the balance is very important because what I'm trying to go say is not -- we shouldn't have the mind-set we cannot release the data. That's very dangerous because otherwise, we cannot use anything because one thing that Wong talked about is, okay, I touched that point. They have differentiate privacy. Right? The concept has been introduced many, many years ago.

When we are trying to add some noise or even trying to synthesize the new data from the original data, it seems to be more safe to -- to people, right? It means that we don't release the actual data. We release all the syntactic data. I think one of the challenging problems -- I mean, the problem of that approach is suppose that someone used noise, added data, to do something, for example, in their start-up, right, can we guarantee the success of the data crunching, the result?

I mean that could be the -- a bad thing for the common as well. If he used that data to do something. For example, we try to predict the behavior of our customers. I think we should somehow protect the private sector as well. So I think that the key is to balance on this topic, not -- not trying to wholly protect the data.

>> THITIRAT THIPSAMRITKUL: Actually, if it's possible to move to my slides.

So actually, what I meant by saying that, I could not think about an example that the state -- the state law itself changed the actual architectures that I think even privacy by designs must be done by many actors, by many players in the industry, not by the state law alone because we know the state cannot control the set of the -- or even the Internet itself.

Okay?

But that -- can they move to the issue of inception? Our privacy by design is the one thing that we should consider being the guide line or becoming the new data protection principle for many stakeholders and we -- when we think about encryption, I would like to put to another legal issue here, is that how do we assess if an encryption sufficient to protect data?

So there are many elements that maybe you can put -- you can add more here but what I can put, identify from the -- at least from GDPR, as the strength -- first, the strength of the encryption algorithm. So this one is definitely depending on the technology itself. Second the length of the encryption key. But I don't know in the future would this be one element enough or not? Okay?

And, of course, one -- the other element that many forget is the security of the key management. So this is not only about the technology itself, but it's also about the management system, the privacy policy within the organization, and yeah go back to the Taiwanese lawyer. That example that you provide before, that actually if you have the same person holding the key, and overlapping -- in overlapping organizations, would that be the good management or not?

And another problem from that example is even if that person holds the key -- so that person will legally access -- be able to access the key when that person acts on behalf of one -- of one organization, right? But on behalf of the other organization, that access would be considered illegal. And now even within the GDPR, we have to debate, should we consider the illegal access as one element to assess the risk of data -- of encryption -- of decryption as well. And that leads to the possibility of future development of decryption. This is also the principle established by the data protection directive of EU that we can't just consider the correct one but also the future.

Some propose that actually if you produce an encrypted set of data and we intend to use this for ten years but in the ninth year that the encryption technology has developed, so in that ninth year, within the first nine years, data was completely anonymous. So it can be considered as anonymous data. So it doesn't fall into the protection -- the data protection principle, the data protection law, but after the data was developed, it has to be reviewed and as personal data, okay?

So I would like to quote this one from another law professor that due to the technical development, the encrypted data, it would only be anonymous for a certain period of time as I used in my example before and that's the level of encryption has to be checked constantly by the controller, not by the -- not by the law enforcer, but by the controller themselves and not only when the controller possessed data for the first time. So it must be reviewed and this can be considered as quite a big burden for, you know --

>> ARTHIT SURIYAWONGKUL: The controller?

>> THITIRAT THIPSAMRITKUL: Yeah, the controller.

>> ARTHIT SURIYAWONGKUL: Do you have anything, Wong? The panel going to ask each other now.

(Laughter).

>> JUGGAPONG NATWICHAI: We have been trying to protect something, the data privacy, right? It seems to me that the approach around the world is really have general law and then the set of practices could be updated by the state or something like that. Is that always the case in Taiwan?

>> ARTHIT SURIYAWONGKUL: Like sunset provision in the states. Like by five or ten years, you should -- we will, yeah.

>>> Wong: Generally you renew the certificate every two or three years, it complies with the standard and even the standard will be revised. So if the standard is revised, you have to check the process again. It's not a law. It's a national standard.

And secondly, you know, most of the data usage is only project based. So the project may only be two or three years and after the project ends, then the data should be removed, totally removed.

So for the question, it should not be too -- right now, it's not too prevalent for data encryption. In response to the audience, how can we do that? Actually, what we do is still far behind the expectation of the civil society, and the industry still says we should move. But right now -- we can have a national standard and we have some kind of mechanism. I think they are two important driving forces -- the first one is that two years ago, we have a very top down policy, that we have -- it comes from Google and he realized the value of data. So he just set in course a ministry how to set up standard and how to assign the duty to a ministry to do that. And so it's across ministry cooperation to attribute.

And secondly, in the team that -- actually the team included technical and also the legal staff and also include other domains receive expert. For example, when we are dealing with the financial data, the text data, then we should have some financial expert, and we are dealing with the healthcare data which includes an expert from the medical industry.

The course domain, cooperation is very important for data.

>> ARTHIT SURIYAWONGKUL: I just want to check with the remote participants. There are any questions? Two questions from here, and probably we have to wrap up.

You first.

>> PARTICIPANT: Yes, I'm from Sri Lanka but originally from Norway. I have a question. Is there anything that could prevent a government agency from using big data to monitor individual students' behavior to oppose or prevent terrorism or even political opposition. I have an article here from Saudi Arabia, big data analytics are behavior monitoring the students.

>> PARTICIPANT: I think as far as the philippines goes, I think I can answer that.

(Laughter).

National privacy commission of the Philippines.

In our jurisdiction, if whatever your governmental purpose is, there's an exception for that government purpose. The moment you stray from that purpose, then you are not covered under the exception provision and you have no demonstrate consent. If you don't have consent and you still process it for that unauthorized purpose, then there's criminal liability.

We talked earlier about the difference between the absolute approach, when it comes to personal data, and the relative approach, and it seems that in the personal approach, then the distinction on whether encrypted data is personal data becomes irrelevant, because then -- but do you see the trend in other jurisdictions moving towards the relative approach or because the ECJ decision came out last year, do you see benches moving toward an absolute approach?

>> THITIRAT THIPSAMRITKUL: So I think for ECJ and ECJ is quite clear that they took the absolute approach, but that before the GDPR became effective. And if we analyze the GDPR -- ours is very, very complicated, we can see both absolute approach and relative approach there.

I think -- I'm not quite sure. I do not have the actual answer to, that but if you look at many others jurisdiction, I think many government which those who are quite keen to encourage private sectors to use more data, to have more innovations, they try to facilitate the use of data and try not -- not to drop the encrypted data out of the personal data, but try to make it -- try to make some rules but more, you know, less flexible rules for them, like for -- if we take the Japanese example, so they have kind of the guidelines that what kind of data you must delete if. If you delete it, you can use it freely but it doesn't mean that you are not subject to the data protection law anyway.

So I think maybe we can go back to the ECJ decision in the sense that even know the data protection law has a wide scope, a wide natural scope, both in the trade reality and also by the content, still there are some exceptions and if you can prove that you follow the standards and have the legitimate reason, still you can use the data. It doesn't mean that labeling the data as personal data does not mean you can't use it in more productive way.

>> ARTHIT SURIYAWONGKUL: We have, like, five minutes here. Any remarks, comments, observations? If not, probably I will pass back to the panelists to probably last words before we go back to the closing plenary.

Oh, there's one. I'm sorry.

>> PARTICIPANT: I just have a comment. This year, the privacy commissioner conference, the international one is happening in Hong Kong in September. So maybe for those of you who know, and I think it's the first -- it's really a very big conference of all the international privacy and data commissioners of the world, from the governments and so on. So it may be quite interesting for civil society to observe.

>> ARTHIT SURIYAWONGKUL: Yes.

>> PARTICIPANT: And there's also the university of Hong Kong is having an Asian privacy scholars conference. So a lot of things are

happening if you are interested in this field.

>> ARTHIT SURIYAWONGKUL: That's very important. Apart from IGF, and regional global one, there's news to be discussed about it. The commissioners conference, the scholars and also AIPP, the international association of privacy petitioners. I think in Singapore as well. Yeah, IPP, right?

So we can actually explore pore and have some discussion across the circles, right? So anything as final remarks from our panelists, please?

>> THITIRAT THIPSAMRITKUL: I have another point to make, since we -- okay. So we talk about the general law or the sectoral law there. Another remark here is that this might be one of local concerns because Thailand is one of the countries in this region that could not develop a general law for the data protection, even though we try for more than ten years now. So the problem here is that in the countries where the government could not develop the general law, still many sectors, many industry, they still need some rules for data protection, because the -- actually the data protection standard is enforced from the West, like we see in the EU, the GDPR is trying to extend the scope of application to any data protection happen in any country in the world, if the data subjects, they are European citizen.

So many industry will develop their own rules and then we will have the sectoral rules anyway, like what happened in Thailand some financial sector trying to determine their own rules and then we can see some discriminatory effect happens in many -- in the country. That's one point.

Okay. (Off microphone comments).

The monitoring the students. There, I think that -- we go back to the Taiwanese case again, that when you have this exception in every data privacy law, that if it is required by law, if it's required by the police or regulatory law enforcement, every principle can be accepted.

So that is the flaw in the law itself and those ones who have the power to abuse this, they are the governments. So what might be something to convince the government not to put that kind of back door, or that kind of laws into law is the exception -- is the example of what we had in the US, EU -- no, the safe harbor principle -- the safe harbor -- the safe harbor agreement that has been, you know, considered as illegal under the EU law by the European court of justice, which makes the American private sector efforts to achieve the data protection level become meaningless, just because of the government-backed, the government effort to abuse those flaws in the law.

And then this is the, you know the power play between the American and the European, but what we can benefit from here is that telling our government that any way you industries must deal -- must sell and buy things from the European customers, in that sense if you put that kind of law or that kind of back door in your own law, what happens is you will be considered as the country's operating with the law data protection level and then the national law or the guylines that you try to issue would be meaningless. So that might be one thing.

Okay. Yeah.

>> JUGGAPONG NATWICHAI: One last thing. Statistic data is not safe. I have to say that. Because we are going to publish an article to say that even under the statistic data framework, you can get all the maximum and minimum, you still can track back to the origin of it.

>> ARTHIT SURIYAWONGKUL: Thanks, Wong, Juggapong, and Thitirat and the remote participant as well.

That's it for this. If you would hike to reflect back to maybe the synthesis or the document regarding the discussion in this room, you are welcome. Go to the website.

Thanks.