Tokyo Symposium on Genomic Med...ne, Therapeutics and Health007

Wed, Apr 10, 2024 12:17PM • 22:03

**SUMMARY KEYWORDS**

genomic data, data, research, safeguard, information, genomic, korea, measures, privacy, stylization, scientific research, identify, genetic, standardization, purposes, guidelines, scientists, treated, talks, nutshell

00:00

give you a surprise. Just a little bit of my background. I'm by your side here. I used to be a decision went to medical school they have to perform your decades ago has since been focusing on legal issues that arise at the intersection of law and life science naturally. This has been God privacy has been one of those topics that really made my heart race inside. focusing a lot on this journey for the past few years there Korean privacy law and I'm saying just skip this here sighs in the interest of time, so stone eyes data are treated as personal data. The approach is the same as the US GDPR under the GDPR pseudonymise data are TDs personal data or personal information. They're not treated as de identified data. But the approach is different from one jurisdiction to another. So Korea, just like I said, Korea and the EU are almost identical. They both treat the de pseudonymise data as personal information, or as the US especially in the research context, doesn't treat synonymous data as personal data to find information under common goals. When you start my data, it starts to be treated as non identifying information. So it's different for each jurisdiction and I think Japan is the is unique in treating treating streamlines data as either personally identifiable information or not personally identifiable information depending on the level of standardization, if I understand correctly. So again, we're different from one jurisdiction to another. And because of this when the US us, we're trying to negotiate their privacy contract after this turns to litigation, this insurance to litigation struck down with previous practices shield between the US and US the new contract is so bad for them. Name is the EU US data privacy framework that can change deceptive or recently mouth trees pseudonymised data as personal information in the research center in the research context. So that was one of the biggest differences between the precuts the Privacy Shield and Dow, the effective EU US state apprenticeship program that all is based on how those students jurisdictions treats diverse data differently. Okay, moving on. This is actually the looking at the core of my talk stay under paid countries. There are always scientific research exemptions or exceptions. So as a rule, you need to consent from the data subject to use that person's data right. But if you're doing research, there exceptions which enable you to use the Christmas data without versus explicit consent or consent, creates a shot under Article 28 dash two previous practice law and I'm just going to read it out. That personal information controller may process stone mice information without consent of data subjects, for statistical purposes, society research purposes and archiving purposes in the public interest etc. Please sounds you will notice that this sounds very familiar to the GDPR under the GDPR or anything nine which is the lawful grounds on which a data controller can't process personal data. There have been multiple rounds which are two years before the very bottom is INFJ which says the data controller will get a lot to process personal data if the Processing is necessary for archiving purposes in the public interest scientific for historical research purposes and so on and so forth, in accordance with Article 89. One, so it is allowed under certain circumstances for those purposes, which include scientific research. Again, there are so to allow ourselves this lady is obviously Korea slaw was modeled after the GDPR and if we could have our lady by her one I'm going to bring up the test because probably they're too small to see either otherwise screen. So the data controller when relying on the scientific research assumption was put in place safeguards. That will ensure that technical and organizational measures are in place in particular, in order to ensure respect for the principle of data minimization. Those measures may include may include stylization provided that those purposes can be fulfilled in that bag. Okay, a lot of abstract concepts apply around here, but in a nutshell, you have to implement some safety measures safeguarding measures, which may include sterilization to minimize personal identifying information, but he says May so it's not a prerequisite is recommended highly recommend that you do to the extent that you do secrecy your research after stylization, so, you know, it's backwards. You don't even have to sign eyes if he keeps you from pursuing your research was. So the prerequisite is a little different entry and salary information is a prerequisite for essentially research accession whereas under the GDPR it is not per se a prerequisite so other Korean law although you will start to be modeled after the GDPR became a huge issue what proper stylization really means and what it entails. Because scientists obviously want to rely on sexy research, essentially without having to obtain consent for the use of cases with consent from the data subjects. So it became a huge issue. And I'm sure you all have you know that there's no one right way of summarizing but there are definitely Bromley south and standardizing and one example goes all the way back to the late 1990s When the Massachusetts School Insurance Commission case, the state of Massachusetts made available some research data was actually top records for hospitals and they sit on rise to help workers by removing the name addresses because we begin with the zip code for a state and the gender and a very clever PhD student back then she has become a professor, Dr. Sweetie. She was able to get a copy of the local voter list, which was also publicly available shifts to answer together and was able to identify a lot of people whose covers work, including the data set and short actually able to find 100% Massachusetts governor will be well became a very famous case in the area of price law. So this was one of the premise behind the Privacy Rule under the HIPAA was you many of you may be aware. So there are always up to insolvency standardization. But what is the right way of doing the standardization? In a nutshell, there's no one right way of doing stylization and there are many principles that you can rely on. But there's no one consensus of stylization. That's the bottom line and one of the principles that can remind you is contained ISOs similar to 5237. It deals with us with our customers from might encounter information. But here if you were doing really talks about things in very broad abstract terms, so doesn't give you a very concrete idea of how to implement civilization, especially when it comes to genetic, genetic or genomic data. Skip just a few slides. But if you're interested I encourage you to need a review. So in trivia, like I said, it became a big issue within the scientific community what standardization means, especially for genetic electronic data, so the government had to come up with some kind of guideline itself initially back in 2020. So he provided guidelines on how to stop disparities and health data, both structured and non structured or even images. But the funny thing is what they said about the genomic data is that they weren't quite sure what the proper way of stuff

09:39

that was said they said, don't do it yet. We don't know how to start my electronic data or genetic data is really tricky, and it's gonna be very hard to find the right balance between like the auto scaling, or learning journey Data and Safety, keeping the genomic for genetic data valuable for research. So that's what we said effective 2020 Don't do it yet. Again, the scientists were not happy with the guidelines. So the government had to come up with another set of clarifying that they did earlier this year, a couple months ago. This time this time. Well, the structured text data images, the guidelines were already there back in 2020. Also free text at this time, what was new was that they provide these guidelines, stuffs anonymizing genomic data, and these are really small text. I'm sorry, I wish I use too fast. But I'll just read the first bullet point, I think we should say the most relevant so sequence informations in this case. So again, DNA or the same files were petitioning information in germline cells and short term reduce which increase the risk of re identification should be deleted or replaced with random sequences in large or small unless they are directly related to the research. Okay, that's the bottom line. The second bullet point is that it deals with metadata. So you have to summarize the meta data to and on the right hand side, the first bullet point, I talked about the first TQ files as a guideline is saying, you cannot see it on my spastically files because it's a raw file and there's no way of summarizing FTP files. So 32 files cannot be used under this entity research exception. You have to obtain consent from the data subject. That's the bottom line. So also the bureaucrats that came up with this idea obviously the government assembler to club scientists, they literally dramatics people to come up with this conclusion. That still even though this was the recommendation for a group of highly qualified scholars and scientists, you know, the opinions are divided over this approach, then this is doing enough or maybe dismissing too much. That everybody has their own opinion service, you have your own opinion. I like to hear that you any session. Now. This is the law of the ground in Korea for standardizing and genomic data to be able to benefit from the secretary researching session. I'm going to skip over my previous criticism. I'm going to optimize my critique on this. So I don't have taken backgrounds obviously why that is so my criticism has more to do with other factors of safeguarding privacy. So we identify validity of genomic information or genetic information. We already talked about this during the presentation, and also probably script presentation, I think the RE identifiability of genomic data sets are largely three sectors. So well the data are set on ways to get what kind of data sets are available for the voting the user or intuitive intruder to be identified. So what kind of data sets are available to cross reference against that also, we tend to the user and create, I think, to the guidelines to the very long talks taking place in Korea have placed too much purpose at the first factor of this free identifiability function I think so there were too fixated on how to technically Siddhanam eyes genomic data appropriately, whereas other factors also come into play. For example, cross references, the data that are available for somebody to cross reference. Some genetic or genomic data yes so if this was in the US, for example, we heard many cases where a piece of general information was used using publicly available genomic databases such as Jed match, to find to solve mystery cases or first game, the famous Golden State killer case. I think there have been many, many similar cases using the same rationale since so in the US, it is not impossible to re identify someone purely with somebody's charity. Or shall be permission. It is possible. Probably you did it for every one of them, but you probably you'll be able to pull off doing it. Yeah, to find the cable person for a large group of people in places like the US and some other jurisdictions. Where TTC genetic testing is widely available and larger than people had previously. Maybe not in Korea, maybe that is juristische itself, the size of a publicly available genomic database comes into play. With the split with this same sample safety that I just mentioned, as well as the intent of the user. So if the context of scientific research exception of the recipient of standardized genetic or genomic data are scientists so we kind of know the integrity and they also connect to the research institutions so we can have not just that person, the recipient, but also his or her institution to all of California. We have many layers of safeguards into institutionally so we can use that as safeguard to protect the privacy of God are synonymous to genomic or genetic data even further. So there are many other factors that you can we can utilize to safeguard the first set of genetic data in addition to pseudonymisation that these other factors didn't come in to consideration enough during during the talks in Korea, when they were discussing how to set up the rules for standardization of information so that that I think that portion was lacking. And I feel like we should is the Korean regulator should start discussing further what these other factors mean and how they can be harnessed to increase or decrease eukaryotic viability of genomic information during scientific research. So on the trail, we'll get off the trail by by requiring central researchers to place their data storage that has a very good track record does safeguard their audit trails, that's another way so in a nutshell, in a nutshell, retrieval and destruction feedback are useless. Also, one measure of doing it so even though shell I think we have to tick off both technical and organizational measures in building a safe environment for sanctuary research purposes where data cannot be perfectly to identify without sacrificing the granularity or value the second CT data, research data, and I think genomic data or genetic data is the prime example. I cannot think of any other example that is harder to summarize or to identify without sacrificing the quality of the data as a scientific data. It is the current example where we are facing that kind of dilemma. So in protecting the privacy of genomic data, we have to think more about not just the technical measures, but organizational measures. What do these terms mean technical and organizational measures that include a ton of literature in the private civil domain, and also a GDPR article 25 If there's too many of these terms, so if you're not familiar, please with the literature, there's a ton of material. What I can say is that the Koreas approach so far, trying to find the best way of summarizing genomic data was focusing too much on the technical aspects whereas the organizational measures that we can use, such as those that are equation that I've gone over in the past Meski sites, were maybe relatively productive. So I think it's time for Korea and press other jurisdictions to put these two measures in parallel. Because technically pseudonymized situated as so difficult, maybe we should play some more focus more attention on the organizational measures. Okay. That said, I think that is the time to pay attention and I'll take your questions.

19:30

Hopefully we have time for some questions. Yeah.

19:34

So, really nice presentation. I actually completely agree with you that the question of protecting data has become too technical and I think it's interesting also out since the GDPR, has been adopted, a lot of the debate on genomics and the ethical issue has been reduced to really looking at privacy and other de identified data and very technical things. So I really commend you from saying, you know, there, there are other things to look into, as you're looking at protecting data. And you mentioned the environment of Korea, South Korea, which is interesting. I know that direct to consumer companies. For example, are not yet as popular as they are in the US. My only word of caution on this current perspective, is that things will change. I think that information just as it has gone 100 countries around the world will tend to spread and if you have that kind of report that will need to be revisited. And also in the context of researchers I completely agree that we should think of researchers as very responsible people and also institutional safeguards as already providing like a layer of protection. But I think we also must not forget that sometimes, you need to be very responsible and well intended yourself. But if there's a breach, another person will have access to the data and that person might not be as well intentioned. So I think that has to be taken into consideration is an exception for doing sessions for research purpose,

21:23

their thirst for your own accounts. I absolutely agree with the two points that's over the objection. I should ask your first point. There was a slide actually that this will change the environment which will change over time, so the issue has to be revisited every so often.

21:41

I absolutely agree. Excellent. All right. No further questions. I want to thank all the speakers for an excellent session and for staying in time. Thank you very much.