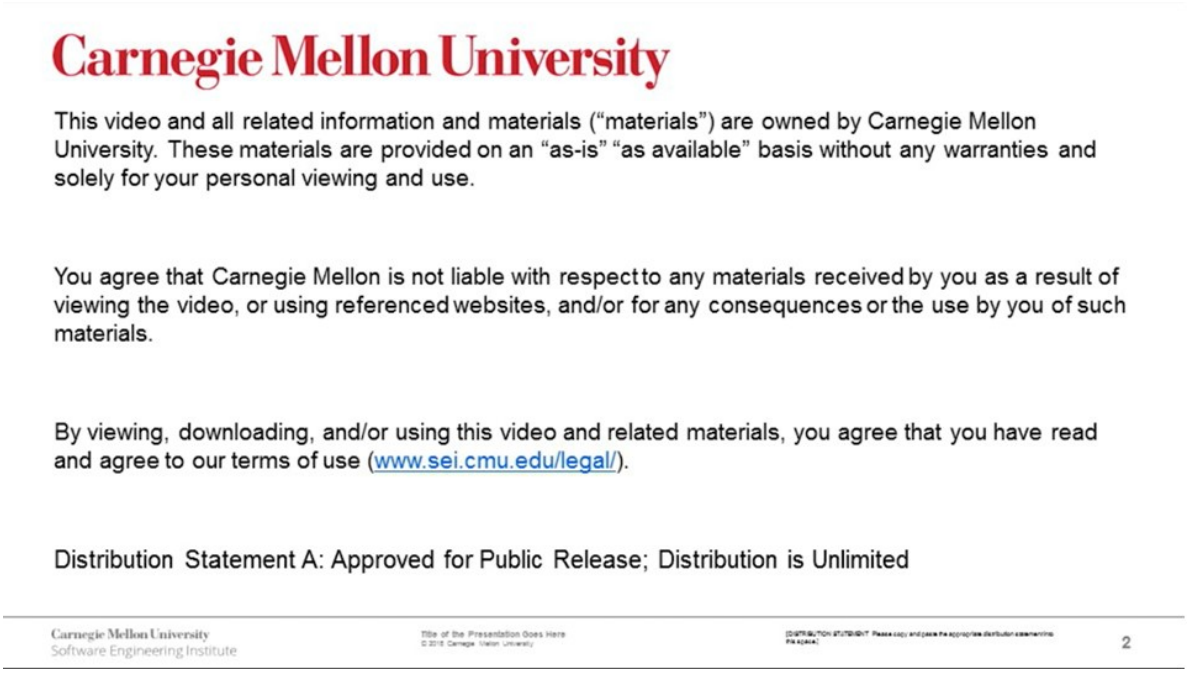
Digital Footprints: Managing Privacy and Security

Table of Contents

Copyright 2018. Carnegie Mellon University



Carnegie Mellon University

This video and all related information and materials ("materials") are owned by Carnegie Mellon University. These materials are provided on an "as-is" "as available" basis without any warranties and solely for your personal viewing and use.

You agree that Carnegie Mellon is not liable with respect to any materials received by you as a result of viewing the video, or using referenced websites, and/or for any consequences or the use by you of such materials.

By viewing, downloading, and/or using this video and related materials, you agree that you have read and agree to our terms of use (www.sei.cmu.edu/legal/).

Distribution Statement A: Approved for Public Release; Distribution is Unlimited

Carnegie Mellon University
Software Engineering Institute

Title of the Presentation Goes Here
© 2018 Carnegie Mellon University

[DISTRIBUTION STATEMENT: Please copy and paste the appropriate distribution statement from the options]

2

Copyright 2018. Software Engineering Institute

**002 Speaker: My name is Shane McGraw. I'll be your audience moderator for today's presentation.

And I'd like to thank you for attending. We want to make today as interactive as possible. So, we will address questions throughout the discussion and again at the end of the discussion. And you can submit those questions to our staff at any time by using the Q and A or chat tabs on the page interface. We're streaming this on YouTube on Ustream as well. So, depending on your platform of choice, use those tabs. And it will come to our event staff.

We also ask that you fill out our survey upon leaving today's event as your feedback is greatly appreciated. And we'll supply a link to that survey in our chat area soon.

Now, I'd like to introduce our speakers for today. Lena Pons holds an MA in computer science from Loyola University, Maryland with a specialization in natural language processing for question answering. Prior to joining SEI, Lena worked at the National Institutes of Health maintaining and improving a text mining system for categorizing research grants by disease category. Welcome, Lena.

Next, we have Matthew Butkovic. And Matt is the technical manager for the cybersecurity assurance group in the CERT division of the SEI. He performs critical infrastructure protection research and develops methods, tools, and techniques for evaluating capabilities and managing risk. Matt, welcome.

And now, we're going to turn it over to Lena to lead our conversation. Lena, all yours.

Speaker: Okay, great. I thought maybe we'd start a little bit talking about sort of metadata and refresh a little bit from part one of the series. And so, to that end, following on some of the interest on metadata with European regulations and the general data privacy rule and also the hearings and Facebook-- there's been a lot of interest in sort of what data are-- is being collected by these sort of social media platforms. How is it being used? Who is it transferred to? And what are the implications of that? How do you keep track and make informed decisions about how you want to manage and control sort of what the access to your-- to your personal data is?

So, to that end, I'd like to talk a little bit about existing Federal Trade Commission authority to regulate and control access to individual, personal metadata that's currently being collected, and then also touch upon some proposed changes that are currently being discussed. So, yeah.

Speaker: Excellent.

Speaker: I think this is certainly a timely topic. And I think there's sort of this general sense that there is this asymmetric understanding that those who are collecting the data typically understand with fine-grained precision sort of what they're looking for and how it's used. Whereas, the

folks that are generating this data, or the owners of this data, often feel ill-informed about how it's being used and collected.

Speaker: Agreed. I mean one of the things that I was thinking about in watching the Facebook hearings and the response to that was there was a lot of focus on Facebook and how you can opt in and out of Facebook's collection and use of your data, but not very much focus at all on secondary users and secondary purchasers. And I really think that Facebook has definitely taken some steps to be more transparent about what the controls that you have access to are, how you can get access and download what data Facebook has currently collected. You can look at some of the advertising categories that Facebook has assigned to you based on the data that they've collected. But what you can't really see is the secondary sales market. And I think that that's a place where the sort of asymmetrical understanding becomes even more pronounced and even more problematic.

Speaker: Yeah, I think that there's a fundamental truth we have to come to grips with. There's a complex ecosystem that exists that isn't going away. We have entire ways of working, entire industries now built on the collection and use of this data. So, we can, and I think we're rightly so, taking steps now to ensure that the custodians of the data, the processors of the data, the owners of

the data all have some understanding of their obligation to protect. But this isn't going away anytime soon. So, I think that in sort of the popular discussion I sometimes hear that this is the start of something profound that kind of ends the way data is collected. No, I don't think that's the case at all. Lena, thoughts about sort of where we are in the evolution of our thinking about personal data?

Speaker: Yeah, so I mean I think, to that end, one way to think about that is that the policy and regulatory definitions that go along with the way data is collected and stored and transferred and used are very much behind the technology curve. And so, a lot of the regulatory context that Federal Trade Commission uses to define sort of what the boundaries are for using and storing and transferring data are based on the assumption that that data is expensive to collect, difficult to find, and difficult to transfer and store for long periods of time. None of those are supported by the current technological context. And so, those assumptions have led to a regulatory regime that assumes that you have to find an investigator to go and do some work to identify and collect and store this information. But the cost of storage has gone down so much. And many, many-- there's many, many sort of like big data-- these machine learning processing algorithms that are just sucking in huge amounts of information. And also, it's stored and collected by many sources. And so, then there's sort of the multipoint

provenance issue. So, you can-- any advertising company could plausibly say that they got any piece of information about an individual from multiple places. And so, dealing with that regulatory challenge is also very much behind the technology.

Speaker: Yeah, I think there's this perennial struggle where technology outpaces law and regulation. And this is just another variation on that theme. When I first had to think about this, think about privacy and data in a serious way, it was in the late 1990s in the banking context with Gramm-Leach-Bliley. And I think back to the challenges we had then, and they seem really quaint compared to the challenges we have now. Chief among them, as you mentioned, is that we now have a borderless approach here. So, you're not looking at data collection or data generation in specific geographies, but rather cross-geography collection. In a sense, with storage technology evolving, geographic location is trivial or non-consequential technically. But regulation and the implications for you, as a collector or processor of that data, are profound, despite the fact that borders don't really matter technology-wise anymore.

Speaker: Yeah, that's a great point. That's a great point about borderless because when you're thinking about trying to set up a regulatory context, then the challenges of identifying and tying down a particular piece of data or a data collecting entity to a

location is very important to what the regulatory context is but isn't very important to most of the technology related to collecting and using that data. And so, then you also have the problem of the companies are involved have very little accountability on that end. They're operating in the environment that the technology supports. And they're trying to provide services and give people things that they want and trying to navigate the challenge of making a personal decision about what services you want to-- are important to you and you want to participate in and what the breakdown of removing your data-- so, you know, there's the challenge of if you opt out from a service, or if you opt out from a data collection, then some services are no longer available to you.

Speaker: Yeah, so let's, if we could, let's explore the idea of giving the people what they want. So, legislation, regulation is a response to the populace-- it should be a response to the populace saying these are the guideposts with which we want to see something-- how we want something conducted in our society, right? It wasn't that long ago that I heard this description of the post-privacy era that people were fine living in a transparent way. Facebook was the example that I heard time and time again. I also heard it tied to generational differences, which I don't entirely think are valid. But we can explore that. But it seems to me that we are

at this watershed moment where some of those assumptions about how transparent folks want to be with their data are being challenged. And certainly, those in marketing and data collection very much would like to see that paradigm because economically, it's the most viable. Thoughts on sort of do people still have an appetite for privacy? The current round of regulation that we see, both in Europe and things emerging in the U.S., are we at some inflection point in this topic?

Speaker: I don't know about an inflection point, but there's definitely- I think that the idea of defining a post-privacy society is a little misleading. And also, I think that one of the things that you saw in the Facebook hearings and the response to the Facebook hearings was that people may have been behaving in a way where they were-- they had provided these social media platforms with implicit consent to use the data but didn't really understand what the implications of its use were. And I think that was, again, something we touched on in part one about not having full understanding of what the implications are of like combining from data different sources or what you can actually discover from metadata. And so, I think that Facebook was often talking about a post-privacy society in the context of people love the service that we're providing. And openness is great. We have sort of this capacity to connect people. And it's all about being open and transparent. But the people who-

but there's two challenges to that. The people who don't participate in Facebook are invisible to that discussion. If you look at the implications of ceding some of your personal privacy and your private life to a company as too onerous and don't participate in that, then you're no longer in the sort of post-privacy discussion.

The other problem is the issue of sort of making a truly informed decision about what kind of use of your data you're actually okay with. And I think that there's not enough clear information for many people to make truly informed decisions. And so, I think that that's a place where improved support from Federal Trade Commission to provide and communicate to people information about what the implications are. That's a great place, I think, for a regulatory program to improve the ability of people to make real choices about what level of privacy they're willing to cede in exchange for services.

Speaker: Sure, I think this is a really key there, which is informed consent, right, that you're giving those-- you're giving people the ability to opt out of the collection of that data. The flip side being that folks sometimes feel deceived. They are maybe in denial about that economic transaction we described. Whereas, the value you represent as a user is in your data. If you want to access a social media platform or storage platform, implicitly, you need to be comfortable sharing something with

that provider. Otherwise, there's no economic benefit for the provider.

Speaker: There's no economic benefit for the provider, and also I mean there's lots of types of-- geolocation data, for example, is a place where it's very, very difficult to make informed decisions about how you want that data to be used because if you say, "Well, I don't want any geolocation data to be stored or tracked for my individual person," well then that creates some limitations to services that can be provided, especially services about proximate things that you might be interested in finding. So, if you say, "I don't want my cellphone to ever record geolocation data about where I am," then you're also saying, "I don't want Google maps to tell me that there's a coffee shop nearby that I might want to visit when I'm in a place that I'm unfamiliar with." And so, that tradeoff is actually a really big tradeoff. And being responsible for keeping track of I want my geolocation data tracked for this purpose only, or in this context only, or at this time, that's a pretty challenging-- that's a pretty challenging thing for the provider to give differentially to individual people. And so, I mean that's also a big question is can these services-- can these companies even provide that level of granular control that an individual might want. And how do you sort of come to a balance that is- that gets everybody the most of what they want, the most preservation of service, and the most granular

control, but also like not putting the burden on Facebook to have to say okay, well we gave you a huge array of checkboxes. And you want through and said, "I'm okay." That's a pretty difficult task.

Speaker: Yeah, there's certainly a question of the technical feasibility, enforceability. I know in this discussion, we'll touch on GDPR and some potential pending changes here in the U.S. One of the things I would sort of remind us that although we've seen a significant fundamental change in the way privacy is being approached in Europe via GDPR, and in the U.S. arguably there's a new focus, let's not forget we're just slices of the world. So, I don't know if we'll have time today, but I think about how differently this is viewed in Asia, for instance, how the Chinese market is very different than the European market. And I just, as we described this borderless world of data, it's important to remember that just because GDPR is the law of the land in Europe, maybe with good reason, and there could be changes coming to the U.S., we're playing in this larger global context, which will then have economic implications, which is if you are a consumer that is untouchable because we cannot conveniently collect data in your market, perhaps you'll be exempted or ignored. Whereas, other large markets may be further exploited for their data because it is easier and without the regulatory burden.

Speaker: Yeah, I think we had

chatted about that before about the concept of does data become more valuable in a market that has sort of more accessibility to that kind of data and what will be the sort of business implications for places where there's more control. My perspective is really trying to navigate the technical challenges of giving people the most capacity to make a decision about how they want to interact with data collections. And one thing that I want to talk about, in terms of technical challenges, was the issue of if you change your mind. That's a very, very difficult complex process.

Speaker: It certainly is, yeah.

Speaker: So, all of the people who went and changed their Facebook privacy settings after the Facebook hearings, they can really only control the data going forward. And everything that was collected up to the point where they changed their privacy settings is now part of a record and part of a product, an advertising product, that it would be technically nearly impossible to claw back. And so, I think that that's a place that's also very difficult is that how do you make a decision about today, and then what happens when, in the future, you have a different relationship to your willingness to have that private data shared because something in your life changed or something in your perspective changed. And how do you deal with that?

Speaker: Yeah, there's a lot of

interesting examples of government struggling with this issue, so the right to be forgotten in the European context or even in the U.S. context where things that are public record, with good reason, there is law. There is precedent. These things must remain available. So, I think that's another interesting facet of this, which is the old adage the Internet never forgets. It's true. Does a citizen have a right to request that a private entity or a government somehow obscure or make unavailable a record? And I think that this is where the discussion, why I'm mentioning this is you described this through the difficulty of toggling on and off these technical means. I think that's a good example where if we have a record of real estate transactions, and in the future decide you don't want your home sale to be public record, well okay, what does that mean? What's the greater implication using a very basic example?

Speaker: Yeah, I mean that's an interesting point too because like with the home sale transaction, there's actually another-- there's other potential consequences to that to the future purchaser of that house. And so, then there is kind of the delicate balance of where does your right to privacy end.

Speaker: The famous story, it's not a pretty story, but I think it's illustrative of what we're discussing. A family who was grieving a loved one lost in an accident, so a terrible auto accident. And the pictures of

this auto accident became available publicly. So, they were being displayed on the Internet as a curiosity. Look at how this car came apart. Well, it wasn't just the car. This loved one had been killed. And the family had to go to court in California and sue the state to find a way to prevent this public record from making the rounds on the Internet as a curiosity. I think it was a really interesting sort of ethical and moral dimension to this as well. So, it's not just-- it's just not collecting data so that something can be sold to you. It's also I think ensuring that these things that are sensitive or disturbing related to you or your loved ones are used only in an appropriate way.

Speaker: Yeah, yeah, that's a good point. There is sort of a challenge of when you think about advertisers, the way they frequently frame the way you're using their data is that you know they're helping you to make a better decision. They're making you aware of products and opportunities that you might not otherwise be aware of. But the reality is that lots of advertising products are used for purposes that are not necessarily aboveboard. And so, this is a place where I think that the potential for changing FTC regulations could be really useful. So, I've seen there is a legislative proposal on the table which would allow you to obtain, from social media companies that you're interested in, a list of all of the advertising-- like secondary sales

entities that have access to your data. On the one hand, I think that that's great. Like in terms of transparent information, I think that that's really-- has potential to be very helpful. In reality, the ecosystem of advertisers, the secondary sales advertisers, is huge. And so, trying to navigate that list of companies that you never agreed to do business with and you don't know anything about their business practices is very difficult because you can research lots of different companies. But if you get a list of hundreds of companies and you don't know anything about their business practices or what kinds of advertising products they're actually selling, it's very difficult to say, "Well now, do I go through and say I want to opt out of all secondary sales? Do I want to opt out of these secondary sales?" And so, I would want to know that my data was being used on secondary sales market with a company that sells a product that's used for tracking your partner and for potentially malicious purposes. And so, I mean I think that that's of great concern is being able to really understand Facebook as a company that you can think about, understand, research their business practices. It's really easy to find out information about like what Facebook does and says. But these secondary sales advertising companies are much, much harder to get that information.

Speaker: So, you just hit on a question I was just going to ask. And chime in again, but Joe chimed in and said, "There's enormous value in

this data to improve our lives and create economic value. How do we get that transparency and control?" Is it government regulation?

Speaker: Yeah, I think that making it explicit and making it required that lists and awareness of what those companies are is great. And I mean it really does say like you can say oh, I do business with these companies. I know who they are. Or I definitely want to know about marketing opportunities for products that I'm potentially interested in. That part is great. It's really just I think the hidden labor of trying to navigate the morass of what are all these companies, and what are they business practices, and are they companies that I really want to do business with.

Speaker: I think in that sense, this is a variation on a classic supply chain problem. So, there's the third party/fourth party problem. I find that many organizations can't classify and understand their own internal data, let alone the data they're sharing or selling to third parties. So, I think that to kind of make sure we're addressing all the stakeholders potentially in this webinar, it's not just, in my mind, there's many facets to this. Those who collect data and resell it, they aren't necessarily doing this-- I would argue they're not doing it out of some malicious intent, but rather with some economic motivation. Understand, though, the potential liability, especially going forward with the rise of new

legislation related to that practice. Get a handle on the data you possess. Get a handle on the data you collect. Only collect the things you need. I think these are all points of advice that I would offer those that are in the business of collecting data.

Speaker: Yeah, I mean that really is just good data hygiene practice, right? And I was thinking a little bit about the issue of where this interacts a little bit with cybersecurity, a little bit with my regular work mission. But there is the potential for pieces of information that can be gleaned from social media outlets to actually be a cybersecurity risk. And I think that that's another place where really understanding the risks is very difficult. I mean it's difficult even for people who have a good understanding of a lot of these pieces of the puzzle, of what pieces of information are valuable, how they can be used and misused and what the potential for deidentifying hypothetically anonymous data. Even if you have a pretty good handle on all of those forces, it's still hard to think about what is the threat model, what is the risk that any individual person is assuming by making that data available.

Speaker: Right. I think that sometimes the practices that are legitimate, economically necessary, feel somehow malicious to those that have submitted the data. So, if I don't know, for instance, that a

party's collecting my data to sell to another party to then sell me something of solicit me for something that seems sensitive or embarrassing or what have you, that may feel malicious, when in fact, it's just part of the economic model. As one of the participants in the webinar said, there's vast economic potential in this data, which Shane, I don't know if there's other questions in the queue, but I was going to offer a question to Lena, which is when I think about GDPR and I think about some of the pending legislation you described, is there a risk that we're overcorrecting in all of this?

Speaker: I think that it will be hard to know for GDPR until we have some experience of how it's been implemented. And I have some questions myself about what the technical feasibility is of actually like complying with the letter of the GDPR rules will be. I think that with the legislative proposals that are on the table in the United States, I think that there is a decent balance of at least structuring it to give FTC the right kind of authority to make good decisions about how to set that balance and say what we really want is to not constrain business's ability to use the economic value that's associated with collecting that data and really capitalize on that, but at the same time, saying can we give citizens their individual rights to determine how they want to participate in that market.

And so, that actually segues into an

example I wanted to mention which was the pretty famous example of Target. And so, Target got kind of burned because they were using metadata and machine learning to predict, based on previous purchases, if a woman had become pregnant and then send her coupons related to things that you might need after you've had a baby. And if you're a happy, healthy pregnant woman, and you get coupons from Target for baby stuff, maybe that's great. And maybe you're a woman who was pregnant and had a miscarriage, and then that becomes a very painful thing for this company to be engaging in. And so, they, in response to that, stopped the practice of sending coupons for that particular purpose. But they still do infer things about what, based on products that you've bought, things that you are likely to buy or might want and send people speculative coupons. And so, that's kind of a touchy-- that's kind of like a touchy place because it's like they want to say oh, here's a service that we can provide. And what could go wrong? And then, of course, in reality what could go wrong is the oops moment of yeah, we should have thought of that.

Speaker: Yeah, there's a just because we can, doesn't mean we should sort of decision that has to be made with this. So, I think that's part of the struggle as well is that we-- the data collection methods are relatively new. The secondary markets for data are relatively new.

So, I think that we're getting our heads around what works and doesn't work. At the end of the day, if those sales tactics are effective, then they'll continue. If they're not, right? So, the efficacy of that is in question. I say that, but I've got a counterexample, which is in the U.S. the do not call list. These unsolicited telemarketing calls are I think universally reviled. I don't know anyone that welcomes them. I don't know anyone that's ever purchased anything based on one of these calls. But they persist. So, we do not call list, and we still get these calls. So, and they must be effective because it goes on. I think of spam email. It's the same thing, right? This must work at some small percentage because someone deems it economically viable to keep doing this. So, I do worry if there's a long tail to this. Even if we have FTC regulation or GDPR certainly is here, how readily will it be enforced. I mean in theory, the do not call list should give us legal recourse to sue in a substantial way those who bother us. But the truth is that's not happening.

Speaker: Yeah, so the do not call list is actually a great example of regulatory programs not keeping pace with technology. And so, the reason that calls persist independent of the do not call list has a lot to do with the move from traditional wired telephones to voice over IP. And so, basically, the calls that you get have a number that appears as a domestic U.S. number, but the call actually

originates outside of the United States. Well, FTC doesn't have the ability to enforce and stop those calls that a being originated outside the country. And so, then that is really technological challenge. So, I mean they are very-- Federal Trade Commission is very, very aware of the fact that the do not call is not providing citizens with the outcome that they want, which is to not get these spam phone calls. But how you address the technical challenge of sort of borderless telephones is a real issue, is a real issue, is a real problem.

Speaker: And it seems to me that that's an excellent example of the technology problem, well I guess rather the regulatory problem, which is the technology is always one step ahead. So, FTC did this in good faith not anticipating that voice over IP would change the game. The end result is the consumer feels like the FTC might not be doing enough to help them.

Speaker: Yeah. Yeah. And so, I think that that's why being really cognizant and careful about how you frame what the legal authority we would give to FTC in order to give people more sort of informed control and informed consent about how their data is being used. It really has to give enough leeway for FTC to build regulations that are changing with the technological frontier. And I mean and it's very difficult because the technological frontier is moving very fast, and especially in sort of this area of when you're dealing with

very large data and modelling. I mean it's hard even for experts to keep up with the frontier of what's happening.

Speaker: Yeah, there's a temptation to fight the last war. So, I'll offer this quick story. My wife teaches high school. And every year, they do a survey of the students regarding social media. And in the last few years, there's been this pronounced distaste for Facebook, and this year by students described as a lame place where people show off their kids. What we see-- and Snapchat's the big winner. Instagram's the winner. Will we see A, a backlash against the social media providers that are collecting data in a substantial way like Facebook? And will we see an organic just sort of gravitation away from those sorts of services based on the preference of the consumer that basically renders irrelevant some of what we're discussing because that data isn't collected by some of these services.

Speaker: I think it's possible, but I think that there is such a strong push-- when Facebook started, they were not primarily a source of information for advertisers. They started as here, we're a social media platform. We provide the ability for people to find their friends and share information about their lives and share news. And then you see, as that company grows and develops, they're looking for what is the way that we-- how do we pay for this service that we're providing to people for free, and how do we make this a

viable business. And then you see an advertising market grow around that. So, the advertising market's not going anywhere.

Speaker: But does the nature of the advertising change is my question.

Speaker: In what way?

Speaker: So, I'm thinking-- and again, this is somewhat-- we're still in the same space here. But I'm just thinking that if I'm the next new social media platform, and I'm contending with GDPR and new regulation from the FTC, maybe I change my practices. And targeted marketing isn't my goal, which runs counter to what I've just described, which is targeted marketing must work because we still do it despite it being an annoyance with the phone calls and spam emails. So, I'm not sure where this is all going. Collectively, we don't know. But I think something to consider is do the economics of this change and obviate some of the need for the things we're discussing today.

Speaker: Yeah, I mean I guess that that's a question of are we less sensitive to privacy as a society because I think that advertisers are going to do what's effective. They're going to use methods that are going to result in what's effective. And if targeted marketing goes a different direction and you see more of, like you said, on Snapchat. Like Snapchat, you see a lot of sort of like sponsored content. And there's a lot

of mixed in with your regular feed, there's an ad that's built in and served to you based on some stuff that they're inferring about you based on what they know. And so, does that come all in platform from Snapchat and your followers and people you follow? And is it all within that ecosystem? Or does it come from secondary advertisers? In the case of Snapchat, I don't know specifically if that's the practice or not. But you could see it running both ways. Like Facebook definitely serves ads that come from within their own targeting mechanism and also come from secondary market. So, they've sold data from-- collected by Facebook to the secondary market. And then the secondary market serves ads to Facebook from analysis that they've done based on agglomerating data from multiple sources. So, I mean there is sort of a balance there. So, does targeted advertising transfer more into sponsored content or something that look like it's built into your followers? And is that less invasive of individual privacy? That's sort of an open question.

Speaker: I think there's another facet of this for those who possess data and may be sharing it. I've heard in HR circles that using social media as a way to determine the suitability of a candidate is a hot topic. And I think that we're probably just on the edge of case law and regulation that allows or disallows your potential employer to look at your Facebook profile or Instagram

or Snapchat and say you're engaged in behaviors we would not want associated with our organization or brand. So, any thoughts about how that factors in, the sort of potential to exclude folks from jobs or from membership based on social media?

Speaker: Yeah, I mean that's a tricky one because that has-- that again has sort of-- you're balancing two interests, right? On the one hand, an employer, if they are accessing publicly accessible information about a potential candidate, and that information leads them to feel like that candidate is not suitable for employment, it doesn't feel exactly fair to say oh well, yeah this publicly available information that this individual chose to put in their public social media presence and had a reasonable expectation that they could be identified, that's a challenging point. Now, on the other hand, I'm a private citizen. And things that I did up to this point, they reflect on me. But you have more visibility into my life than you would reasonably have had before social media. When you were judging a candidate at a different time, you had-- you could-- you had interview and references. And you could basically contact these people and say is there information about this person that you know that you could tell me that should drive my decision about whether this is a good candidate for this job.

Speaker: Yeah, it's a really tricky issue. In the EU, there was a case

where you had a man that had to collect bankruptcy. And this was a matter of public record. And there was a legal action taken on his behalf to be forgotten. So, that yes, you can't erase the fact that he's had a bankruptcy, but can you make it harder to find that information so when you Google his name, actually like literally Google his name, the first thing that doesn't come up is bankruptcy in this jurisdiction in, I believe it was, Spain. It's an interesting question. And I don't know that we have the answer. And I don't know that legislatively in the U.S., we've come anywhere close to kind of striking that balance yet.

Speaker: Yeah, I mean striking the balance is challenging. And there's also sort of the technical challenge of Google has a proprietary algorithm that is uses for deciding where in the search results a result like that will fall. And so, then that's a tricky regulatory question of does an entity like Federal Trade Commission have the right to say to Google, "Well, you need to change the way your search algorithm works such that a result like that will appear in a different position on the search results." And I'm not sure, in terms of right to be forgotten, I'm not sure if the effect that's intended is to push a result like that farther down the result list or actually prevent it from appearing at all. But that is, to my mind, a fairly invasive regulatory practice. And so, I think that that would be a pretty tough sell in the United States, something that has such a strong

effect on a company's specific business practices.

Speaker: Yeah, so to kind of build on that, it's fascinating to see how Twitter's become central to the political discourse in the U.S. And recently, we had a court ruling that said that the president can't block people on Twitter because if you're going to use Twitter as a way to make announcements and convey your thoughts, then you can't stop a citizen from consuming those. So, I think that is a shining example of where politics, regulation, and the public's ability to consume the thoughts of our leaders really with this jumble, we haven't really figured out exactly what's appropriate and what's not appropriate yet.

Speaker: Yeah, I mean that's an interesting point, the blurring of the line between social media and traditional news media. The way a lot of people use Facebook and Twitter is as their primary source of news. And so, then when you think about how you would treat journalism in terms of there's lots of rules about what is an appropriate thing to say about a private citizen versus a public citizen or somebody in the public eye. So, what the Washington Post can print about the president is very different than what the Washington Post can print about your or I. And that kind of standard, that kind of journalistic standard, does not extend down into social media because it's not considered journalism. But it functions in that space. And so, then

the question of what's the appropriate-- that's a very interesting point about sort of the sliding scale of privacy is that there is a difference in how we treat the privacy of sort of a private individual versus somebody who's in the public eye.

Speaker: Yeah, and I think in a commercial context, it makes me think about what constitutes a business record. If I were an attorney, and I were looking to bring action against an organization, and they tweeted something that sounded like a pronouncement of opinion or position, I could see you could say that that is tantamount to a business record, which then invokes a bunch of other things regulatory-wise. I don't think we've got answers for these things I guess is the point that we're making.

Speaker: Yeah, not only do we not have good answers, but the technical challenges around how those things are stored, and who's responsible for them, who assumes the risk, and who assumes the responsibility, how you define the harm. That whole ecosystem of how do you define what the legal implications of something is are not well-defined or well-understood. And so, then what you see is that the judicial branch has been relatively hesitant to make these definitions because they are standing at the point of saying well, we're working with law, legal authority, that's several decades old. Much of the FTC authority that exists now that covers sort of personal data

privacy dates back to the 1970s. And so, that standard, now the judicial branch is stuck with law that's relevant-- started in the 1970s, regulations that might be maybe twenty years more recent than that. And then they're having to stand in the spot of making that decision. But judges are not necessarily best positioned to be making those kinds of technical determinations. That is really the onus of the regulatory agencies because they have the mix of legal and technical expertise to really say okay, we have people understand both the legal ramifications and the technical ramifications, and we can say we know how we're going to enforce this. We know that we can enforce this. We can inform the public about what the implications of that enforcement is. So, I think that that's really interesting.

Speaker: So, we had one clarification of one question coming through. The clarification was-- this is from a little bit earlier. "If you are saying that the past data is impossible to erase, then are you saying the provider won't comply with GDPR?"

Speaker: I'll take that. I mean I think-- please, if you have an opinion as well. I think that the provider will make every attempt to comply with GDPR if GDPR provides adequate incentive to do something. So, I think much like other regulation, I'll point to HIPAA as an example. There is the effort you'll make, and then the effort

you'll make depending on the consequences.

Speaker: Yeah, I mean I think that like impossible-- past data is impossible to erase is sort of-- there's a couple steps to that, right? So, GDPR says to the provider that you have to erase past data. And at the collection source, it's technologically feasible to say and prove that you've done that. What the challenge is is with the secondary market. And so, when that data's been transferred and how it's been transferred and where it's stored and how it's stored and how it's tagged and how it's tied to the original provider, there's not a lot of clarity to say that every advertising company that buys a data product from Facebook can then tie back to an individual person what that data is. And then you can-- can you prove that the data that's been transferred possibly multiple times, combined and recombined into other data products, can you say it's been deleted in every instance? I wouldn't say that it's technologically infeasible for Facebook to prove that they deleted on you individually, but really just does that mean that it's gone from the Internet.

Speaker: Exactly, and what is that burden of proof? That's the other thing that needs to be decided. Once the genie's out of the bottle, I think it's both technically and economically unlikely that we fully erase something that's made its way into the wild.

Speaker: Yeah.

Speaker: And I think there's-- the world is rife with examples of that. So, I think that absolutely, organizations will try to comply with a sincere desire to do so, but it's very difficult to say conclusively something's been forgotten or erased.

Speaker: Then next, we have any thoughts on businesses like Equifax that profit from personal data that individuals didn't share and don't benefit from its use?

Speaker: Yeah, so Equifax falls in the category of data brokerages. And I had sort of intended to keep this discussion constrained to sort of social media and metadata contained there. Data brokerages are-- have separate regulatory status. They operate in a different way. But it is actually quite a relevant and adjacent issue. Data brokerages are typically these companies that you don't do direct business with. And you don't have the opportunity to opt out of participation in a data brokerage. And so, then with respect to Equifax, then keeping data away from Equifax becomes a very difficult process of controlling all of those potential sources that Equifax may-- or other data brokerages may be using to produce the products that they're producing. And so, it's a very complex issue. And then that goes into sort of the informed consent piece and the implications of opting out. Like if you opt-- if you could opt out of Equifax, how negatively would

that impact your life? How difficult would it be for you to say I can no longer participate in the credit market? That's a huge burden to an individual.

Speaker: Right, and what's the unintended consequences, right? I mean I could foresee this future where, if you allowed folks to opt out of those data brokerages, either you have to compel businesses to make credit decisions without the benefit of those sources or have some other kind of compensating way to determine the creditworthiness of someone. So, I think it's a really good example where I don't think there's an easy fix for that. Perhaps, we should ask more of those brokers in the way of protecting the data. The root of that question was sort of-- I think Equifax is probably used as an example because you had a breach. And this data was released in a way that was unintended.

Speaker: Yeah, I mean the data breach issue is something that I guess we haven't really touched on but is of course of extreme importance, right? When you have a data sharing relationship with an entity, like if it's Facebook or it's Equifax or it's somebody else, you want to feel like you can trust them to protect that data so that you aren't harmed by it. And the difficulty is that companies don't actually feel that motivated to do a good job of providing the kind of cybersecurity assurance that you would want because they don't actually perceive

it to have an ill effect on them in the end. So, there's a breach and then apology and an offer for like okay, we'll-- credit monitoring's on us. Like I don't know who even pays for credit monitoring at this point because everybody's been exposed at this point in some way.

Speaker: The enroll rate for those who are offered credit monitoring after a breach is only fourteen percent. So, I think there's a fatigue with all this, right? I do think breaches still matter. Right, I know there's a school of thought that says they don't. I had a really interesting discussion where someone said to me, "What about the Facebook breach?" I said, "There wasn't a Facebook breach." So, I think in the public's mind, there's a conflation of your data is out there, and there's been a breach. It's kind of interesting to me the way that's playing out.

Speaker: Yeah, I mean that's a good-- that goes back to the what is people's actual appetite for privacy because I mean if you assume that the Facebook hearing was about a breach, and that the data is out there because they were attacked and that data was put out there by somebody with malicious intent, rather than really just no, this is just you interacting with this company and then them sharing your data for sort of secondary purposes. But yeah, and also on the sort of fatigue issue of people sort of-- a lot of people have the assumption oh well, it's out there anyway. Anybody can get anything

on me from dark web. But also, the consequences of identity theft are real.

Speaker: Right.

Speaker: And expensive and challenging, and it can be very, very difficult to put your life back together. If you don't have a personal experience with that, it's very hard to see what the cost and the consequence is.

Speaker: Right, and to kind of bring the two discussion together, certainly, there's examples where social media data aids in identity theft. I mean social media data aids in physical theft, right? I mean there's the stories of you say you're in Cabo, so I know you're not at home. And I rob you. So, I think that there's-- it's all part of this larger dynamic of the information you share comes at a cost.

Speaker: Right.

Speaker: And it also has a value. That value can be by those who are using that data for some legitimate purpose or by those looking to rob you. So, and that's not going to go away. There's always going to be this tension in the way we share data in my opinion.

Speaker: Yeah, no, I would agree with that.

Speaker: So, we got this question/comment from part one of

our discussion last month. And I'll just let you guys chime in on it. It says a quote from Scott McNeely. He said it in 1999. There's no such thing as privacy any longer. Get over it. B.J.'s comment was, "We've known about this situation for a long time. Since you're showing us how gaining understanding of our data in this manner is so easy, it seems that by using social engineering, the bad guys can figure out our authentication credentials easily. How can anyone feel their data is safe at all ever?" It's all gloom and doom.

Speaker: Yeah, a couple of-- I'm sorry, Lena if you had-- so, I would say this, be deliberate in what you share. And be deliberate in the way you protect it. So, that is a truism I think in all facets of information security and life, frankly. So, yeah, can you be confident your data is not compromised, it's not being shared in some inappropriate way? No, I think the sort of easy answer is no, you can't. But I do think there's steps you can take as an individual, as a private citizen, to ensure that your data is as protected as possible.

Speaker: Yeah, and I mean I think that that attitude makes it seem like this is something that you can't have any control over. But really, you should be able to take some control over what you're sharing and thinking a little bit about the risks. And I think that what the end goal of changing sort of the regulatory context is is to make it easier for people to make good decisions to engage in the kind

of practices that Matt was talking about.

Speaker: Can we get one more, and then I'll let you guys wrap up final thoughts? We have about eight minutes left. One from Ellen asking, "Can you provide a little more detail about what the FTC can or will do?"

Speaker: That's tricky. I don't want to speak for what the FTC would do. I have not seen a new regulatory proposal saying here's our plan for changing regulations. Any legislative proposals that are on the table have been introduced but haven't been taken up by committee. And so, we don't know yet sort of what the new legal authority might be. So, if new legislation were passed, and it's consistent with some of the proposals that we've seen on the table, then the kinds of new protections that you might see from FTC coming forward would be notification of data breaches, would be stronger opt out requirements, and giving people sort of more opportunities and more transparency about what the opt out requirements are, and then more insight into what the sort of the secondary market is. In terms of what FTC could do without new legal authority, I mean they do have a fairly broad authority, existing authority, to regulate these things. And they may choose to take up sort of new rules consistent with things that they think would be sort of important. And so, yeah.
Speaker: Matt, anything to add?

Speaker: No, I think Lena's explained it very well. I would offer this. So, it'd be interesting to do this five years from now because I think we'll be likely talking about the same issues in different forms. I'm curious to see how effective GDPR really is. I'm curious to see what the appetite is in this country legislatively for doing something of substance on the subject. So, I think that-- that's not meant to be kind of an expression of pessimism. I just think that it's a very difficult subject. And the technology will continue to outpace our ability to understand it from a legislative perspective for a long time.

Speaker: Yeah, I mean I have heard some discussion that data privacy is likely to be a campaign issue for the upcoming election. And I would be surprised if it doesn't get discussed in that context of sort of it being something that people are definitely interested in. It's always interesting when you think about the regulatory context for an emerging industry, right? And so, the sort of classic examples of like the railroads, and--

Speaker: But this is my concern though, right? So, in a political science sense, this is a bit of a valence issue. No one is anti-privacy, really, right, at least publicly. It's like saying playground safety. No one's anti-playground safety. But what does it mean to be pro-privacy? So, my caution is I hope we're not too reductive in our thinking about this in the political context.

Speaker: Yeah, I mean I think that, again, the thing that I would want to stress is that as we think about what legislation is appropriate and what regulations might be appropriate is to really be wary of the problems of fighting the war that you're already in and not being cognizant enough of the pace of technological advancement and really making sure that any legislative proposal should really have a requirement to revisit the regulations on a periodic basis to just ensure that there is always sort of revisiting of what the technical challenges are and the technical context.

Speaker: Just one quick one from me. So, you guys have showed a lot of expertise in this area today. What are you guys doing in your day to day jobs at CERT that relate to this stuff? Is research going on? What work are you guys doing that can help this whole situation?

Speaker: Start first.

Speaker: Sure, in terms of like individual-- in terms of individual data privacy, I don't necessarily think about that so much. But in terms of data management and data hygiene, exposure, risk exposure that comes from not understanding the kinds of data that you might be releasing, there's lots of examples where that comes into sort of our more traditional customer context. And so, we are always thinking about how do you build a model that you can understand. How do you understand

where data is coming from and tracking where data is coming from? And data provenance, which is sort of the process of tracking the chain of custody of a piece of data is a huge issue for our customers. And it might not-- the data that they're transferring might not be sort of the kinds of things that we had been talking about in this discussion, but they're things that are of tremendous importance. And being able to do that is very challenging.

Speaker: I mean I think about this mainly from-- in the context of operational resilience for an organization, not so much from the data-- the individual perspective. So, in our work constructing models like the CERT resilience management model that helps you understand how data is one piece of an asset set that's required for mission success-- and as Lena said, I think all the fundamentals apply. So, data hygiene, cyber hygiene, governance, all of these feed into a larger collection of things that are required in the organization. So, I would suggest that if anyone's interested, there's a number of items available freely from the CERT web presence that could start as a point from which you can start to analyze and make decisions for your organization or as a consumer of data or as someone-- as an individual that possesses data as to what's rational to do in this situation.

Speaker: Yeah, and this work isn't relevant to my personal work, but

CERT definitely does a lot of work on insider threat. And then there's definitely an intersection here between balancing being able to address and think about insider threat problems and what data may or may not be available.

Speaker: Personally, as researcher, I'm very interested in this as a facet of overall business strategy. So, if we-- I'd love to hear from the folks that are watching the webinar. And if you are interested, we'd love to engage and get your thoughts on how we better tackle these challenges.

Speaker: Yeah, absolutely.

Speaker: So, about a minute left. I'll let you two wrap it up and take it from here.

Speaker: Okay.

Speaker: Anything else to wrap up the event-- today's event.

Speaker: Sure, I would just say, like all of these problems, don't despair. There are solutions out there. And just be a smart possessor of data and a smart consumer of data.

Speaker: Yeah, I guess I would just stress that when we think about the policy context to really just be aware of what the technical challenges are and make sure that we have good communication between technical experts and policy experts to make sure that we're crafting policy

proposals and environment that really get everybody what's intended and what they want out of those proposals.

Speaker: Great, Lena, Matt, thank you very much. Thank you for your expertise, great conversation. Thanks, everyone for attending today. An archive of this event will be available later today. We'll send out to everybody part one and part two. So, we hope that you share that with your colleagues. As I mentioned earlier, our survey is available. The link to that survey is in our chat tab now. And we ask that you fill that upon exiting as your feedback is always greatly appreciated. Thanks, everyone. Have a great day.