FTC PrivacyCon 2017
January 12, 2017
Segment 1
Transcript

KRISTIN COHEN: Good morning, everyone I just want to welcome you all to our second Privacy Con my name is Kristin Cohen and I am an attorney in the Division of Privacy and Identity Protection at the Federal Trade Commission. My co-organizers for today's event are Pader Magee, also in the Division of Privacy and Identity Protection, and Justin Brookman, who is in the Office of Technology Research and Investigation. Before we get started, I just need to review just a few administrative details that you've heard at every one of our workshops. Please silence any mobile phones. If you need to use them during today's event, just please be respectful of the speakers and other audience members.

Please be aware if you leave this building at any point during the day, you will have to go back through security. So please leave enough time. The restrooms are right outside the auditorium. The plaza east cafeteria will be open today until 3:00. If you eat there, you do not need to go back through security. But please keep in mind that only water is permitted in the auditorium. Most of you received an FTC event lanyard, and we do reuse those, so please return them at the end of the day. If an emergency occurs that requires you to leave this conference room but remain in the building, please just listen to the announcements over the PA system.

And if you need to evacuate the building, please leave through the 7th Street exit, and after leaving the building turn left on 7th street, and across E street to the FTC emergency asse

Whitney Merrill. And we also want to thank those moderating panels today, including Mark Eichorn and Jessica Rich.

We want to thank those who are heading up our poster session, Whitney Merrill and Tina Young. And of course those who put this whole conference together, Fawn Buchard, Crystal Peters, and Bruce Jennings, alongside our paralegal support from Carry Davis, Jonathan [INAUDIBLE] Joseph Kennedy, David [INAUDIBLE] Jennifer [INAUDIBLE] Kethan Dahlberg, and Omar [INAUDIBLE] And support from our division of privacy-- I mean our division of consumer and business ed, Jessica Skretch, and from our office of public affairs, Nicole Jones. So thank you all for all your hard work today. And now it is my great honor to welcome the chairwoman of the Federal Trade Commission, Edith Ramirez, to give opening remarks.

EDITH RAMIREZ: Thank you very much Kristin, and good morning, everyone. And welcome to our second Privacy Con. When it comes to privacy, technology has always presented a challenge. How can we make the use of the tremendous benefits of technological innovation while ensuring that our privacy is protected? This has been true from the snap camera of Warren and Brandeis' time, to the drones of today. The last several decades have brought change at a breakneck pace. The rise of the personal computer in the 1980s, the internet in the 1990s, the smartphone in the 2000s. And this decade, the internet of things. This dizzying array of technological advances is only going to continue to grow.

Last week, I had the opportunity to be in Las Vegas at the consumer electronic show, and had the chance to walk the showroom floor. There were smart cars that use technologies to sense driver emotion and deploy sensory outputs like sound, scent, temperature and light, in an effort to promote mental awareness and potentially reduce incidents of road rage. There were organic light emitting diode TVs as thin as cell phones that are capable of controlling the light of each individual pixel. There were passenger carrying helicopter drones that can be used to transport organs for transplants. Drones that can fold up and fit inside your pocket, and others that are outfitted with connected virtual reality goggles that promise a whole new experience.

From the robotic vacuum cleaner that also serves as a mobile home security camera and an air humidifier, to the smart trash can that can scan barcodes of disposed items in order to build a shopping list of items that need to be replaced. Almost all of these technologies however will rely to varying degrees on the collection of consumer information. And data collection is growing exponentially. Experts estimate that by the (y)20(t)-on banf2020,ate werl

Some of the risks of these new technologies are similar to ones that we've seen before. For example, traffic management technologies might only prove useful if they use data that includes a person's geolocation information. Now, we've long recognized that geolocation information is sensitive, and should not be collected or used without a consumer's opt in consent. Risks of unauthorized exposure of geolocation information include , stalking revelation of political, health, and religious affiliations, and even burglary. As this example shows, the possibility of unexpected uses for information must be weighed against the benefits.

But in addition to some of these familiar challenges there are also new ones. One is the ever-growing number of actors that have a role in collecting, compiling, interpreting, and using data in a world that relies and operates on big data, IoT, and AI. There are consumer facing companies, a device manufacturer, a smart hub platform, or a publisher website or app. There are behind the scenes technology companies, software vendors that connect IoT products to the internet. And of course, the numerous analytics and advertising companies. This vast array of entities makes it difficult to provide consumers with informed choices. And this challenge is exacerbated when non-consumer facing entities increasingly handle consumer data. This also raises concerns about whether all of these actors are appropriately protecting the security of consumer's personal information.

Second, with the new technologies, privacy and security failures aren't simply about threats to personal information. They can also include threats to health and safety. Particularly in relation to certain health devices and connected cars. For instance, the failure of security of IoT devices, in particular the ease with which IoT devices can be recruited into vast botnets to be used in DDoS attacks, could pose substantial risks. To meaningfully thwart potential botnet armies, a significant majority of manufacturers would have to act collectively to improve security.

Third, by relying on algorithms based on big data techniques and machine learning, companies may disadvantage certain populations. As we note in our big data report issued last year, even large data sets may be missing information about certain populations. Such as those who have unequal access to technology, or are less involved in the formal economy. And big data analytics can reproduce existing patterns of discrimination, or reflect the widespread biases that exist in our society. For example, an algorithm that isolates attributes of good employees or good students, may simply be replicating biases that existed in previous hiring or admission decisions.

The only way to keep this balancing act in equilibrium is to earn and maintain consumer trust. And this is where the FTC comes in. So what do these emerging technological developments, and the challenges that they present, mean for the FTC? It means that we have to continue to be nimble and smart to keep Pace And we have to leverage our resources. At the FTC, research and data play a key role in helping to guide our work. This is precisely why Privacy Con is so important. The research this event generates directly informs three critical areas of our privacy and security agenda.

First, we use research-- both research presented at Privacy Con as well as other research-- to identify potential areas for investigation and enforcement. For instance, tech researchers brought to our attention the practices of InMobi and Turn, two companies that were the subject of recent we alleged that

academic, tech, and policy worlds. We'll continue to learn from this event to enhance our understanding of consumer expectations, to inform how practices in this dynamic economy align with those expectations, and how devices and data can be secured in this new landscape. Today's forum, which is going to feature discussions on IoT and big data, mobile privacy, consumer privacy expectations, online behavioral advertising, and information security, will undoubtedly provide valuable insight on these and other issues. And it will help the FTC address emerging privacy and security challenges in a complex, dynamic marketplace.

So just to close I really want to thank our panelists for sharing their expertise, and all of you for joining us as we seek to study these important issues. And I really also want to take this opportunity to thank the FCC staff who organized today's event. And in particular Kristin Cohen, Pader Magee, Justin Brookman. And Mark Eichorn. I also want acknowledge Lorrie Cranor, our chief technologist who unfortunately will be leaving us very soon. But she's been an incredible addition and asset to us at the agency, so thank you very much Lorrie, for everything that you've done. So thank, you for being here, and now I want to turn the floor over to I believe Peder Magee. Thank you.

PADER MAGEE: Good morning. Thank you very much Chairman Ramirez, and thanks to the rest of the audience for coming out to Privacy Con Two. My name is Pader Magee, I'm an attorney in the FTCs Division of Identity-- of Privacy and Identity Protection and I'm going to be moderating our first session this morning, which I think is a good segue from the chairwoman's remarks. The first session is entitled, Internet of Things and Big Data. And I'd like to ask my panelists to come on up if they would. We have five researchers presenting on four separate and very interesting projects. They're each going to have 10 minutes to discuss their work and then we'll have a discussion period.

To get that started, I'll pose a few questions and then open up for audience questions. If you have something you'd like to ask, please line up behind the microphones after all the panelists have finished presenting their research, and then we'll take your questions. So let me start out by introducing Noah Apthorpe and Dillon Reisman from Princeton's Center for Information and Technology Policy. If you'd like to go to the podium and--

[INAUDIBLE]

PADER MAGEE: Thank you.

DILLON REISMAN: So, hi everyone. We'd like to thank the FTC for having us here today. My home. A smart home is a home in which devices-- or rather, in which traditionally analog appliances have been replaced by computers. You're probably already familiar with examples of this. Maybe some of you have a Nest thermostat for instance. Thermostats adjust the temperature in your home. A Nest thermostat is in some sense nothing more than a computer in the shape of a thermostat. It does the same thing, but it does it intelligently. It learns your preferences.

But there's also a second category of devices we're considering here today, and those are appliances that are brand new. They don't have a non-digital analog. An example of that would be an Amazon Echo. A smart personal assistant. You can ask it questions, it can answer them.

Each Bluetooth device announces its presence via advertising packets transmitted through the

Is this tracking happening? Are app developers already using this? It's a bit hard to answer, because the operating systems don't actually make it easy for you to analyze what is happening. And even if it was easy, you can't know if the apps are just collecting data, or making inferences based on it. But what we do know is that there is nothing currently to stop app developers from engaging in this kind of tracking and profiling. Well, the next question is, what can individuals do to prevent it? Well, not very much. They can turn off the Bluetooth on their phones, but that's about it. And if we are actually interested in taking advantage of the smart device revolution, turning off the Bluetooth on your phone is not really a functional solution.

To conclude, what we've done is we've identified a new type of privacy risks, profiling and tracking, that can happen using the nearby Bluetooth enabled devices. And we've shown that this new type of attack is feasible. And in the course of this study, we've also discovered some of the happening to their data in this type of environment and with these type of products. There is no easy way for the consumer to navigate the privacy and security of the new digital world. So we see a need for a consistent and accessible standard to be able to measure these products comparatively, to be able to determine what are better and worst performers for these sort of items.

Now in order to do this, we want to leverage the deep knowledge and the expert knowledge of many in our community. So for this initial effort, we're working with several well-known organizations, including Ranking Digital Rights, Disconnect, and Cyber Independent Testing Lab. Starting about six months ago, we got this group together to leverage the diverse expertise of the group, of this core group that we have, to start the process of putting together a proposal for a digital standard. So we met first to compile a draft of various criteria. We split up the work then to exercise tests for the various parts of the proposal against a few products from three different product verticals. We tried it against browsers, against some mobile apps, and against some connected devices.

And the purpose of the testing wasn't necessarily to investigate these particular products, but it was to vet what we had done, vet the proposal for sense, that it was feasible to do testing on it, and to improve based on what we had found from that course of that testing. So we're now refining the proposal, and we're working on getting ready to launch it. So before we can test of course, we need a shared definition of what is good, so we know what we're looking for. How do we define what is goodness in this space, in this digital world? So we started by structuring the work around four organizing principles: Security, privacy, governance and compliance, and ownership.

So security answers the question, is it safe? It includes topics related to encryption and security, security updates, passwords, things of that nature. Privacy is, is it private? Deals with permissions, over permissioning, and data sharing, and consumer control of their data. Governance and compliance answers, are the policies strong for consumers? It covers how well companies may protect consumers' privacy, and also freedom of expression. And ownership, is it mine? This covers right to repair, and covers things like permanence of functionality.

So for each of these four topics, we define several criteria which are anchored on consumer expectations. So this is what's interesting. We didn't start from technical requirements, we started from what the consumers would expect. So for example, instead of saying this room should be between 68 and 70 degrees, and with such and such humidity, I'm saying, the room should be comfortable. That's the consumer expectation. The room should be comfortable. Then we would dive down into that and develop indicators, which were attributes or behaviors that would achieve the criteria. So the attributes that relate to, the room would be comfortable, would be temperature and humidity.

And then we would define test procedures. And we'd take a thermometer and place it in four sections of the room at these different times and measure the temperature. But always we wanted to go back to what the consumer was expecting, what the consumer needed. So we've got the-- so here's an example. That's the first thing I want to go to. So this is an example of criteria that we had come up with.

The criterion, what does the consumer expect? The product should be protected from known vulnerabilities that presents a danger from attackers. That's what the consumer wants from their product. This criterion has several indicators that we developed. One of them, is the software secure against known bugs and types of attacks? And then a procedure, how you would test this. Well you could launch activities from the user interface and monitor communications to and from the device. So we could apply this method to an investigation we had done last year on a mobile app that is named Glow. And that app, we had done that investigation before we developed this standard, these criteria. But it shows that it applies, and it shows that the criteria is relevant.

So for example, in the case of this Glow app, the Glow app is a women's health and fertility app. It's a mobile app on your phone, and we found that an attacker-- we were monitoring traffic to

was a criteria that Facebook developed, and is not meant to be a direct measure of race. But is been to sort of be a proxy measure of what are your interests. So this was announced at-- well, it's been around since 2014, but was announced at South by Southwest as a major achievement in marketing. The press responded with some concerns, and then later pro publica, placed an ad raising more concerns for the government. It caught the attention of the Federal Housing authority and HUD. And subsequently Facebook actually altered its practice to no longer allow advertisers to target advertisements based on-- advertisements for credit, housing, or jobs, based on people's ethnic affinity. So you can see that it's just not a hypothetical example, there are real consequences in the world if people are unhappy with how you're personalizing your work.

So here's our research design. We looked at a couple of different domains, really common areas for personalization, advertising, search results, and retail pricing. We looked at different kinds of data types. So, what criteria are they using to personalize this to people? And then we looked at what the source of the data was. So this is a little bit more abstract for users, but did you provide the information? Like you tell some sources your gender. Did they guess it? Did they watch you and sort of see what you're doing? And if they did guess, if they infer the information, was it right or wrong? So using these three pieces, we assembled hypothetical scenarios that we gave to the respondents.

Here's one that actually corresponds to the Facebook ethnic-- ethnic affinity example. The ad is shown to you based on your race, which was inferred, and is accurate. So we gave people one from each section of domain, and they ended up with 18 different options that respondents saw. They saw one of these, where they were able to answer, how fair is this? They were able to give an answer. And they were also asked to give an open text response. The survey was designed through Qualtrics. It was administered through Amazon Mechanical Turk. And then you can see a little bit here about the demographics. I won't go into it too in-depth, there's some criteria that we got a pretty good measure of diversity, and some where we didn't. So, not perfect.

Normally I would give like a longer lead in to this, but here's the response for location. So this is city or town of residence, pretty non granular location criteria. You can see that for the most part, people felt like this was fair, this was above board, right? The middle line represents neutral, give or take. Abov neke a

And then the last one I think is really the most important result from this section. People were pretty OK after they thought about it for a while. At first you think, well it's not really fair to charge pretty acceptable across all the contexts. People felt like they understood why that would happen. They saw some benefits to that, in fact, for their own behavior. The third one I think is really important, especially at the FTC, thinking about the FIPS, is that the data quality was meaningful. Accuracy did matter to folks, and you saw some mixed results on it. But generally speaking, inaccurate data was perceived with a more negative feeling.

And then personalized pricing should mirror offline practices, as I mentioned. And the last one I think is really something that's going to become just more and more important as we go forward in the next few years, is that personalization based on race is really controversial. And I think what we saw in the quotes that people responded to was it was controversial because people weren't sure it was relevant. Which is really interesting because a lot of the goal of personalization from the company point of view, is to increase relevance. But they're using sort of a different version of the word relevant, than I think a lot of people are, when they're sort of casually thinking about what's relevant to me as a person.

So I think that this actually adds a lot of really important information. And you saw how the Facebook ethnic affinity example played out in real life. That should help folks who are doing startups or any sort of data intensive analytics think about how to improve the quality of the products they're offering and how to reach the audience respectfully, and yet reach a relevant audience. Thanks.

PEDER MAGEE: Thanks. Thanks very much. And thank all of the researchers for their great work. What I'd like to do now is, we've got a little bit of time, around 20 minutes, to have a discussion. I'll start out by posing a question for each panelist, and if anyone in the audience has a question, please line at the microphones and we'll call you in turn. So why don't I start out with the first project, Noah and Dillon. Your research indicates that a passive network observer such ane a ryzike aAeuoarc(a) 4( )-,uclth p]TJ   0 -(e)4(l)-14((a) 4(e)4(l)-12(e)t   2(e)4(5r2(ks)-1 out)-2(w)2c)-2(-6(ks

PEDER MAGEE: I know that one pressure on research is funding. Have you thought about developing-- partnering with people, or are developing it through Consumer Reports-- a bug bounty type program where you incentivize researchers to look for these types of vulnerabilities?

MARIA RERECICH: So, what's interesting is that one of the items in the standard we're working on is a company that's open to-- we think it's beneficial if a company is open to getting reports of vulnerabilities in their product. And so if they have a bug bounty program, that's a good thing for a company to have, because they feel it shows that they're open to that kind of information.

PEDER MAGEE: So that would be reflected in how you rank, or comparatively rank--

MARIA RERECICH: It could be. It could be, right. It could decide to be.

PEDER MAGEE: OK, great. Alethia, as the scope of big data gets bigger and automated personalization using machine learnin should be, start thinking more clearly about what you want and what you need, and what people expect and what they like. And then only get those parts and let go of the sort of blunt instrument stereotype categories that are causing probably more trouble than they're worth.

PEDER MAGEE: But doesn't that concept of getting rid of the blunt categories sort of encourage additional data collection so that you can refine it, make it more precise? There seems to be a little bit of tension there.

ALETHEA LANGE: I think you can learn stuff-- you can drop the underlying infrastructure that we pulled forward from offline advertising, where you had to be blunt because there we ha

so much. We're going to take a break until 10:35. Refreshments are available for purchase in the cafeteria, which is down the hall. Regrettably, neither food nor drink may be brought back into the auditorium.