Jules and Michelle's Big, Fat Data Journey

A Day in the Life of Two Data Wonks and Their Good Friend Mr. Fox

Michelle Dennedy: Your every move is being tracked. Every show you watch, purchase you make, link you click, tap, or swipe is being recorded. We willingly, often blindly, trade access to our contacts, photos, location and who knows what – just for a little online entertainment or for the convenience promised by the latest application. But what do we really know about the app developers and companies protecting and collecting this data? More importantly, what are they doing with it?

Join us as my guest, the lovely Jules Polonetsky, and I race through a typical day to find the answers on our big, fat data journey. Along for the ride is Mr. Jonathan Fox, director of privacy engineering and strategy at Cisco.

Jonathan Fox: Just going to do some backseat driving.

Michelle Dennedy: Cybersecurity, data protection, privacy. You like to stay ahead of the curve and listen to experts who are leading the way and deriving greater value from data with a more organized approach to data privacy. You're like us, just a few deviations past the norm. You are a privacy sigma rider.

Hi, I'm Michelle Dennedy, chief privacy officer at Cisco. Jules Polonetsky, my wonderful long-time friend amongst many other things, served as Chief Privacy Officer at AOL and DoubleClick. Today, he's the CEO and Founder of the Future of Privacy Forum. The FPF promotes the advancement of Principle Data Practices (this is why you have to move your mic back when you're talking on a podcast), and is supported by more than 130 leading companies and foundations, including Cisco. Some of the leading privacy thought leaders and academics in the US, also serve on its advisory board.

Jules and I thought it would be fun to talk about the various ways that people are being tracked, and the types of data that are being recorded, as people encounter tracking throughout their typical day. The types of projects that Jules works on don't just involve big data, they involve big, fat data. So let's get started with Jules and Michelle, and Jonathan's, big, fat data journey.

Welcome, Jules.

Jules Polonetsky: Delighted to be with you and Jonathan. You know, almost 10 years ago when Christopher Wolfe and I started the Future of Privacy Forum, our thought was, the world is awash in data and it's getting busier, and all the tracking that we are aware of (or that we suspect) absolutely raises concerns. But it also ... when used in a reasonable way, powers all the services and tools we like. Maybe gives us the data that we'll need to research and solve new challenges, diseases, artificial intelligence. So on the one hand, when we do it wrong, when we're not sure what's happening, it is tracking. When we maybe lean into it ... (apologies to Sheryl Sandberg), when we lean into it and tell people what's going on, then engage them, then on issues like the online advertising

type of issues that we work on, Smart Grid, connected cars, locations, wearables, health, all of these issues hopefully play out in a way that advances the kinds of things that most people would like to see happen in the course of a day.

Michelle Dennedy: It's pretty heavy stuff and I think ... I'm going to take a little bit of a step back because this is what our producer, Susan likes to say: The funniest think tank meets the funniest podcast. We both deal with really heady issues, some of which you've highlighted. How did you get into all this Jules, and what really excites you to go down all of these paths? These are really tricky problems to solve.

Jules Polonetsky: I spent the first third of my career in consumer protection, I was a state legislator, a consumer, a Ferris commissioner for the city of New York, a congressional staffer. Always represented and served sort of the working person. Blue collar constituents, people like my parents who never made a lot of money, but when they spent money, they wanted to be treated fairly. It's not that they didn't want to spend or they didn't want data used; they wanted [no] question a fair deal. They wanted us to know what was going on and to know that they weren't being taken advantage of, and I carried that over.

I think I learned that there was a gap. There were people very worried about data use, often maybe unaware of what was happening behind the scenes or policy makers who are concerned. Civil society folks worried about the bad things that could happen and sometimes those bad things did happen, and then there were lots of folks. Like you and the Fox and our colleagues who are chief privacy officers and security folks who certainly believed in the mission of their companies. They worked there. They're committed to helping solve the world's problems with tech and data, but they realize that you've got to do it right or you will be broken down, you will be criticized, you will run into legal issues.

There wasn't a place for those, let's say, middle-of-the-road people. It's dangerous in the middle of the road, right? You've got to dodge things on every side. That crowd, the folks who were optimistic about data, needed a place to work with academics, to work with advocates, to work with each other, and try to solve those problems to figure out how to take an optimistic view about what can be done with tech and data, but figure out the really hard things. How do you do this right? Not so easy to do some of this right.

How do you design things in ways people actually understand? How do you have large amounts of data and do large-scale research without creating giant databases of ruin? I think it's one of the biggest challenges society faces today.

Michelle Dennedy: I agree, so let's break this down a little bit. Let's go on a great data adventure. Why don't we start with waking up in the morning? So just in the earliest morning, what kind of information are we giving away from the minute we open our eyes and how is this data collected? And is data really free? Are we getting free services or are we getting served up for free?

Jules Polonetsky: First of all, when you wake up and hopefully you slept your seven-and-a-half hours, in fact you might check your fitness app, which probably reports that you actually slept only four-and-a-half hours last night. So how long I've slept may be available to a company or an app developer. If I've got a bad back and I'm worrying about posture, I may be sleeping on some sort of pad or

some sort of fitness thing that will maybe remind me not to snore, maybe will help with my sleeping posture ...

Michelle Dennedy: Is that a thing? I would like that.

Jules Polonetsky: Indeed. How you slept, how long you slept, and of course, maybe who you slept with is also available because frankly, the unique movements, if you're sleeping with that wearable as many of us do because we want to ... hey we want to get credit for any exercise possible and that counts, or tracking how long we've slept. So the unique patterns of how you move throughout the day are very telling. Then I wake up and I do some exercise and again, the algorithms that your wearable uses have gotten more sophisticated. People were upset that they only got credit for fast jogging, meanwhile they were sitting and perhaps doing some strenuous yoga or maybe riding a bike but their wrists weren't flapping.

So over the years, those devices have matured and the big data analysis of those devices really yields some useful information. Now some of this, let me give you the upside, right? If perhaps, Parkinson's is coming on, if there are links between your behavior that could be correlated later to disease, well that would be nice for us to know and help us advance medical science. So there is a positive. Some for the companies that actually work with researchers, but certainly where you've slept, you didn't sleep alone, how long you slept, and maybe your morning exercise routine, that's available and if you're choosing a fitness app, you do want to look and figure out what the rules are. Some of them may be free, which might mean if supported, meaning your data might be used, usually for marketing to you, and possibly for research.

Michelle Dennedy: There is so much in this data journey and I don't think in the next few minutes we're going to even get through the whole day because there is so much, as you're illustrating, even before you get out of bed, all of this information. Some of which is ... or can be very useful. So people that have sleep apnea, for example, obviously there are medical applications. But I think it's also very interesting about who gets which side and when. So you talk about your partners in bed. I happen to have slept with a gorgeous young blonde last night; she happens to be my 11-year-old daughter. Who gets to choose who knows about relationships, or status, or age? Who in the bed gets to decide choice?

So we're starting before you even open your eyes in the morning, you're talking about some information that's deeply personal and I think where the Future of Privacy Forum is most interesting and interactive with us is, we want to figure out as technologists and consumers, want to figure out as people, how are we interacting with this data, Jules? How are we influencing the build? How are we influencing the outcomes and will we ever really have control over our data? I'll ask you three giant questions as we're more than halfway through our 20 minutes together.

Jules Polonetsky: Let's take a natural, ethical issue that I think we already flagged. Certainly companies have for a long time said, "Look, I'm giving you a free product, or I'm giving you a subsidized product, so I'm going to use the data to make my product better and I'm going to use it for advertising and marketing. Hopefully they're not selling it, or if they are, you know about it and you're opting out or opting in.

Michelle Dennedy: Let me stop you right there. What is it? How do we even know where the beginning of it is? The edges of this data?

Jules Polonetsky: We're in a world that is awash in data. We're walking around with our phone, which is routinely pinging cell towers, Wi-Fi networks, so our location is almost invariably available for a whole range of reasons. We're sharing data on Twitter and on social media, so our relationships, our networks, things that can be inferred, where I go tells the world the kind of person I am. The kind of shopping I do, and again, some of this is going to power the next stages of our lives. I drive to work, I'm delighted that my car is aware of other cars nearby, can warn me that there is somebody in my blind spot. We're on the verge of cars communicating to each other to advise ... it's been very icy today in the area. Right now, I either need to see a car slipping or maybe if I have one of the newest vehicles, in a way a LIDAR is viewing and seeing this, but why shouldn't the car down the block alert me to the dangerous conditions? And that is increasingly becoming available and that only works, that safety measure, if everybody is in.

So one of the real hard questions that chief privacy officers and others end up citing is, look, clearly there are places where we want to give people choices. Opt in, opt out, do I want this tracking or not? There are other areas where it doesn't work unless everybody is in the mobility system that is starting to emerge, for example. A lot of research projects, a lot of health research, if only the people who choose to participate are in the mix, you get a very selective audience maybe only of the people who have the problem, or only people who don't have the problem. So the million-dollar question for most researchers or most companies is when is it fair to say I have the right, I have the obligation, or it's acceptable, ethical for me to use this data because it's going to be good for you, or it's going to be good for society? That's a big challenge.

Jonathan Fox: I think that maybe you were going to this Jules, but I think also the issue has changed just for advertising. Just for product improvement they have different definitions today than they did 20 years ago just because of the ability to micro-target or the ability to personalize to a greater degree. It's no longer just demographics on your TV and the time slot.

Jules Polonetsky: Let's take the smart home devices that increasingly are showing up. The Google Home, the Amazon Echo. On one hand, there was some thoughtful privacy by design, shall we say in setting these devices up because recording everything you're doing in your home 24/7, A, who's got the space for all that mostly empty data and B, that really would be scary and/or maybe my security system ought to be doing that, but I don't want an open microphone on my TV just because I want to control it remotely, or on those devices. So today, those are designed to be listening locally for a wake word, and dumping every couple of seconds, sort of the background noise that they encounter and not sending it out of your home. If I grabbed your Echo or Google Home or the Apple devices or probably the others that have been released at the consumer electronics show. There is nothing that is going to be on that device other than the last seconds of data.

However, when it does open up, now the commands that I'm providing are being sent back to those companies. Think about what they ought to be doing with that data? Presumably, my fast-talking Brooklyn accent, we want them improving their voice analysis so that every language in the world, every accent, somebody with a lisp, disability, any other challenge is going to improve and be more effective, but clearly, there ought to be boundaries as well. Because there is background noise that's detected. Maybe we want them learning how to screen that out, but we don't want them using that information, perhaps beyond our expectations.

Michelle Dennedy: So you said a bunch of things, which are ... excited my electrons over here. Both design, which of course, Jonathan and I have been hunting for many years, but also, I think it's fascinating and people should know that the FPF has a philosopher on staff and I think in the hunt for requirements, unless we are going to agree as a society that everyone can spy on everything at all times, the government can spy on us, and I don't think that any society has made that decision yet. If not, I think the decision as you've pointed out is not can we record, can you observe, but should we?

So can you tell us a little bit more about this philosopher you have and how they are adding to the hard decisions of design and even harder decisions of engineering on that design?

Jules Polonetsky: I think where philosophers are really useful is in two really interesting framings. [Evan] Selinger, who we work with, is really phenomenal, but people probably are aware as well as people like Helen Nissenbaum who has been a philosopher of privacy, or Luciano Floridi, a philosopher at Oxford who is sort of an information ethics philosopher, and I think they've helped us in a number of ways. They help us frame what privacy means. I don't think we want to think about privacy in terms of kick boxes or do I send you too much email. When you get naked with your loved one, if you are naked in front of your doctor, your privacy is not being invaded, even though you're naked. But on the other hand, if someone posts a picture of you, even if you're clothed but you're scratching your nose in an embarrassing way, and then distributes it in a way that embarrasses you, you certainly might feel that your privacy is invaded.

Michelle Dennedy: Yeah, I do a good enough job of embarrassing myself; I certainly don't need help.

Jules Polonetsky: Well, understanding that context can frame whether or not something is private more than what the actual data or the state is, I think, incredibly useful. Philosophers are also good at helping us evaluate risk and benefit. What is benefit? And in one framing that we worked on with the input of some philosophers, we helped try to assess when we say benefit, who? Whose benefit? The company that's marketing? Their benefit? Is it beneficial to you? Maybe it's not beneficial to you but it helps your community. Maybe it's pure research and it doesn't really help you today, but we might be advancing knowledge about climate control or about some value that maybe your culture holds dear and maybe doesn't.

So privacy has become far more complicated than the legal regulatory work that it was when many of us start out, and really a matter of capturing the broader impacts on different communities and philosophers are really helpful at defining those types of issues.

Jonathan Fox: And Jules, would it be too far to say that it also helps bring us back a nice full circle to what is fair?

Jules Polonetsky: What's the point of all this privacy if not to make sure that we treat people fairly? We're not defending this value because ... and some I think do see it as ... hey, let's just lock up all the information and make sure everything is private, but I think the broader thinking is really saying, why are we protecting this? What are the harms? We just did a chart of possible harms involved with algorithmic decision making and we said, what are all the possible things anybody may worry about, you may not agree with every one of them, but what are all the things that somebody might worry about?

Are they being exposed? Is there ... are the things that they are paying attention to narrowed from their view? Is there a fear of discrimination? What are all those factors and how do I go through, when I'm making a decision about guiding a product, shaping a feature? Do I take into account that broad range of issues, all of which can be enhanced or limited or affected by data and tech design.

Michelle Dennedy: So I'm going to ask you the big question, even bigger on ... as we close off our short journey together. What gives you hope, Jules? What gives you hope that we won't go Orwellian and that society or a number of leaders in society are going to choose wisely?

Jules Polonetsky: The frontier I think we're all grappling with today is machine learning and artificial intelligence. And when I talk to the lead researchers in those areas, people who want and need access to vast amounts of data to train algorithms, which they hope will be used from trivial things like better recommendation to analyzing the world's health data to come up with new ways to identify and treat disease. When I talk to them, they are all determined, they recognize the power of what they're working on, and they're determined that data be available to not just the largest companies, that the threats that people worry about, will this affect jobs, will it affect our economy, they get it. So I think we've come to a point where anybody can recognize that what we're playing with is both the ability to dramatically change, hopefully for the better society, or to dramatically harm, and it's not a matter of debating anymore over whether someone is doing a little bit too much marketing and whether the opt out works well enough, it's now a matter of how do we ensure that the power of data supports an optimistic future for human kind, and does it lead it to an Orwellian future.

We see some countries in the world where you do see the power of this data being used to surveil and to track and to limit and minimize, so it's not theoretical. You don't have to convince any business person or any technologist who might have scoffed at you in the past as being a [inaudible] or just being overly sensitive to marketing. They get it today and it's a matter of figuring out the path to that right direction. Rarely do you have to convince anybody that there is a great risk and a great benefit involved in the decisions we make.

Michelle Dennedy: It's heady and exciting stuff and I want to be respectful of your journey that you're about to take in an airplane and thank you very much for coming in and joining Jonathan and myself in the studio today and I look forward, watch this space, Future of Privacy Forum is in Cisco, [and] we have some really cool stuff planned for this year.

Jules Polonetsky: Concurred.

Michelle Dennedy: Thank you, Jonathan Fox, as well.

Jonathan Fox: Thank you.

Michelle Dennedy: Almost every company has designed decisions built into the questions they ask their users and specific reasons for the access permissions required to use their apps. As a consumer, you may not appreciate how a company categorizes you or makes assumptions, or you may wonder why a flashlight app may need access to your location. For app developers and companies, if you don't want to creep out your customers, it's important to make it easy for them to understand the why's behind your permission requests and be transparent with the way you use their data.

Arm them with the information to decide for themselves whether to opt in to your service, and you'll have a happier, more satisfied customer for the long term.

Jules, I didn't ask you before, how can our listeners read more about this and get in contact and follow you on your crazy data journey?

Jules Polonetsky: Lostinformation@fpf.org, and of course you can find us on Twitter and Facebook. Future of Privacy Forum.

Michelle Dennedy: Excellent. Well thank you so much, and Mr. Fox, do you have a Twitter or LinkedIn account people should be following?

Jonathan Fox: I'm on LinkedIn.

Michelle Dennedy: All right, follow Jonathan Fox on LinkedIn. He's a wild one.

You've been listening to Privacy Sigma Riders, brought to you by the Cisco Security and Trust Organization. Special thanks to Kory Westerhold for our original theme music. Our producers are Susan Borton and David Ball. You can find all our episodes on Trust.Cisco.com, or subscribe wherever you listen to podcasts, then please, take a moment to review and rate us on iTunes. To stay ahead of the curve between episodes, consider following us on Facebook, LinkedIn, and Twitter, and you can find me, Michelle Dennedy, on Twitter @mdennedy. Until next time.