Judith Kelly: Imagine that you are going shopping and you are going through the grocery aisle, and you're going through the bakery aisle. You're getting everything you need, but something's different in the store. You noticed a new aisle called personal information. So you're a little taken aback, you're confused. You go down the aisle and you say, "What is this aisle," and it seems like there are boxes of things you can buy. One of them says Age, Gender, Educational Level. Another one says, Political Beliefs, another one says Military Personnel Data, and one seems to be related to Government Employees, and there's one that's really interesting. It says Whereabouts, Real-Time GPS Locations. Confused, you are ready to leave this aisle and you look up, and all of the sudden you realize there are a lot of people in this aisle, a lot of them you don't know, and they seem to be buying some of these things. Some of them are, Oh, there's a person that might be the person who is building that new department store in town, and those folks over there, are, those are the heads of the political party that's currently running your town, and well, you're ready to get out of here, because, Oh, there's your ex, so you are bolting. This is really creepy. Is this legal? What is going on? My guest today is here to answer some of the questions about what might actually be going on, even if it's not happening in the grocery store. I'm Judith Kelley, Dean of the Sanford School of Public Policy at Duke University. Welcome to Policy 360.

My guest is Justin Sherman. Justin is a cyber policy fellow at Duke University's Technology Policy Lab here at the Sanford School, where he directs the data brokerage research for Duke's Privacy and Democracy Project. He's also a fellow at the Atlantic Council's Cyber Statecraft Initiative and a contributor to Wired among other roles. Welcome to Policy 360, Justin.

Justin Sherman: Great to be here. Thanks for having me.

Judith Kelly: It's great to have you here. Before we get into the details, let's talk for just a bit about data brokering. What is data brokering?

Justin Sherman: Data brokering is an industry whose entire business model is buying and selling our data. We may never have heard of the industry before. We may never have interacted with these companies before. But I think we're all familiar at this point after Cambridge Analytica, after any number of privacy breaches and data leaks that companies like Facebook or Twitter, companies we interact with directly are collecting our data.

A lot of consumers have some general understanding of that fact. What a lot of people don't know is that once you interact with that business directly, there's an entire second layer of data sharing going on. And that's what data brokerage is. Those companies then buying and selling and sharing and licensing all of that data around in this global multi-billion dollar pretty unregulated industry.

Judith Kelly: You say pretty unregulated, but when I am on websites, I remember seeing... They say, "Oh, we don't sell your data." Shouldn't I feel safe?

Justin Sherman: Right. A lot of the websites will tell you that they're not selling your data. There's no way to check that. There's no way to enforce that. We see all the time that companies make all kinds of claims about cybersecurity or privacy that are exaggerated or not verifiable and end up being not true. The fact is in the US, there are very few to virtually no restrictions on this industry I just mentioned, where all it's about is buying and selling our information.

Judith Kelly: You said in the US. Why did you say in the US? Is it different elsewhere?

Justin Sherman: Yeah. Well, sort of. The US has a particularly weak privacy regime. If we think about, for example, the European Union, a couple of years ago, the GDPR, the General Data Protection Regulation, went into effect, which put a bunch of controls around what companies could do with your data, and also significantly made sure that companies had to disclose to users that they were collecting data on them. There are other countries with privacy laws of varying strength and quality. Brazil passed one last year. India in the process of building one out.

For all the Chinese government spies, the Chinese government is even putting controls on Chinese companies. But in the US, we don't have a strong federal privacy law. That's why a lot of these data broker firms that operate here and are buying and selling Americans' data really are not operating with any serious restrictions.

Judith Kelly: When I'm on a website, is it just that website collecting data, the company whose website I'm on collecting data about me? Or might it also be third parties that the company might not even realize are roaming around on their webpage?

Justin Sherman: There are tons of third parties on those websites, and this is a really important point. When you go onto let's say Amazon, there might be a number of other companies whose code, whose app services Amazon is plugging into its website. They might also get information about you. When you open Uber, to give another example, Uber might use a bunch of plugins where advertisers or other companies are getting your GPS location data.

That's a really important point is that it's not just the companies we interact with directly, but sometimes there's a whole network of other code providers, advertisers, the third parties that are getting that information as well.

Judith Kelly: All right. You have taken a good look, Justin, at some of these larger data brokers. What were you able to find out about what they are doing and collecting?

Justin Sherman: For this report, which is the first out of our project here at Duke, and excited to say we'll be having many more of these come out, I looked at 10 large data brokers and did, as you referenced, a pretty wide survey of the kinds of data they were advertising on American citizens. This was not going sort of behind the scenes and seeing what they might be doing that they're not talking about. This was just saying, let's just look alone at the things that they're openly saying that they have, the kinds of data that they have.
What I found was pretty disturbing. I found companies advertising sensitive demographic information, so things like your race, your gender, your income level, your marital status. I found data brokers advertising your political preferences and beliefs. There was one broker, for example, Affinity Answers, that has a whole list saying, "We can tell you if somebody likes Bill O'Reilly or Arianna Huffington or Anderson Cooper. We can tell you if they support the NAACP or the ECLU.

Judith Kelly: We're not just talking here how you're registered to vote.

Justin Sherman: Yeah, exactly. We're talking about data on politicians you support. There was also data on if people supported a party organization in their state. So yeah, it's definitely a lot of information about voting behavior.

Judith Kelly: What are the names of some of these companies?

Justin Sherman: One of the largest of these data brokers is Acxiom, which not many people have heard of, but it's based in the US. It advertises data on literally billions of people around the world. And that's an important point, right? When we think about surveillance and data and companies, we think of big platforms a lot of the time like Facebook and Twitter, which have global user bases. But a company like Acxiom with data on billions of people, I mean, that's actually at the same scale, if not even greater than some of these platforms.

But Acxiom is one. Equifax, which people are likely familiar with from the Equifax hack a couple years ago, is another. But then there are a bunch of smaller companies, as well as companies like Oracle, which some folks might associate with cloud computing or software, but have in the last five or six years realized the amount of money to be made in this data buying and selling and have shifted their business models almost entirely to brokering data.

Judith Kelly: How granular does it get? Can people also find out like our shopping habits and things like that?

Justin Sherman: They can. These companies, as I detail on the report, talk about knowing what you buy, where you buy it, when you bought it, how much you spent. Off of that, they can do all kinds of things like extrapolate your favorite stores and brands. They can perhaps make estimations about how much money's coming into your household. They also advertise real-time GPS location, quite literally your GPS location from your phone, which is obviously a very intimate piece of information.

Another example of how knowing that directly enables you to do all kinds of things, we can talk more about, follow people around, things like that, but you can also learn sensitive information about people from it. The data in a lot of cases does get very granular.

Judith Kelly: That is kind of creepy. The following round part seems particularly invasive. Could an ex, somebody who's been divorced from somebody, or somebody who wants to stalk somebody for some reason, is this real-time, kind of like Find My Phone? If you are willingly sharing the same data with somebody, can somebody get that data even if you haven't consented to it?

Justin Sherman: They can and it has happened. And you're right, that it's both disturbing and there are real risks of physical harm. I'll give two sets of examples. Another great project that the Sanford School has been running is looking at cyber policy and gender violence. There are many cases of abusive individuals getting access to broker data and then using that to stalk, harass, hurt, even kill. People have quite literally been murdered, particularly women and members of the LGBTQ community, because let's say they move to a different state.

They change their address. The abusive individual goes online. These data brokers have this data out there. They find out where they live, where their kids go to school, and they hunt them down. There are lots of real cases of this, even if they're not publicized for privacy of reasons for survivors in many cases. That's one real risk of this kind of data. Another which is a little more recent in the news is listeners may have seen that there was a Catholic website called The Pillar that bought real-time GPS location data from a data broker that got it from Grindr.

What these people at this website did is they then stalked this priest, this closeted priest, who was using Grindr, followed him around, tracked him as he met up with men through the application, and then they outed him. Since then, they've actually said in The Times in other words have reported this that they have more data and they might out more priests. All to say, right, you can, as you said, picture these scenarios which have really happened where you get someone's real-time GPS location. That person has no idea.

And you're actually able to in real-time track them around as they go to their home, as they go to a bar, as they go to an abortion clinic even, right? There's a lot of real potential harms that can occur.

Judith Kelly: You say this is not very regulated, but this is pretty scary stuff. Aren't there other laws that can be used, like consent laws and things like that, that could be used to go after some of these companies and say, "You can't really share this information?"

Justin Sherman: There should be laws to say that. Unfortunately, the way it is in this country, beyond a few very narrow cases when we're talking about health data in HIPAA or credit information in the Fair Credit Reporting Act or FERPA and student data is another one, but there aren't a lot of categories of data in this country that are subject to the kinds of consent requirements you just mentioned.

When it comes to things like what do you buy, where do you travel, which political figures do you support, what's your GPS location, as sensitive as that stuff is, as dangerous as that could be in the wrong hands, it's not something that our policy makers have regulated.

Judith Kelly: I just picked up my phone. I'm just going to turn my eye allocation services off. And then there's a button says tracking. It says, "Allow apps to request to track," and it says, "Allow apps to ask to track your activity across other company's apps and websites. When this is off, all new app tracking requests are automatically denied." Am I good?

Justin Sherman: This is exactly part of the problem, and this speaks to why your point about third party code is so important. We can tell an application to stop collecting our location data, for example, by turning it off in the settings. That doesn't always mean they're doing it, again, because that's a control on the software side that your phone has put in place and not something they can necessarily be punished for legally. But there's lots of applications, as we mentioned, that plug into other applications and get data.

One particularly shocking example is that a couple of years ago, it was discovered, it was reported by journalists that Facebook had all of this data on women's menstrual cycles and women's menstrual health. It turns out that even though people had literally no idea, a lot of women were using a health app that was not made by Facebook, that was not owned by Facebook, that was a completely separate thing, but Facebook was plugged into it and getting the data.

That's the kind of case where, like you just said, when you pull your phone up and you see that it says other apps might track you, you can't really be too sure how many advertisers and software code development kits and other things are plugged into a single actor using.

Judith Kelly: Another thing that you had mentioned in a tweet of yours was about data on military personnel. It seems to me that that would introduce some national security concerns. No?

Justin Sherman: Absolutely. Multiple of these brokers in the ecosystem advertise data on all kinds of classes of individuals. There are brokers that advertise data on first responders or healthcare workers, brokers that advertise data on students, on people of particular races or sexual orientations, right? Another category of data that these brokers advertise is military personnel, and that's one of the things I tracked in this report.

But as you said, when we're talking about all of this information, we've been saying GPS locations, who your family members are, where you live, what you buy, what you think about politics, all that stuff is potentially useful to say a foreign intelligence service who's trying to learn more about a senior member of the military, for example.

Judith Kelly: Right, or the government. You said also for government officials. It's the same issue.

Justin Sherman: Yes. We've obviously heard a lot in the news in the last five or six years about the US government being concerned that the Chinese government in particular is building large intelligence data sets on our diplomats and civil servants. When the Office of Personnel Management was hacked several years ago, that, of course, was the main concern from our intelligence community.

To think that all of this information is sitting out there, you can imagine lots of scenarios where the things that someone buys online and their internet search history is not something that an intelligence agency already has and It's just sitting there. Maybe they're just going to go buy it through a shell company or they're going to hack it from the data broker. There's lots of ways that data can be used, as you said, in ways that threaten security.

Judith Kelly: Now, one of the companies I think that was in your report, if I remember right, was Equifax. Is that right?

Justin Sherman: That's correct. Yes.

Judith Kelly: Now, that's a company name that's familiar to a lot of us because of the breach of our credit data some while ago.

Justin Sherman: Yes, they were hacked and millions and millions of Americans had all kinds of sensitive information, credit information leaked. Equifax is a large data broker as well. It doesn't just do credit stuff. It sells all kinds of other data about individuals around the world, but especially Americans. When we think about, again, what are the potential harms, well, one, here's a company that we've already seen has very intimate data on people. That was just sort of one sliver. There's a lot that they have as a broker that's really fine grained on individuals.

But the second point, again, is about what happens when these companies have giant databases on millions of people and they don't secure them properly. We saw this with Equifax where the... This was not a sophisticated hack with Equifax. This was some serious security failures that people knew about and didn't act on and things of this nature. Again, to think as well that these brokers have these giant spreadsheets on all kinds of things about you and that they're keeping it secure is just not always the case.

There have been examples of data leaks from brokers. There was one last year where a data broker that tracked people's social media profiles had 200 or 250 million people's social media information just sitting unprotected on a server with no password. You get these cases and Equifax is a great case study where the broker doesn't have good security, and so there's an additional issue with the data getting leaked potentially.

Judith Kelly: But what's so ironic about that whole thing was Equifax hadn't taken good care of my social security number and my other data, and therefore their solution was, "So now we will offer you free credit monitoring for X number of years." I have no choice because somebody's got to monitor it. I can't, right?

Justin Sherman: Right, right. I mean, tangentially, they even had a whole other issue where instead of having people submit complaints about data leak issues or identity theft issues through their website, they set up some other website. I forget what it was. It had a ridiculous...

Judith Kelly: I do remember that. Yes.

Justin Sherman: Some ridiculous name like reportyourequifaxbreachhere.com. I mean, I'm being a bit hyperbolic.

Judith Kelly: Or checkwhetheryourcreditscorehasbeenviolated.com.

Justin Sherman: Right. What happened is, of course, immediately a bunch of cyber criminals bought domains that sounded very similar and then actually stole many more people's information. Again, I mean, some of these companies are not... I know they changed their security leadership and stuff like that since then, but some of these companies are not even protecting the data well that they're holding on us.

Judith Kelly: Yeah. Now, I mean, as long as you live in a society where people are respecting each other's rights and things like that, we may be able to get by. But if we start to think about some of these new laws that have started to get past where you have citizen enforcement of laws, like the abortion on Texas, for example, where you can actually make money on turning in your fellow citizens for various behaviors, in this case, going to an abortion clinic or even giving somebody a ride to another state to get an abortion.

We could imagine that type of citizen vigilante rules pertaining to other domains, and you cross that with this kind of ability to get data on people or observe their whereabouts. Does that have you worried?

Justin Sherman: It does. Yeah. There are brokers who advertise information about people's health. It's long been known that some of these brokers will have data where they're saying or knowing or predicting that people are pregnant. There have been cases even where there's a younger individual or something of this nature who's at home, and all of a sudden, they're getting maternity clothing flyers in the mail. The news is out now, or maybe they got an abortion and then they get such an advertisement. Companies do have information on...
It's extremely intimate, but they do have information on things related to people's sexual activity. I mean, it's very disturbing. That combined with all kinds of other things, facial recognition, like you said, GPS locations, when we think about something like this very just disgusting I have to say law, you really do enter into a space where there's a real opportunity for physical harm.

Judith Kelly: Physical harm, which, of course, is the ultimate concern, but it seems like that these systems could also be a threat to just our system of governance to democracy itself.

Justin Sherman: Absolutely. One way this happens, I think, is government purchasing of data brokers. We've talked a lot about individuals using data broker data. We should talk more about companies that buy it, but another big customer for these brokers is the US Federal Government. And the reason is that there are a bunch of controls in place around what federal law enforcement and security agencies can gather vis-a-vis data on Americans, warrants, Fourth Amendment controls, all kinds of things.

The reality is a lot of that can be circumvented by just buying data from a data broker. The FBI, Immigration and Customs Enforcement, DHS, a bunch of federal agencies have been reported or caught really, in some cases, buying phone location data, buying data on smart thermostats and other devices in your home to be able to see is someone and living in a particular residence, for example. There's all kinds of these cases, and that I think is an example of why this is undemocratic.

We, as a country, have spent years very imperfectly and we're still far behind right with privacy, but trying to figure out, okay, we have these boundaries, we've set up between the government and citizens and spying on citizens, and we have these Fourth Amendment protections we need to keep updated. And here comes the data brokerage industry offering a lot of these agencies away around it. The FBI, for example, there's prohibitions in place where they can't just go to a cell company and willy-nilly they'll hand over data on let's say a hundred million people's phone records.

However, what they can do is they can go to a data broker that got it from the cell company and just buy the data from the data broker. They don't need a warrant, and they don't need to disclose it, and there's no oversight. So stuff like that I think is part of why this is really just an undemocratic sort of ecosystem.

Judith Kelly: Justin, I know you said there isn't a lot of legislation out there. It's virtually unregulated, you said. If you could pass one law, what would that be?

Justin Sherman: That's the eternal question. We've spent a bunch of time thinking about this. I have to give credit to David Hoffman and Ken Rogerson and Joe Andellinger and so many other great faculty who work with our data broker project. But there are three recommendations I make in this report for what Congress can do. The first is that any federal privacy law that's going to be passed has to address data brokerage.
If the privacy is only about, like we mentioned earlier, the relationship between you and one app or you and one commerce platform, we're missing that whole second layer of data sharing and buying that's going on. The second thing Congress needs to do is to give the executive branch the ability to put some export controls around data sales. To say, if you're selling to an entity that has a particular tie to let's say a foreign military organization or a foreign intelligence

service, we're going to block you from selling let's say a hundred million Americans' GPS locations.

And then the third thing is to give the FTC more authority and more resources. The big thing here is it's not just the ability to investigate data brokers themselves. It's also making sure the FTC can investigate misuses and abuses of data broker data. Because it's not just the fact that there are brokers that advertise data on people with Alzheimer's and dementia.

It's the fact that, and this has actually happened, that scammers have then bought that data, figured out all of these Americans with brain health issues, and then used that to target a bunch of Medicaid scams let's say. Again, the harms are real, but it's not just the brokers themselves, it's also the uses of data off of that.

Judith Kelly: Yeah. I mean, already, Justin, a lot of the things that are going on are nefarious in the sense that they are illegal. Are there code that can be written? Are there actual actions that can be taken to enforce it other than criminal penalties?

Justin Sherman: It's never going to be perfect. There's always people who will try to circumvent any law. Particularly with technology, people will try and write workarounds and code. But one thing you can do, I think, is technical audits of an application to try and see where it's sharing data. The Consumer Protection Authority in Norway actually has done a couple of these, where they've looked at an app and they've done an evaluation to see how many other advertisers, third party code providers, that that application is sending its data to.

They actually did this with Grindr last year, and they found, you can read, it's quite disturbing, dozens and dozens of companies that are plugged into just that one app getting information. If you had some kind of say control where you said, "You need to disclose the list of everyone that's getting the data used by your app," and an app listed three people, three companies, and then you did an audit and found actually there's like 57 advertisers getting the data, that's an example of a technical check you could do.

Again, it's not perfect. It's like an inspection that you'd have to do, but I think that is at least one potential option.

Judith Kelly: Well, it sounds like at least we might want to think about leaving the phone behind a couple of times when we go for walks or do other things. Maybe we don't need to have it with all the time.

Justin Sherman: Yeah, never a bad idea. We joke, right? If you told a spy agency from 50 years ago about an iPhone, they would probably start salivating. I mean, I don't think they would've ever believed that people would carry around a device that enables all those kinds of data collection. Obviously, we get benefits from them, of course, but it is important I think time and time again to be reminded of all of the data that's collected on us because of it.

Judith Kelly: What makes you most optimistic, if anything, in this regard?

Justin Sherman: There are a lot of really great people on the hill, in the FTC, and other parts of the government that have long been paying attention to the data broker problem and have a really deep understanding of the industry. The challenge, I think, is the political momentum issue. Obviously, as we've said, the big platforms, Facebook, Google, Amazon, get the vast majority of the attention with privacy. I am optimistic because I think that political momentum is growing again.

Even in the past several months through this project talking to policymakers, it has been heartening to see just how many people are dedicated to addressing these issues. That's really the main thing going forward is building visibility for this issue and making sure people understand it's not just the business you interact with that has your data, but a whole ecosystem behind it.

Judith Kelly: All right. We're going to stop right there, Justin. Thank you so much for being on the show today.

Justin Sherman: Yeah, thank you for having me.

Judith Kelly: My guest has been Justin Sherman, and Justin is a cyber policy fellow at Duke University's Technology Policy Lab here at the Sanford School. And right now he is also trying to earn his master's in security studies from Georgetown University. Good luck with that, Justin.

Justin Sherman: Thank you.

Judith Kelly: Thank you so much for joining me. We'll be back soon with another conversation. I'm Judith Kelly.