FTC PrivacyCon 2017
January 12, 2017
Segment 4
Transcript

KRISTIN KRAUSE COHEN: You came back from lunch and that there's a lot of you here. We have a great panel ahead of us on online behavioral advertising. And I hope many of you were able to check out the poster session. And thank you to all of those folks who put up posters. I found it really informative and really appreciated that.

Now he's wearing his mask so he can rob banks to fuel his adrenaline-powered lifestyle. But others wear a mask in anonymity to, we engage in anonymity. One of the things we get from that is autonomy. There's a zone where we can engage in behavior, we can make decisions free from observation, free from interference.

Autonomy, this notion of privacy, this zone of autonomy underlies a lot of 14th Amendment jurisprudence. It goes to reproductive rights. We're given a space where government can't intrude on personal and intimate decisions. So autonomy is very basic in our constitution.

And there are also some utilitarian reasons we may want to engage, to protect autonomy. You know, first of all, you may, if you're being watched all the time, if you can't be anonymous, you just may be embarrassed about what you're doing. So there is dignity harms.

If you're being watched and this is an important thing and what the paper goes to, you then may refrain from engaging in certain things that are crucial to your development. Things, everyone needs a zone to explore or to figure out who they are. So if you can't engage in that personal development, that's going to retard your growth as a human.

That can also have the spillover effects on society, right? I mean, so this idea of self-discovery has a lot to do with creating variety and which some argue may be crucial to a functioning democracy. So we need this kind of variety.

And finally, if you can't be anonymous. If you can't go check that book out at the library and with the mask on, maybe you won't do it. Maybe if you've got some condition that you want to learn about, but you're embarrassed that someone will know, you'll engage in what's called privacy-protective behavior. That also is harmful.

So there are a lot of reasons why, this concept of anonymity, which is one of the concepts of privacy, is important. So what am I doing here? What is the study about?

Well, we've seen from other panels and people who are familiar with the privacy scholarship, one of the big questions here is to figure out how consumers make trade-offs with privacy and other values.

Now, there are a lot of work out there that tries to get at that directly by conducting experiments to figure out how much people are willing to pay or are willing to trade off in some sort of hypothetical trade between money and privacy.

Well, what I wanted to do is to go and look in the real world and it's really hard. And one reason a lot of this work is done experiments is it's kind of hard to find natural experiments or find real world experiments where privacy's been changed and gauge consumer reaction. And that's what I wanted to do.

I wanted to, even though I'm not measuring directly the valuation of privacy, I wanted to measure how people change their behavior because of a reduction in anonymity. And

Google Trends is something that Google constructed from 2012, I'm sorry 2010 through present day, that's an index. And it's publicly available. You go, Google Google Trends and you'll go to their page and you can see all this metric of search groups.

Now it's not a search volume, it's an index of relative popularity of search terms and somewhat complicated and somewhat opaque exactly how it's done. But even though it doesn't capture direct search volume, it does accurately capture trends.

So if a search term is going up or going down, you'll get the directions right. You may just not, the magnitudes aren't there because it's an index that ranges from 0 to 100, all right?

So I'm going to do this in a difference-in-differences framework. And what does that mean? It means I'm going to have a treatment group and a control group and I'm going to look at them before and after the experiment. The experiment here being the 2012 policy change.

So in this experiment, the treatment group is sensitive search. The kind of search that I may be deterred from, that I think that this reduction in privacy is going to impact. All right? And the control group is non-sensitive search. Things like weather. That I don't care if people know I'm searching about weather. I may care if they think I'm searching about porn.

And basically the unit of observation is the Google Trend score for each of these searches. I have 20 sensitive and 20 non-sensitive searches. And for a week in a state, depending on the window I look at I have between 13,000 and 108,000 and I use lots of different controls. Week, term, state, fixed effects, as well as I search specific trends, things like that.

So what are the main findings? Well, I find in a short window, basically around March, and there are some things that I've done some work since that is not in the paper that was submitted. So it'll be coming out probably in the next couple of weeks on SSRN with some updated stuff.

But I find depending on the window you look at and the controls you put in, somewhere between a 5% to 10% reduction in sensitive search relative to non-

this, right? They purchased BlueKai, they purchased AddThis, and now both of those tracking data sets you can imagine being combined.

The second point I want to make is that trackers can impede HTTPS adoption. So the HTTPS is basically a secure connection to a server. If you want to make sure you're connecting to that server and you want to make sure that the data you're getting back hasn't been tampered with, you'll do it securely through HTTPS.

And if that server happens to load resources from a non-secure server, it's called mixed content. So basically, some of the content on the page isn't secure.

And if that happens, some browsers like Firefox will show actual, will show a warning, a downgrade to the security indicator that in my opinion looks worse than just a normal HTTPS page which wouldn't have any kind of yellow warning on it.

So sites which may not normally want to adopt HTTPS because they're not handling credit card data, they're not handling log-

And I want to focus on the fact though, that we're seeing browsers respond to this. So a year before we did our study there was a paper called The Leaking Battery, which actually looked at the use of the Battery API for fingerprinting. And then we went and measured the use of that in the wild and we found that sites were actually doing it, or scripts were actually doing it.

And in response to both of those papers, both Firefox and Safari ended up removing the Battery API from the browsers. Firefox unshipped it and Safari just never shipped it.

And I think the point here is that browsers are

KRISTIN KRAUSE COHEN: Thank you so much, Steve. I'd like to introduce Zubair Shafiq, who is from the University of Iowa, who's going to tell us about anti ad blocker.

ZUBAIR SHAFIQ: Thanks, Kristin. So my name is Zubair Shafiq, and this is joint work with Zhyiun Qian at UC Riverside, who is also in the audience. Today I'm going to talk about our research about the arms race between ad blockers and anti

unfortunately, this program does not give users any explicit control to opt out of the tracking which enables these targeted ads.

Moreover, many publishers and advertisers are not part of this ad choices program. There was another interesting effort called Do Not Track, which was in 2009.

This was supposed to be another voluntary program. It was supported by major browsers, but unfortunately the online advertising industry did not catch up and they did not support this voluntary opt out service.

Very recently the FCC has passed some regulations barring broadband providers to monetize user information without explicit opt in. And this is, obviously, very commendable.

FTC has been, I've heard, interested in something similar, but the recent political climate, uncertain political climate makes it unlikely that any media regulation would be passed in this regard.

So really the privacy conscious users and the research community have started to look at some technical countermeasures to get, to solve some of these privacy problems. There are these privacy enhancing, privacy preserving tools, which are commonly used.

For example, Privacy Badger, which specifically targets publishers that do not respect, do not track. Or Ghostery, which is a proprietary tool to block trackers. There's also increased adoption of the so-called ad blocking tools, which are generally open source and they use public filter lists to block ads and trackers on websites.

These ad blockers have become very popular. According to a recent estimate by PageFair, more than 600 million people around the world use ad blockers. According to a recent academic study by comScore, more than 18% of users in the US, use ad blocker.

In the male demographic, between the ages of 18 and 34, 50% of users in Germany use ad blockers. In the US this percentage is around 30%. So clearly, hundreds of millions of people are using ad blockers to protect their privacy.

Unfortunately, the online advertising industry sees these ad blocking tools as a growing threat. So they resorted to these so-called anti ad blocking techniques. So these websites employ these anti ad blocking scripts, which detect users who are using ad blockers. And then they force these users to disable their ad blocker or whitelist the website.

Many popular online publishers, such as the Washington Times, Wired, Forbes, have recently tried to interrupt and block ad block users. Such attempts to undermine ad blockers could mean a return to the status quo.

Therefore it is very important that we develop effective and long lasting countermeasures to circumvent anti ad blockers and strengthen ad blockers. So the goal of our project is twofold.

So I would like to really point out that ad blocking is not a problem, it's a symptom of a deeper problem in the online advertising ecosystem. Right now it's all about data. Companies want to money monetize user information, user data.

Right now you have to balance between users, society and economy. But right now these companies are putting the economy first and users at the very end. And this has to change. You have to put users first in this trade-off.

Our research aims to put users first. Our research is to make a stealthy ad blocker. We'll give users control over which ads and trackers are OK. So if they want to support a website, they can choose to whitelist a website and allow advertising on a particular website.

In the long term, there is a need to make better and more robust privacy preserving tools. Thank you.

KRISTIN KRAUSE COHEN: Thank you, Zubair. So we have a short period of time for some question and answer. And I encourage any of you in the audience who have questions to please come to the microphone, because we'd love to take audience questions. I'm going to start it off though, with James. I have a question about your research.

It seemed as if your paper concluded that because there was not a long term change in consumer sensitive search behavior that basically privacy choice was working. And I wondered if that was necessarily the case. Do consumers really have a choice in terms, if they want to stay connected

price goes up, demand curve flipped down and that's what happens. And that's kind of how I view this.

KRISTIN KRAUSE COHEN: OK. Steve, I wanted to ask you about your research. What do you see as the next, what you hope to measure next with your tool. You did a lot of measurement and there was a lot discussed in your paper, but what do you see as the next area?

And for example, I noticed in your paper you looked at a lot of different kinds of canvasing, fingerprinting type of techniques. And this morning we heard from Aleksandra Korolova about the possibility of Bluetooth fingerprinting. Is that something that you would be able to look at with your tool? And what other areas would you want to look into?

STEVEN ENGLEHARDT: Yes sure. So I think one of the next steps coming is this notion of sensor based fingerprinting. You can imagine, if you want to re-ide

And some of them were asking consumers to change the settings. Were you able to measure if they were successful in that? If they were able to show the value proposition to consumers of changing those settings?

ZUBAIR SHAFIQ: So in our study we were not able to measure that, because we were doing an active study. So we did not have data from actual users. But there have been anecdotal reports that users do actually sometimes respond to these messages. And some fraction of users do actually whitelist the site or turn off their ad blockers.

So there is another thing which has also becoming, which has become increasingly popular. These are these alternate revenue models. So websites are recognizing that it's not just ads that they can use to monetize their services. Some of these websites are asking users to subscribe or

So it'll be interesting to see how that works out if users like it, if users are willing to pay per page visit instead of monthly fee for a specific site.

KRISTIN KRAUSE COHEN: Yes, we have an audience question.

BERIN SZOKA: Berin Szoka, Tech Freedom. It might be interesting, it might also be catastrophic. I mean you're talking about a model that has worked to deliver free content across the Internet, and that has allowed new start-ups and new media sources in the long tail to deploy and reach users.

So how do you think about, I know you can't study everything in one report, but how do you think about we could study in the future the effect on publishers, which in turn means the effect on revenue that's available for media outlets. That are dependent on the ad publishers for revenue.

So I personally think it's not a win-win or win-lose. It's not a win-lose situation. Everyone can benefit. But we need to make sure that we put users at the forefront of this conversation on any new revenue model that gets widely used.

JAMES COOPER: I just want to, I guess maybe follow up with what Berin says. I mean, I wond

payment scheme. You could also move back to more contextual ads. Or even, there's a bunch of

KRISTIN KRAUSE COHEN: We are running out of time, but I think we have time for maybe one, maybe two questions. Go ahead.

BRAD WELTMAN: Thank you, appreciate the research. I'm Brad Weltman with the Interactive

How about if we ask the questions and we can get their responses offline?

SPEAKER: I had a question for Professor Cooper. Running experiments in the wild is really complicated and I'm wondering how you took account for all the potentially confounding variables that could have influenced the results that you had?

Specifically, having done research, for example, looking at how the implementation of surveillance cameras might interact with crime patterns, one has to look at whether or not crime patterns change over time. So historically there's always a drop in March, for example.

One would potentially want to look not just at the search patterns on Google, but on Bing and on Yahoo and on other things to make sure that, in fact, you weren't just seeing a pattern that was happening globally. One might want to look at other things that were happening that might deter people from making a search request for a term like porn.

And so I'm just wondering if you could talk a little bit about how you're able to understand that something was not just a correlation, but there was actually some causal connection.

JAMES COOPER: I mean, since that was, can I respond publicly as opposed to--

KRISTIN KRAUSE COHEN: If you can do it in one minute.

JAMES COOPER: Yeah, I can. I used, I mean, I try to make causal claims. Like I said, it's a difference-in-differences. I have state fixed effects. I have week fixed effects. I have term fixed effects. I have term trends fixed effects.

I also look at different types of control, different windows of the control group to make sure that they're, I sort of look at the dummy interaction.

So I think, I mean this is all in my, I use a lot of econometrics, I'm not trying to use jargon, but I'm very aware of what it takes to make causal claims. And I think I control to the extent possible with everything I have.

Now using other websites or other search engines, I think that'd be great. They just don't have the data. And I am looking at the response to a Google privacy policy change. And so that's why I focused on Google.

I mean so so having data from other websites I think would be a fantastic control. I do do placebo checks, too. And like I said, I do do my randomized. Out of, in the distribution of 100 randomized runs that this regression leads me to believe the empirical distribution is kind of what I'm estimating with my point estimate.

KRISTIN KRAUSE COHEN: Thank you all so much. I'm sorry, Sebastian.

SEBASTIAN: I just have a very quick comment. Just for Zubair, I think what I'm always wondering is, if you consider the privacy policy to be a contract. Then what actually does this kind of self help lead to?

That's just one additional thought. And I know many people don't think of privacy policies that way, but might be worthwhile to do consider that as well.

KRISTIN KRAUSE COHEN: So we have a very short break. Please come back at 3:40. We have a great panel on information security and I'd like to thank our panelists for being here.