CAN CONGRESS — OR ROBERT MUELLER — ORDER FACEBOOK TO DIRECT ITS MACHINE LEARNING?

The other day I pointed out that two articles (WSJ, CNN) — both of which infer that Robert Mueller obtained a probable cause search warrant on Facebook based off an interpretation that under Facebook's privacy policy a warrant would be required — actually ignored two other possibilities. Without something stronger than inference, then, these articles do not prove Mueller got a search warrant (particularly given that both miss the logical step of proving that the things Facebook shared with Mueller count as content and not business records).

In response to that and to this column arguing that Facebook should provide more information, some of the smartest surveillance lawyers in the country discussed what kind of legal process would be required, but were unable to come to any conclusions.

Last night, WaPo published a story that made it clear Congress wanted far more than WSJ and CNN had suggested (which largely fell under the category of business records and the ads posted to targets, the latter of which Congress had been able to see but not keep). What Congress is really after is details about the machine learning Facebook used to identify the malicious activity identified in April and the ads described in its most recent report, to test whether Facebook's study was thorough enough.

> A 13-page "white paper" that Facebook published in April drew from this fuller internal report but left out critical details about how the Russian operation worked and how Facebook discovered it,

> according to people briefed on its contents.
>
> Investigators believe the company has not fully examined all potential ways that Russians could have manipulated Facebook's sprawling social media platform.
>
> [snip]
>
> Congressional investigators are questioning whether the Facebook review that yielded those findings was sufficiently thorough.
>
> They said some of the ad purchases that Facebook has unearthed so far had obvious Russian fingerprints, including Russian addresses and payments made in rubles, the Russian currency.
>
> Investigators are pushing Facebook to use its powerful data-crunching ability to track relationships among accounts and ad purchases that may not be as obvious, with the goal of potentially detecting subtle patterns of behavior and content shared by several Facebook users or advertisers.
>
> Such connections — if they exist and can be discovered — might make clear the nature and reach of the Russian propaganda campaign and whether there was collusion between foreign and domestic political actors. Investigators also are pushing for fuller answers from Google and Twitter, both of which may have been targets of Russian propaganda efforts during the 2016 campaign, according to several independent researchers and Hill investigators.
>
> "The internal analysis Facebook has done [on Russian ads] has been very helpful, but we need to know if it's complete," Schiff said. "I don't think Facebook fully knows the answer yet."

> [snip]
>
> In the white paper, Facebook noted new techniques the company had adopted to trace propaganda and disinformation.
>
> Facebook said it was using a data-mining technique known as machine learning to detect patterns of suspicious behavior. The company said its systems could detect "repeated posting of the same content" or huge spikes in the volume of content created as signals of attempts to manipulate the platform.

The push to do more — led largely by Adam Schiff and Mark Warner (both of whom have gotten ahead of the evidence at times in their respective studies) — is totally understandable. We need to know how malicious foreign actors manipulate the social media headquartered in Schiff's home state to sway elections. That's presumably why Facebook voluntarily conducted the study of ads in response to cajoling from Warner.

But the demands they're making are also fairly breathtaking. They're demanding that Facebook use its own intelligence resources to respond to the questions posed by Congress. They're also demanding that Facebook reveal those resources to the public.

Now, I'd be surprised (pleasantly) if either Schiff or Warner made such detailed demands of the NSA. Hell, Congress can't even get NSA to count how many Americans are swept up under Section 702, and that takes far less bulk analysis than Facebook appears to have conducted. And Schiff and Warner surely would never demand that NSA reveal the extent of machine learning techniques that it uses on bulk data, even though that, too, has implications for privacy and democracy (America's and other countries'). And yet they're asking Facebook to do just that.

And consider how two laws might offer guidelines, but (in my opinion) fall far short

of authorizing such a request.

There's Section 702, which permits the government to oblige providers to provide certain data on foreign intelligence targets. Section 702's minimization procedures even permit Congress to obtain data collected by the NSA for their oversight purposes.

Certainly, the Russian (and now Macedonian and Belarus) troll farms Congress wants investigated fall squarely under the definition of permissible targets under the Foreign Government certificate. But there's no public record of NSA making a request as breathtaking as this one, that Facebook (or any other provider) use its own intelligence resources to answer questions the government wants answered. While the NSA does draw from far more data than most people understand (including, probably, providers' own algorithms about individually targeted accounts), the most sweeping request we know of involves Yahoo scanning all its email servers for a signature.

Then there's CISA, which permits providers to voluntarily share cyber threat indicators with the federal government, using these definitions:

> (A) IN GENERAL.—Except as provided in subparagraph (B), the term "cybersecurity threat" means an action, not protected by the First Amendment to the Constitution of the United States, on or through an information system that may result in an unauthorized effort to adversely impact the security, availability, confidentiality, or integrity of an information system or information that is stored on, processed by, or transiting an information system.
>
> (B) EXCLUSION.—The term "cybersecurity threat" does not include any action that solely involves a violation of a consumer term of service or a consumer licensing agreement.
>
> (6) CYBER THREAT INDICATOR.—The term

> “cyber threat indicator” means information that is necessary to describe or identify—
>
> (A) malicious reconnaissance, including anomalous patterns of communications that appear to be transmitted for the purpose of gathering technical information related to a cybersecurity threat or security vulnerability;
>
> (B) a method of defeating a security control or exploitation of a security vulnerability;
>
> (C) a security vulnerability, including anomalous activity that appears to indicate the existence of a security vulnerability;
>
> (D) a method of causing a user with legitimate access to an information system or information that is stored on, processed by, or transiting an information system to unwittingly enable the defeat of a security control or exploitation of a security vulnerability;
>
> (E) malicious cyber command and control;
>
> (F) the actual or potential harm caused by an incident, including a description of the information exfiltrated as a result of a particular cybersecurity threat;
>
> (G) any other attribute of a cybersecurity threat, if disclosure of such attribute is not otherwise prohibited by law; or
>
> (H) any combination thereof.

Since January, discussions of Russian tampering have certainly collapsed Russia’s efforts on social media with their various hacks. Certainly, Russian abuse of social media has been treated as exploiting a vulnerability. But

none of this language defining a cyber threat indicator envisions the malicious use of legitimate ad systems.

Plus, CISA is entirely voluntary. While Facebook thus far has seemed willing to be cajoled into doing these studies, that willingness might change quickly if they had to expose their sources and methods, just as NSA clams up every time you ask about *their* sources and methods.

Moreover, unlike the sharing provisions in 702 minimization procedures, I'm aware of no language in CISA that permits sharing of this information with Congress.

Mind you, part of the problem may be that we've got global companies that have sources and methods that are as sophisticated as those of most nation-states. And, inadequate as they are, Facebook is hypothetically subject to more controls than nation-state intelligence agencies because of Europe's data privacy laws.

All that said, let's be aware of what Schiff and Warner are asking for, however justified it may be from a investigative standpoint. They're asking for things from Facebook that they, NSA's overseers, have been unable to ask from NSA.

If we're going to demand transparency on sources and methods, perhaps we should demand it all around?