

Shadow profiles are the biggest flaw in Facebook's privacy defense

By [Russell Brandom](#) | [@russellbrandom](#) | Apr 11, 2018, 3:53pm EDT

Part of

The Cambridge Analytica scandal

Called before Congress this week, Mark Zuckerberg tried to present Facebook's approach to user data as open and transparent. In question after question, he focused on the privacy choices available to users, and their ownership over all the data they share — and it wasn't all wrong. Facebook has data because users share it (mostly). Users control that data and can review it or delete it whenever they want (with a few exceptions). And if you delete your account, (almost) all of that data will disappear from Facebook's servers within 90 days. None of it's false, but as the parentheses should tell you, it is incomplete — and by the second day of hearings, members of Congress were starting to catch on.

The most powerful example came from Rep. Ben Luján (D-NM), who confronted Zuckerberg on the company's use of shadow profiles — a term for non-user data collection that Zuckerberg was apparently unfamiliar with.

"It's been admitted that you do collect data points on non-Facebook users," Luján asked. "So my question is, can someone who does not have a Facebook account opt out of Facebook's involuntary data collection?"

"Congressman, anyone can opt out of any data collection for ads, whether they use our services or not," Zuckerberg said. "But in order to prevent people from scraping public information, we need to know when someone is trying to repeatedly access our services."

"MY QUESTION IS, CAN SOMEONE WHO DOES NOT HAVE A FACEBOOK ACCOUNT OPT OUT OF DATA COLLECTION?"

“You’ve said everyone controls their data, but you’re collecting data on people who are not even Facebook users, who never signed a consent or privacy agreement and you’re collecting their data,” Luján continued. “And you’re directing people who don’t have a Facebook page to sign up for Facebook in order to get their data.”

In the exchange, Luján seized on a serious flaw in Zuckerberg’s consent-driven vision of Facebook, one that could have regulatory consequences in the months to come. The fact is, even if you’ve never signed up for Facebook, the company still has a general sense of who you are, gathered through uploaded contact lists, photos, or other sources.

Facebook’s collection of data on non-Facebook users opens up a world of questions about what data is and isn’t covered by Zuckerberg’s vision of user consent and control. Zuckerberg repeatedly said that Facebook deletes all your profile data if you delete your account, but what about shadow profile data that pre-dated your account? Zuckerberg also cited the ability to download your Facebook data, but not only would a non-Facebook user not have access to that data trove, the download tool omits data Facebook clearly collects and uses, whether it’s data from [Facebook’s analytics Pixel](#) or location data pulled from a phone.

The most concrete example of a shadow profile comes from Facebook’s People You May Know service, [studied in detail by Kashmir Hill at Gizmodo](#). Even if you’ve never signed up for Facebook, you’ve appeared in the contacts lists of people who did. When users connect their email account or texting data with Facebook, countless non-users are swept up. Instead of discarding their information, Facebook keeps non-user data attached to something Hill calls a shadow profile — a reliable bank of information held in reserve so that, if you ever *do* sign up for Facebook, the company will know exactly who to recommend as friends.

“EVEN AFTER I’VE LOGGED OUT OF FACEBOOK, YOU GUYS STILL HAVE THE ABILITY TO FOLLOW MY INTERACTIONS ON THE WEB”

If that were all, it would be easy enough to wave away, but shadow profiles have become a stand-in for all the data that doesn’t make it into a person’s official profile. Facebook says that when you delete your account, all your data is gone from company servers within 90 days — but it’s hard to believe that applies to shadow profile data,

which exists even without an official profile. Today, Zuckerberg assured Congress that Facebook's data download tool included all the information on a given user — but it's missing much of the web-based tracking that Facebook performs through the Like button embed, only showing the general interest categories that are created as a result of that data. How can we be sure there isn't similar data being collected on non-users, or that it doesn't remain associated with them after deleting their account?

Rep. Kurt Schrader (D-OR) tried to get an answer from Zuckerberg about the extent of Facebook's tracking of users off the platform, but the answer was ambiguous.

"It's my understanding based on the testimony here today that even after I've logged out of Facebook, you guys still have the ability to follow my interactions on the web," Schrader asked Zuckerberg.

"YOU HAVE TRACKERS ALL OVER THE WEB."

"You have control over what we do for ads and the information collections based on that," Zuckerberg replied. "On security, there may be specific things about how you use Facebook, even if you're not logged in, that we keep track of to make sure you're not abusing the systems."

This zone of questioning is particularly tricky for Facebook because, as Luján pointed out, all of Facebook's controls count on a person having a Facebook profile. You can't change your ad settings or download your data unless you're a Facebook user, even though we know the company is still holding some information relating to you. That catch-22 may soon cause problems in Europe — where the GDPR requires data-portability for all citizens, not just Facebook users.

In the meantime, Facebook's data protection tools mostly serve to distract users from the more aggressive data collection happening behind the scenes. That point was driven home by a heated speech from Rep. Debbie Dingell (D-MI) toward the end of the hearing, taking Zuckerberg to task for a lack of information.

"As CEO, you didn't know some key facts," Dingell told Zuckerberg. "You didn't know what a shadow profile was. You didn't know how many apps you need to audit. You did not know how many other firms have been sold data by Cambridge Analytica... You

don't even know all the different kinds of information Facebook is collecting from its users."

"Here's what I do know," Dingell continued. "You have trackers all over the web. On practically every website, we all see the Facebook like or share buttons, and with the Facebook Pixel, people may not even see that Facebook logo. It doesn't matter whether you have a Facebook account. Through those tools, Facebook is able to collect information from all of us."

THE CAMBRIDGE ANALYTICA SCANDAL

WHAT HAPPENED

Facebook suspended Donald Trump's data operations team for misusing people's personal information

Academic who collected 50 million Facebook profiles: 'We thought we were doing something normal'

The shady data-gathering tactics used by Cambridge Analytica were an open secret to online marketers. I know, because I was one

Cambridge Analytica's Facebook data abuse shouldn't get credit for Trump

Cambridge Analytica reportedly still hasn't deleted Facebook user data as promised

Facebook says Cambridge Analytica data collection affected nearly twice as many users as previously thought

Mark Zuckerberg's personal Facebook profile was part of the Cambridge Analytica leak

FACEBOOK REACTS

8 

THE WORLD REACTS

13 

THE APOLOGY TOUR

6 

ZUCKERBERG TESTIFIES

14 

WHAT YOU CAN DO

5 
