

Google's File on Me Was Huge. Here's Why It Wasn't as Creepy as My Facebook Data.

Chen, Brian X . New York Times (Online) , New York: New York Times Company. May 16, 2018.

[ProQuest document link](#)

ABSTRACT (ENGLISH)

Google collects far more information about us than Facebook. But the Google data that our tech columnist downloaded on himself contained fewer surprises and was more easily deleted.

FULL TEXT

Google has far more data about us than Facebook. Yet unlike Mark Zuckerberg's social networking empire, which has been under fire for improperly leaking user data, Google has sidestepped controversy .

You may wonder: Why is that? After all, we turn to Google for not only our internet searches but also for our emails, calendaring, maps, photo uploads, video streaming, mobile phones and web browsers. That's far more pervasive than the baby photos and comments that we post on Facebook.

To help get an answer, I downloaded a copy of all of the information that Google has on me. Then I compared the trove to all the data that I already knew Facebook had obtained on me .

What I found was that my Google data archive was much larger than my Facebook file —about 12 times larger, in fact —but it was also packed with fewer unpleasant surprises.

Most of what I saw in my Google file was information I already knew I had put in there, like my photos, documents and Google emails, while my Facebook data contained a list of 500 advertisers with my contact information and a permanent record of friends I thought I had "deleted" years ago, among other shockers.

Whenever I was perturbed by parts of my Google data, like a record of the Android apps I had opened over the past several years, I was relieved to find out I could delete the data. In contrast, when I downloaded my Facebook data, I found that a lot of what I saw could not be purged.

Aaron Stein, a Google spokesman, said the company had spent many years developing tools for people to download their information.

"It should be easy for people to understand and control their Google data," he said. "We encourage everyone to use these tools so they can make the privacy choices that are right for them."

That's not to say that we should be complacent. Tech companies like Google and Facebook have an incredible amount of power over us that only increases with the more that they know. So downloading and analyzing your Google data, and determining what information you want to keep around or delete, is an exercise I highly recommend. Here's how I did it —and what I learned.

Wading Through Google Takeout

The tool for downloading your Google data is called Takeout, which was released in 2008. Go to [google.com/takeout](https://www.google.com/takeout) and select the information you want to download. You can choose everything or home in on certain things, like your location history from Google Maps, your email conversations in Gmail, your viewing history on YouTube and photos you have uploaded into Google Photos.

If you download your whole archive, your file will probably be enormous. Mine was eight gigabytes, enough to hold about 2,000 hours of music. After I requested my archive, it took Google about half a day to send an email with links to download my files.

Here's what jumped out to me:

•The most noteworthy folder is labeled My Activity, which is an overview of what you have done on Google's products, including Android, Google Maps and Google News.

Inside My Activity is a subfolder labeled Ads. This record contained a history of many websites I had visited, including those I had reached without the help of Google.com. Even sites I opened through Twitter or links I clicked on through a text message were recorded in the Ads folder.

What gives? Google said that many web articles load advertisements served through its ad network, and when you visit sites loading Google ads, you are contributing to an advertisement-related profile that Google is building about you. That's why it's logged in the My Activity folder.

In my conversations with Google, the company argued that it was better to be transparent about the information being collected as opposed to not showing it at all.

Brian Fitzpatrick, the former Google manager who led the team that created Takeout, put it another way.

"Companies are gathering this data about you," he said. "This is just an honest way to look at it."

My take: It is helpful to see that this data is collected. But it should be labeled precisely —like "Pages We Know You Visited" —so people can find this data more easily and decide whether to delete it.

•A subfolder labeled Android contained a detailed history of the Android apps I had opened over the past three years, including the time and date I had launched each app. For example, my log showed I opened the Instagram app in March, the Gmail app in December 2015 and the Google Play Music app in January 2016.

Google uses this information for a feature called app suggestions. The company studies which apps you use, and how often and when you open them, to recommend apps you might want to use throughout the day. For example, if you regularly open Instagram during your lunch break, Google will show a shortcut for the Instagram app at around 12:30 p.m. in a list of suggested apps.

That is a thoughtful feature, but it gave me pause. That level of logging is almost as creepy as a company monitoring all of my keystrokes. Also, retaining this app data for several years feels like an unnecessarily long time. I ultimately opted to turn app suggestions off.

•Many files in my archive were odd formats that were not easy to open or read. For example, some files included the extension .JSON. My Google Maps location history was stored in a .JSON file, and it displayed an unintelligible list of GPS coordinates and time stamps.

Google explained that Takeout was designed for people to be able to easily remove their data from Google and use it elsewhere. Files like those with the .JSON extension are common formats designed to be machine readable so that other programs and tools can make use of the data, according to Google.

That makes sense—but our data should be readable by us, too.

Deleting Objectionable Data

After poring through your Google file, ask yourself what personal data you are uncomfortable with having the company keep. Everybody is going to have a different answer.

For me, I was troubled by Google keeping a history of websites I had visited even when I wasn't using Google products. I also found the company's log of my Android apps usage overly intrusive.

Once you have determined that, then you can delete the data. The place to start is the My Activity tool, located at myactivity.google.com.

I opted to purge the entire history for Ads, where my web-browsing activity was being tracked even when I was not using Google products. I also deleted all my Android data. Once I got going, I also deleted my history of requests made to Google's voice assistant and the web-browsing histories for sites I visited through Google News and Google Chrome.

But that all raises a larger question. What does it mean to delete something from Google? Is it just hidden from plain sight or actually expunged?

A Google spokesman referred me to a webpage summarizing the company's data retention policy, which says different types of data are held onto for different periods of time until it is removed from Google's servers or "retained in an anonymized form."

In other words, some of the data you eliminate will actually be deleted eventually, and some will not. That may not offer much solace. But short of ceasing to use the web entirely, occasionally purging parts of your account data, much like you would do with unwanted junk in your home, is the best you can do.

Credit: Brian X. Chen

DETAILS

Subject:	Electronic mail systems; Social networks; Archives & records; Privacy
Location:	United States--US
People:	Zuckerberg, Mark
Company / organization:	Name: Facebook Inc; NAICS: 518210, 519130; Name: Google Inc; NAICS: 334310, 519130
Identifier / keyword:	Google Inc Facebook Inc Social Media Computers and the Internet Online Advertising Android (Operating System) Mobile Applications Archives and Records Corporate Social Responsibility
Publication title:	New York Times (Online); New York
Publication year:	2018
Publication date:	May 16, 2018
Section:	technology
Publisher:	New York Times Company
Place of publication:	New York
Country of publication:	United States, New York
Publication subject:	General Interest Periodicals--United States
Source type:	Blog, Podcast, or Website
Language of publication:	English
Document type:	News
ProQuest document ID:	2039491472
Document URL:	https://www.proquest.com/blogs-podcasts-websites/google-s-file-on-me-was-huge-here-why-wasn-t-as/docview/2039491472/se-2?accountid=15006
Copyright:	Copyright 2018 The New York Times Company
Last updated:	2018-06-13

Database:

Global Newsstream

LINKS

[Linking Service](#)

Database copyright © 2021 ProQuest LLC. All rights reserved.

[Terms and Conditions](#) [Contact ProQuest](#)