

WEB SITES google, |

'Liberating' Your Data from Google, and What That Really Means

131

Tweet

38

By [Christopher Null](#), PCW

May 21, 2012 6:10 PM

If you listen to the grov company, Google's no becoming [increasingly](#)

44

Share

most recent move--a [sweeping change](#) that consolidated most of its privacy policies under a single umbrella--

immediately drew umbrage from critics who felt that Google was on its way to taking all the data it has collected from its users through its dozens of services and building an exhaustive dossier on each of us that would be used mercilessly in efforts to sell us things.

Google wants you to keep using Search, Docs, and Google+, so it's trying to play nice, and last June Google introduced a service designed to let you see, in part at least, what Google knows about you with a single click.

Called [Google Takeout](#), the service is so simple that it is completely undocumented when you visit the site. You sign in, and then see an offer to "Download an archive of your data from" a variety of services (outlined below), and that's it. You can grab it all in one click, or choose specific services from which to download, but unless your usage of these services is exhaustive (think thousands of Google Docs or Picasa photos), the one-click approach is easiest.





Takeout

All of your data

Choose services

Downloads

Download an archive of your data from:

 +1s

 Buzz

 Circles

 Contacts

 Docs

 Picasa Web Albums

 Profile

 Stream

 Voice

CREATE ARCHIVE

Getting It To Go

Here's what Takeout currently offers:

- A list of URLs to all the +1s you've handed out.
- Your Google Buzz history, presuming you have one.
- A list of contacts from your Circles in Google+.
- A list of the contacts you have saved in Gmail. (These are kept separate from your Circles contacts.)
- Copies of all the Google Docs you've uploaded.
- Copies of any photos you've uploaded to Picasa. (These may include photos

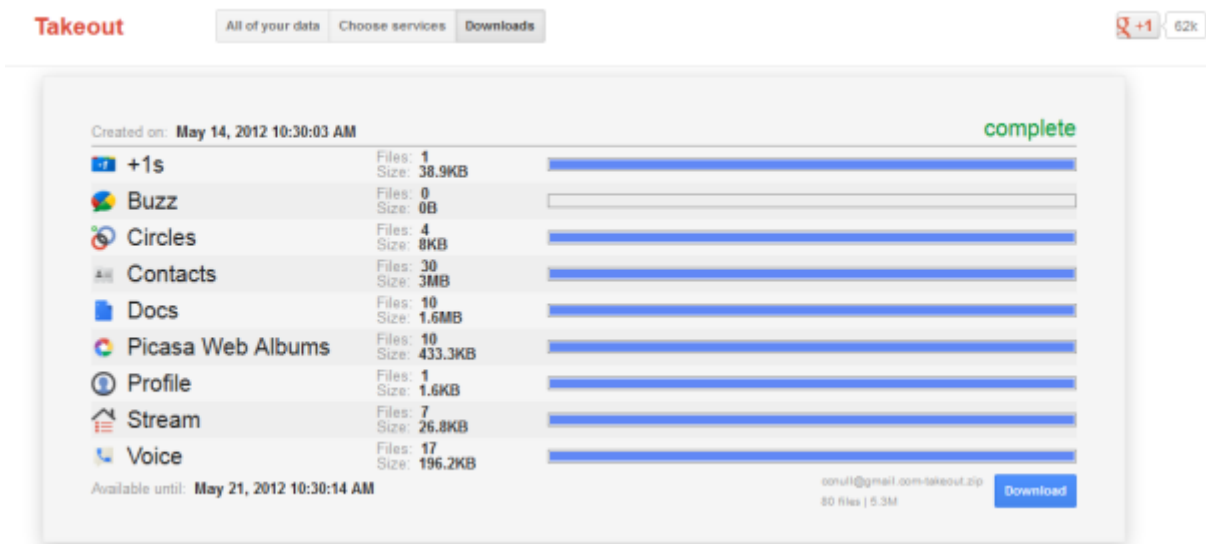
uploaded for use on a Blogger site, if you've ever had one.)

- Some basic information about the personal data you include in your Google+ Profile
- Links to each entry you've personally shared on your Google+ Stream. (Other people's streams that show up in your feed are not included.)
- Your full Google Voice log, including a list of all attempted and completed calls and texts, MP3s of each voice mail, and Google's transcript of each message.

Everything arrives in a single zipped file that you unpack, revealing a separate folder for each Google service. The formatting of this material can be inconsistent. [Google Voice messages](#) are saved as individual HTML and MP3 files, but your +1 bookmarks are amalgamated into a single file. Picasa photos are well organized into folders, but Google Docs are delivered en masse regardless of how you use Collections on the site.

Circles and Contacts information comes in the form of several VCF files, each containing information for contacts split into each of the default categories for Google+. These are easiest to open with Windows Contacts (right-click on a file and you'll see the option; no, I didn't know this application existed either). Information on my Circles contacts was limited to a name and a malformed link to their Google+ profile.

Finally, your Profile info is delivered as a JSON file, coded in a type of Google-centric JavaScript that is not readily openable, but you can get the gist of it by opening the file with Firefox.



What Your Takeout Order Does Not Include

What's most surprising about Takeout isn't how exhaustive this data is, but rather how much of your Google life it completely excludes. Although Google has said that it will continue to add services to Takeout, here's a (partial) list of what you *don't* get now with the Takeout system.

- Your Google Search history.
- Your Google Talk chat history.
- Google Wallet and Google Checkout details, including credit card information and a history of purchases.
- YouTube materials, including videos you liked, shared, or uploaded.
- Posts created with Blogger, or comments you've left on Blogger sites.
- Google Calendar entries.
- Google Health data.
- Bookmarks stored or synced with Chrome or the Google Toolbar.
- Google Latitude location information.
- Anything related to your Android phone, including your account or your Android Market downloads.
- Anything involving Orkut, AdWords, Google Finance, and more.

That list surprised me, not just because it's so long, but also because Google does retain data for most (if not all) of those services. Your Google Search history can be

[accessed here](#), for example. Why doesn't Takeout let you download this information instead of shipping you off to another site?

If you want to get an offline copy of any of this information, your best bet is to check out the comprehensive list of how-tos at the [Data Liberation Front](#), managed by the group of Google engineers that coded the Google Takeout service. Here you'll find detailed instructions on how to manually get your data out of another two dozen Google-operated services not covered by Takeout.

Liberation vs. Deletion

It's important to remember that with Google Takeout you are getting a *copy* of the information stored on Google's servers, and are not [removing the originals from Google's clutches](#). There's no way to delete anything at all via Google Takeout.

[If you want to delete information from Google](#), you'll need to visit each service you use and delete the data or the account manually. In Blogger, for example, that means visiting the blog administration tool and using the "Delete blog" link to remove it from the Web. (There's a "nuclear option," too, available at the bottom of your [Settings](#) page.)

One of the most popular subjects for deletion is your [Google web search history](#). You can turn history recording off or on [here](#), remove specific history entries, or delete your entire history. Most of Google's services offer ways to delete accounts, and Google's [Privacy Policy](#) offers more detail on what exactly this means and entails:

Whenever you use our services, we aim to provide you with access to your personal information. If that information is wrong, we strive to give you ways to update it quickly or to delete it – unless we have to keep that information for legitimate business or legal purposes.

We may reject requests that are unreasonably repetitive, require disproportionate technical effort (for example, developing a new system

or fundamentally changing an existing practice), risk the privacy of others, or would be extremely impractical (for instance, requests concerning information residing on backup tapes).

Where we can provide information access and correction, we will do so for free, except where it would require a disproportionate effort. We aim to maintain our services in a manner that protects information from accidental or malicious destruction. Because of this, after you delete information from our services, we may not immediately delete residual copies from our active servers and may not remove information from our backup systems.

Those concerned with Google's ability to keep tabs on you may want to pay special attention to a few of those clauses: Notably that Google can reject requests to delete information that "require disproportionate technical effort" and, more importantly, that backup copies of your data are not likely to be deleted promptly, if ever.

Takeout: Still Too Limited

So what does all of this really mean?

To date, user commentary has been surprisingly muted about Takeout. Those who have written about it mainly seem thrilled to have one-click access to their Google Voice records, their Google Docs, and their online contacts. In these respects, Takeout is actually a useful tool: Downloading this information piecemeal is a pain, and Google Takeout makes it considerably easier. Painless, even.

But even at a year old, Takeout is still a long way from offering users a legitimate way to get a handle on how exhaustive the information Google has about them really is. The number of services included in Takeout is paltry compared to the vast number of offerings that Google has available, particularly given that those services typically make this information available directly to the user. If Google can figure out a way to consolidate its myriad privacy policies, surely it can figure out a way to

consolidate the downloading of collected user information, too. As engineering challenges go, this doesn't seem like a toughie.

Ultimately, Takeout is a good first step toward giving users more transparency about what the company does with their data, but if Google wants to prove it is serious about privacy, Takeout needs to be radically expanded--and to imbue users with the ability to delete materials they don't want Google to be sitting on.