

Your Memories. Their Cloud.

Google, Apple and Meta offer near-limitless digital basements in which to store photos, videos and important documents, but you should keep a copy of what you hold most dear.



By Kashmir Hill

Published Dec. 31, 2022 Updated June 20, 2023

I have many fears as a mother. My kindergarten-age daughter recently learned a game on the school bus called “Truth or Force.” My youngest refuses to eat almost anything but Kraft Mac and Cheese. Added to the list this year, alongside outside influences and health concerns, is the possibility that my daughters could inadvertently lock me out of my digital life.

That’s what happened to a mother in Colorado whose 9-year-old son used her old smartphone to stream himself naked on YouTube, and a father in San Francisco whose Google account was disabled and deleted because he took naked photos of his toddler for the doctor.

I reported on their experiences for The New York Times, and as I talked to these parents, who were stunned and bereft at the loss of their emails, photos, videos, contacts and important documents spanning decades, I realized I was similarly at risk.

I am “cloud-complacent,” keeping my most important digital information not on a hard drive at home but in the huge digital basement provided via technology companies’ servers. Google gives all users 15 gigabytes free, a quarter of what comes standard on an Android phone, and I have not managed to max it out in 18 years of using the company’s many services.

I did fill up Apple's free 5 GB, so I now pay \$9.99 a month for additional iCloud storage space. Meta has no max; like scrolling on Instagram, the allowed space is infinite.

If I were suddenly cut off from any of these services, the data loss would be professionally and personally devastating.

As a child of the 1980s, I used to have physical constraints on how many photos, journals, VHS tapes and notes passed in seventh grade that I could reasonably keep. But the immense expanse and relatively cheap rent of the so-called cloud has made me a data hoarder. Heading into 2023, I set out to excavate everything I was storing on every service, and find somewhere to save it that I had control over. As I grappled with all the gigabytes, my concern morphed from losing it all to figuring out what was actually worth saving.

Data Harvesting



The author's sister and father tailgating before an ice hockey game in November 2007. Kashmir Hill

I find nearly 100 photos from one November night 15 years ago, out with my family at a Tampa Bay Lightning game when my sisters and I were home for the holidays. We're tailgating with a mini-keg of Heineken. My dad is posing by the car, making a funny face at the ridiculousness of a parking garage party. Then, we're posing in the stadium with the hockey rink in the background, toasting with a stranger we sat next to. Had we bonded with him during an especially close third period? The metadata in the Google Photos jpg file didn't say.

The photos transported me back to a tremendously fun evening that I had all but forgotten. Yet I wondered how there could be so many photos from just one night. How do I decide which to keep and which to get rid of?

This kind of data explosion is a result of economics, said Brewster Kahle, founder of the Internet Archive, a nonprofit library based in San Francisco that saves copies of websites and digitizes books and television shows. Taking a photo used to be expensive because it involved film that needed to be developed.

"It cost a dollar every time you hit a shutter," Mr. Kahle said. "That's no longer the case so we hit the shutter all the time and keep way, way too much."

I had captured the 2007 evening in Tampa pre-smartphone on a digital Canon camera that had a relatively small memory card that I regularly emptied into Google Photos. I found more than 4,000 other photos there, along with 10 gigabytes of data from Blogger, Gmail, Google Chat and Google Search, when I requested a copy of the data in my account using a Google tool called Takeout.

I just pressed a button and a couple of days later got my data in a three-file chunk, which was great, though some of it, including all my emails, was not human-readable. Instead, it came in a form that needed to be uploaded to another service or Google account.

According to a company spokesman, 50 million people a year use Takeout to download their data from 80 different Google products, with 400 billion files exported in 2021. These people may have had plans to move to a different service,

simply wanted their own copy or were preserving what they had on Google before deleting it from the company's servers.

Takeout was created in 2011 by a group of Google engineers who called themselves the Data Liberation Front. Brian Fitzpatrick, a former Google employee in Chicago who led the team, said he thought it was important that the company's users have an easy "off ramp" to leave Google and take their data elsewhere. But Mr. Fitzpatrick said he worried that when people store their digital belongings on a company's server, they "don't think about it or care about it."

Some of my data landlords were more accommodating than others. Twitter, Facebook and Instagram offered Takeout-like tools, while Apple had a more complicated data transfer process that involved voluminous instructions and a USB cable.

The amount of data I eventually pulled down was staggering, including more than 30,000 photos, 2,000 videos, 22,000 tweets, 57,000 emails, 15,000 pages of old Google chats and 16,000 pages of Google searches going back to 2011.

It was such an overwhelming amount of digital stuff that I wasn't surprised to see that Google had hired Marie Kondo as a spokeswoman for the paid version of its storage service — starting at \$1.99 per month for 100 GB. Ms. Kondo suggested better labeling and organization of emails, photos and documents to make it "easy to find the memories that spark joy."

The Missing



The author took this terrible photo at a museum exhibit in 2011 and held onto it for more than a decade. Kashmir Hill

The trove of data brought forgotten episodes of my life back in vivid color. A blurry photo of my best friend's husband with a tiny baby strapped to his chest, standing in front of a wall-sized Beetlejuician face, made me recall a long-ago outing to a Tim Burton exhibit at a museum in Los Angeles. I don't remember what I learned

about the gothic filmmaker, but I do remember my friends' horror when their weeks-old son, now 11, had a blowout and they had to beg a comically oversized diaper from a stranger.

The granularity of what was in my digital archive accentuated the parts of my life that were missing entirely: emails from college in a university-provided account that I hadn't thought to migrate; photos and videos I took on an Android phone that I backed up to an external hard drive that has since disappeared; and stories I'd written in journalism school for publications that no longer exist. They were as lost to me as the confessional journal I once left in the seatback of a plane. The idea that information, once digitized, will stick around forever is flawed.

"We often say the internet never forgets, but it does," said the web historian Ian Milligan. Companies shut down, as happened to GeoCities, an early, popular place for hosting personal websites, or a service cuts back on the amount of free storage it's offering, as when the new owner of Flickr announced in 2019 that free accounts had a limit of 1,000 photos and anything more would be deleted.

Margot Note, an archivist, said her profession thinks a lot about the accessibility of the medium on which data is stored, given the challenge of recovering videos from older formats such as DVDs, VHS tapes and reel film. Ms. Note asks the kinds of questions most of us don't: Will there be the right software or hardware to open all our digital files many years from now? With something called "bit rot" — the degradation of a digital file over time — the files may not be in good shape.

Individuals and institutions think that when they digitize material it will be safe, she said. "But digital files can be more fragile than physical ones."

Where to Put It



The author on a camel in Egypt with a friend in 2007.

Kashmir Hill

Once I assembled my data Frankenstein, I had to decide where to put it. More than a decade ago, pre-cloud complacency, I would regularly back my stuff up to a hard drive that I probably bought at Best Buy. Digital self-storage has gotten more complex as I discovered when I visited the DataHoarder subreddit. Posts there with technical advice for the best home setup were jargon-filled to the point of incomprehension for a newbie. A sample post: “Started with single bay Synology Nas and recently built a 16TB unRAID server on a xeon 1230. Very happy with result.”

I felt as if I'd landed on an alien planet so I turned instead to professional archivists and tech-savvy friends. They recommended two \$299 12-terabyte hard drives, one of which should have ample room for what I have now and what I will create in the future, and another to mirror the first, as well as a \$249 NAS, or network-attached storage system, to connect to my home router, so I could access the files remotely and monitor the health of the drives.

Archivists regularly cited the “3-2-1 rule”: three copies of everything, two copies on different cloud services and one at home. Some also said to keep yet another copy “offsite,” i.e. at a relative’s house or in a bank lockbox, depending on your level of paranoia. History is awash in tales of lost data, including the burning of invaluable master recordings of famous musicians in a Universal Studios fire. John Markoff, a technology journalist who writes for The Times, mined the extensive personal archives of the internet pioneer Stewart Brand for a biography. He found that even Mr. Brand, who meticulously preserved his communications, was missing several years of early emails because of the loss of back-up tapes and had hundreds of thousands of others on an old Macintosh that were a jumble of data that was largely impossible to read.

Getting all your data and figuring out how to securely store it is cumbersome, complicated and costly. There’s a reason most people ignore all their stuff in the cloud.

What to Keep

I noticed a philosophical divide among the archivists I spoke with. Digital archivists were committed to keeping *everything* with the mentality that you never know what you might want one day, while professional archivists who worked with family and institutional collections said it was important to pare down to make an archive manageable for people who look at it in the future.

“It’s often very surprising what turns out to matter,” said Jeff Ubois, who is in the first camp and has organized conferences dedicated to personal archiving.

He brought up a historical example. During World War II, the British war office asked people who had taken coastal vacations to send in their postcards and photographs, an intelligence-gathering exercise to map the coastline that led to the selection of Normandy as the best place to land troops.

Mr. Ubois said it's hard to predict the future uses of what we save. Am I socking this away just for me, to reflect on my life as I age? Is it for my descendants? Is it for an artificial intelligence that will act as a memory prosthetic when I'm 90? And if so, does that A.I. really need to remember that I Googled "starbucks ice cream calorie count" one morning in January 2011?

Pre-internet, we pared down our collections to make them manageable. But now, we have metadata and advanced search techniques to sort through our lives: timestamps, geotags, object recognition. When I recently lost a close relative, I used the facial recognition feature in Apple Photos to unearth photos of him I'd forgotten I'd taken. I was glad to have them, but should I keep all the photos, even the unflattering ones?

Bob Clark, the director of archives at the Rockefeller Archive Center, said that the general rule of thumb in his profession is that less than 5 percent of the material in a collection is worth saving. He faulted the technology companies for offering too much storage space, eliminating the need for deliberating over what we keep.

"They've made it so easy that they have turned us into unintentional data hoarders," he said.

The companies try, occasionally, to play the role of memory miner, surfacing moments that they think should be meaningful, probably aiming to increase my engagement with their platform or inspire brand loyalty. But their algorithmic archivists inadvertently highlight the value of human curation.

Recently, my iPhone served me "Waterfalls over the years," which, as promised, featured a slide show with instrumental music and photos of myself and others in front of a random assortment of waterfalls. Like the British war office during World War II, the technology saw the backdrop as the star of the show.

"I don't think we can simply rely on the algorithms to help you decide what's important or not," Mr. Clark said. "There need to be points of human intervention and judgment involved."

Paring It Down



When deciding what to delete or keep, screenshots might be the first thing to go.
Kashmir Hill

Rather than just keeping a full digital copy of everything, I decided to take the archivists' advice and pare it down somewhat, a process the professionals call appraisal. An easy place to start was the screenshots: the QR codes for flights long ago boarded, privacy agreements I had to click to use an app, emails that were best forwarded to my husband via text and a message from Words With Friends that "nutjob" was not an acceptable word.

There were some clear keepers, a selfie I took in Beijing with the artist Ai Weiwei in April 2015; a video of my eldest daughter's first steps in December 2017; and a shot of me on a camel in front of the Giza Pyramids in 2007, a photo I had purposely

staged to recreate one we had on my childhood fridge of my great-grandmother in the same place doing the same thing, but with a disgruntled expression on her face.

Then there's the stuff I'm ambivalent about, like the many photos with long-ago exes, which for now I'll continue to hoard given that I'm still on good terms with them and I'm not going to fill up 12 terabytes any time soon.

There was also a lot of "data exhaust," as the security technologist Matt Mitchell calls it, a polite term for the record of my life rendered in Google searches, from a 2011 query for karaoke bars in Washington, D.C., to a more recent search for the closest Chuck E. Cheese. I will not keep those on my personal hard drive, and I may take the step of deleting them from Google's servers, which the company makes possible, because their embarrassment potential is higher than their archival value. Mr. Mitchell said super hoarders should pare down, not to make memories easier to find, but to eliminate data that could come back to bite them.

"You need to let go because you can't get hacked if there's nothing to hack," said Mr. Mitchell, the founder of CryptoHarlem, a cybersecurity education nonprofit. "It's only when you're storing too much that you run into the worst of these problems."

Inactive Accounts



Footage of the Boston Pops Orchestra being digitally preserved at George Blood, LP, a company that recovers data from obsolete media like VHS tapes, floppy disks and reel-to-reel tapes. Mark Makela for The New York Times

Right now, it's cheap to hoard all this data in the cloud.

"The cost of storage long term continues to fall," said George Blood, who runs a business outside Philadelphia digitizing information from obsolete media, creating 10 terabytes of data per day, on average. "They may charge you more for the cost of the electricity — spinning the disk your data is on — than the storage itself."

Big technology companies don't often prompt people to minimize their data footprints, until, that is, they near the end of their free storage space. That's when companies force them to decide whether to move to the paid plans. There are signs, though, that the companies don't want to hold onto our data forever: Most have policies allowing them to delete accounts that are inactive for a year or more.

Aware of the potential value of data left behind by those who euphemistically go “inactive,” Apple recently introduced a legacy contact feature, to designate a person who can access an Apple account after the owner’s death. Google has long had a similar tool, prosaically called inactive account manager. Facebook created legacy contacts in 2015 to look after accounts that have been memorialized.

And that really is the ultimate question around personal archives: What becomes of them after we die? By keeping so much, more than we want to sort through, which is almost certainly more than anyone else wants to sort through on our behalf, we may leave behind less than previous generations because our accounts will go inactive and be deleted. Our personal clouds may grow so vast that no one will ever go through them, and all the bits and bytes could end up just blowing away.

Audio produced by Kate Winslett.

Kashmir Hill is a tech reporter based in New York. She writes about the unexpected and sometimes ominous ways technology is changing our lives, particularly when it comes to our privacy. More about Kashmir Hill

A version of this article appears in print on , Section BU, Page 1 of the New York edition with the headline: Your Memories. Their Cloud.