

Only you can see this message



This story's distribution setting is on. [Learn more](#)

User Data Export Approaches of SaaS / Cloud Service / Social Media Providers — An Overview



Daniel Rosehill

May 1 · 8 min read

danielrosehilljm / CloudBackupApproaches

Unwatch 1

Unstar 1

Fork 0

<> Code

Issues 0

Pull requests 0

Actions

Projects 0

Wiki

Security 0

Insights

Settings

In this repository I will endeavor to maintain up to date documentation regarding how to backup your (user) data from various commonly used cloud services providers. [Edit](#)

[Manage topics](#)

23 commits

1 branch

0 packages

0 releases

1 contributor

Branch: master

New pull request

Create new file

Upload files

Find file

Clone or download

danielrosehilljm Added table to homepage Latest commit 59261b5 6 minutes ago

images	Green	15 minutes ago
Cpanels.md	Formatting	34 minutes ago
Facebook.md	Green	15 minutes ago
Github.md	Green	15 minutes ago

Continuing with the backup theme of this week, I decided to create a [GitHub](#) repository documenting the user backup export options of common SaaS and cloud providers.

These are all the services I have taken my own backups of this week — and they vary quite significantly in methodology and ease.

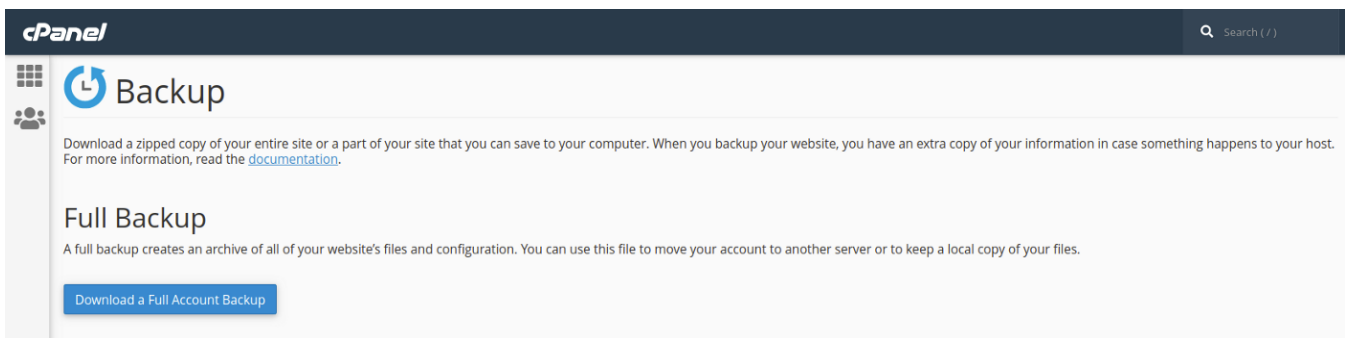
Here's an alphabetical run through:

• • •

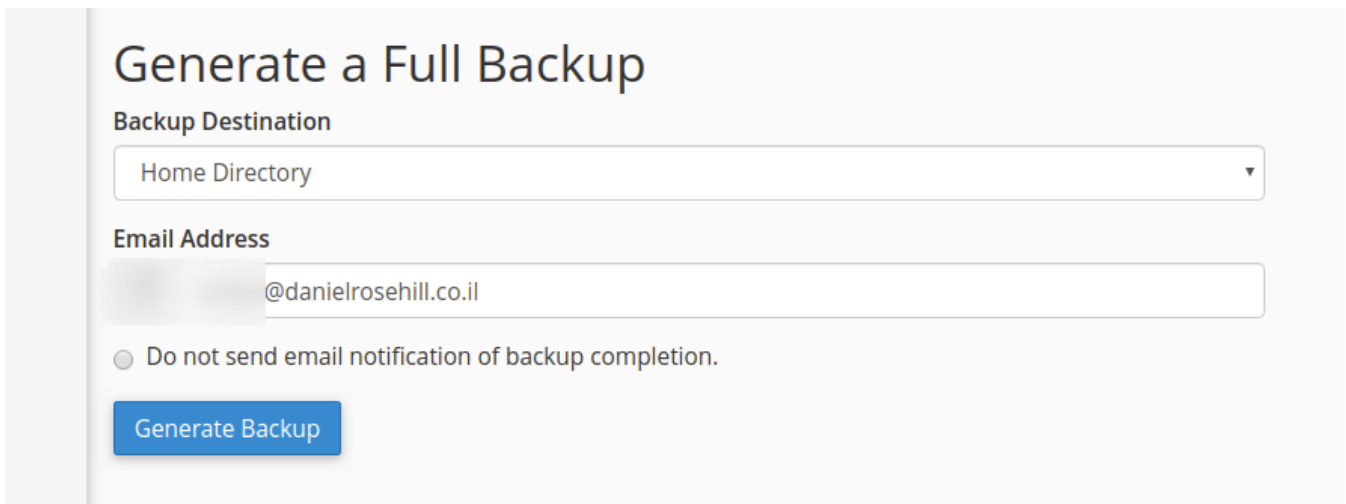
Cpanels

Backing up web hosting Cpanels is relatively straightforward.

Although there a variety of proprietary tools on the market — like JetBackup — Cpanel includes a default option for initiating full account backups including files and MySQL databases.



These can be taken cloud to cloud or delivered to email and moved elsewhere:



• • •

Facebook

I give Facebook top marks for the granularity and ease of its data export options.

Like [G Suite](#) Takeouts, Facebook users are able to control exactly what they want to include in their export archive, including:

- The date range
- The format (HTML/JSON)
- The media quality

Download Your Information

You can download a copy of your Facebook information at any time. You can download a complete copy, or you can select only the types of information and date ranges you want. You can choose to receive your information in an HTML format that is easy to view, or a JSON format, which could allow another service to more easily import it.

Downloading your information is a password-protected process that only you will have access to. Once your copy has been created, it will be available for download for a few days.

If you'd like to view your information without downloading it, you can [Access Your Information](#) at any time.

Request Copy

Available Copies

Date Range: All of my data ▼







Format: HTML ▼

Media Quality: High ▼

Create File

Users can also use simple checkboxes to indicate which services they want Facebook to compress into your archive on the backend:

Your Information ⓘ Deselect All

	Posts Posts you've shared on Facebook, posts that are hidden from your timeline and polls you have created	<input checked="" type="checkbox"/>
	Photos and Videos Photos and videos you've uploaded and shared	<input checked="" type="checkbox"/>
	Comments Comments you've posted on your own posts, on other people's posts or in groups you belong to	<input checked="" type="checkbox"/>
	Likes and Reactions Posts, comments and Pages you've liked or reacted to	<input checked="" type="checkbox"/>
	Friends The people you are connected to on Facebook	<input checked="" type="checkbox"/>
	Stories Photos and videos you've shared to your story	<input checked="" type="checkbox"/>

Users receive a Facebook notification and email when the compression is over and the archive is ready to be downloaded:





Hi Daniel,

The file you've requested containing your Facebook information is now ready for you to download. You can find your file in the [Available Files](#) tab on the [Download Your Information](#) page. For security reasons, the file will be available to download for only a few days after you receive this email. If your file is no longer available by the time you read this email, you can create a new file from the [New File](#) tab.

Because this file may contain private information, you should keep it secure and take precautions when storing or sending it, or when uploading it to another service.

- The Facebook Team

The archive is neatly organized — although trying to decipher statuses from the HTML clutter can be a little tricky!

Name	Size	Modified
about_you		22/04/20 11:16
ads_and_businesses		22/04/20 11:16
apps_and_websites		22/04/20 11:16
comments		22/04/20 11:16
events		22/04/20 11:16
facebook_gaming		22/04/20 11:16
files		22/04/20 11:16
following_and_followers		22/04/20 11:16
friends		22/04/20 11:16

• • •

GitHub

Exporting your repositories from [GitHub](#) is pretty straightforward.

Users can clone their repositories on local or cloud storage by using the Git CLI.

Alternatively, repositories can be manually downloaded as archives and uploaded to another cloud storage repository.

9 commits 1 branch 0 packages 0 releases 1 contributor

Branch: master New pull request

Create new file Upload files Find file Clone or download

Clone with SSH Use HTTPS

Use a password protected SSH key.

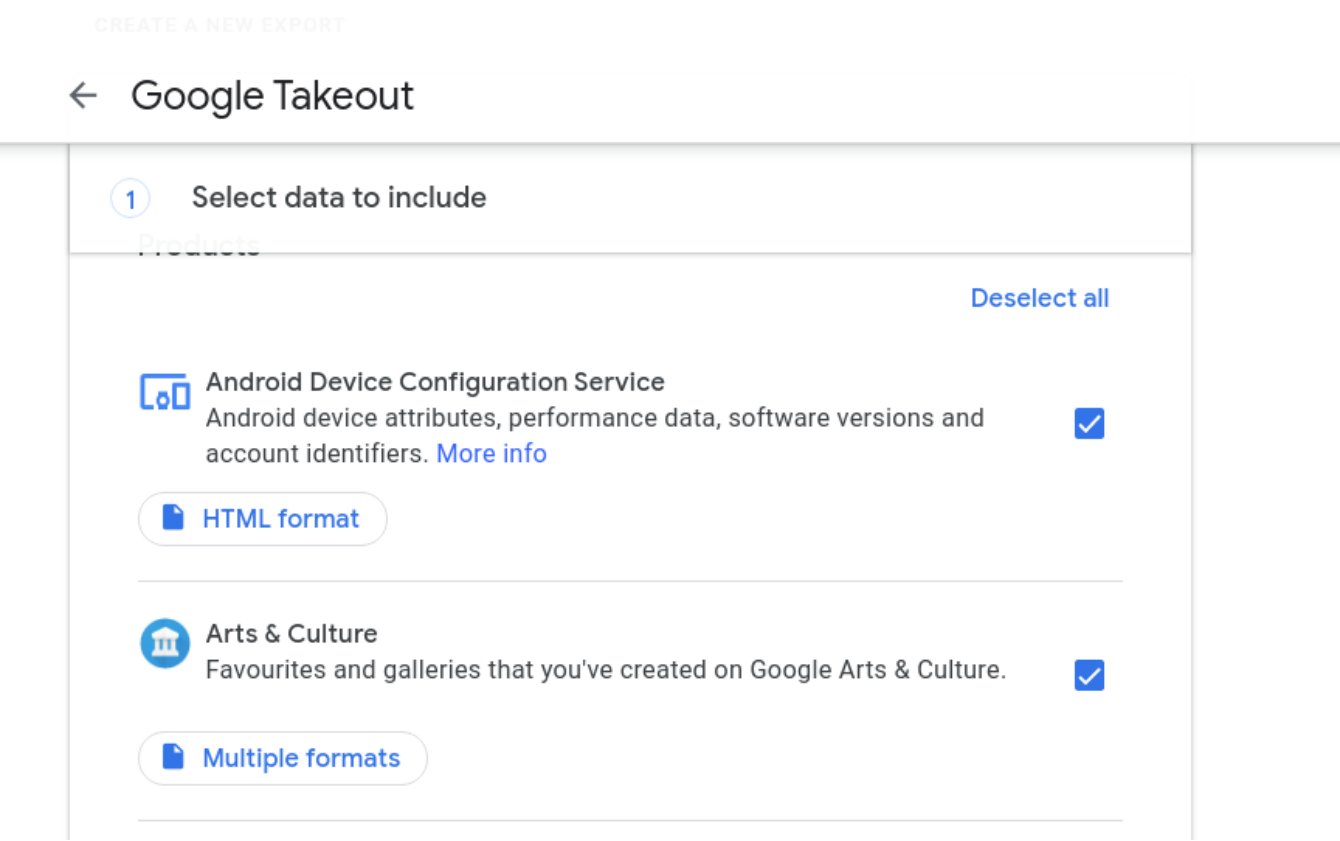
git@github.com:danielrosehilljlm/ToDoist

Download ZIP

. . .





G Suite

Thanks to its Google Takeout backup engine Google makes it very easy for users to export a copy of their data.



Google users can select precisely which services they wish to include in the export — although be warned, when YouTube videos are added to the mix these files can easily end up constituting tens of gigabytes.

Users can retrieve their generated Takeouts for some time after they have been created:

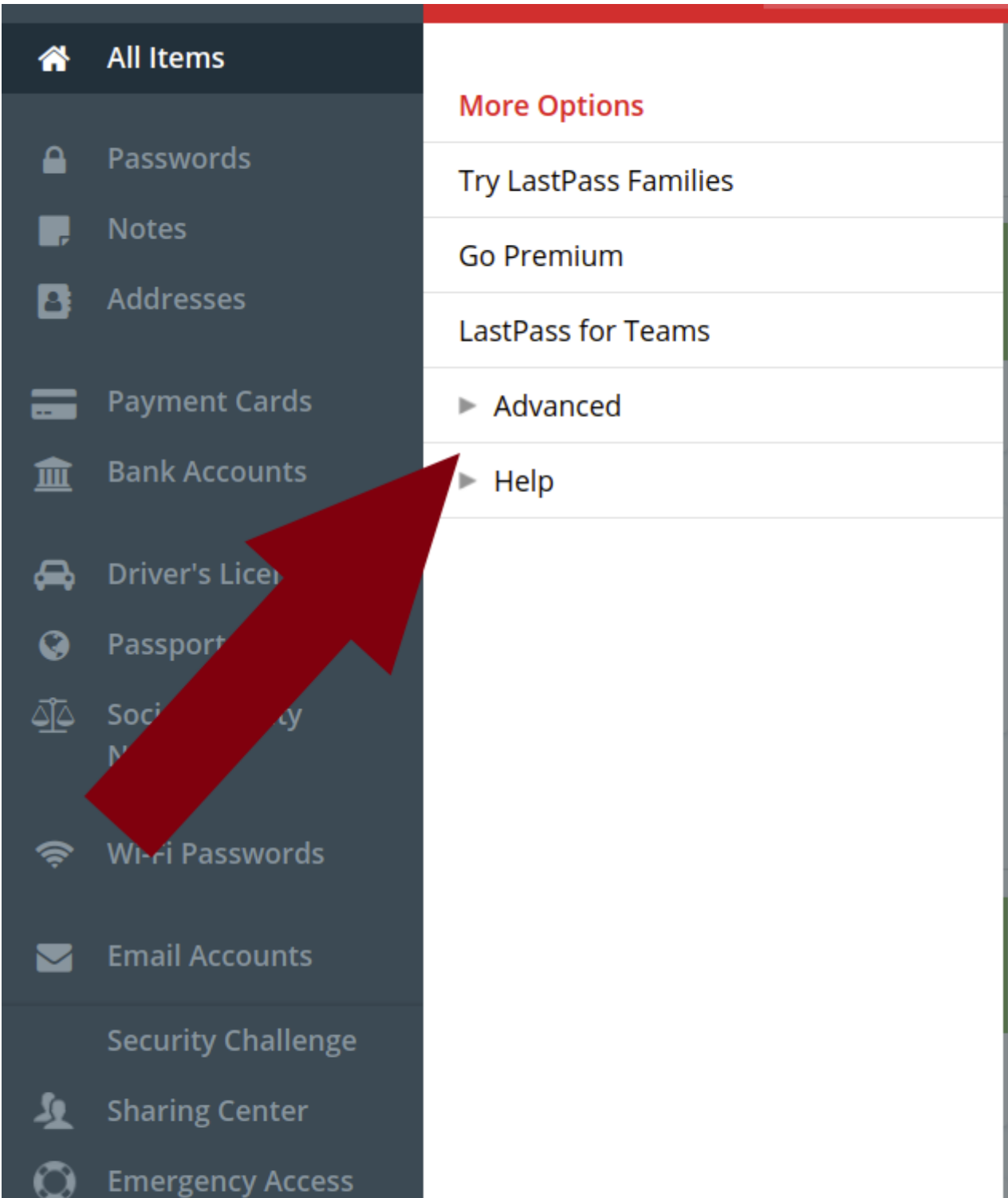
Export	Created on	Available until	Details
YouTube and YouTube Music 29.35 GB		4 May 2020	<input checked="" type="checkbox"/> Download 
50 products 55.44 GB		Expired	

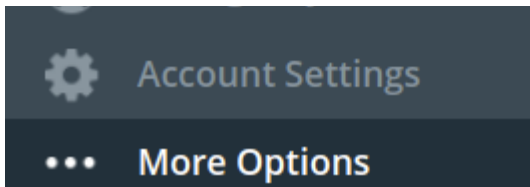
Create new export

...

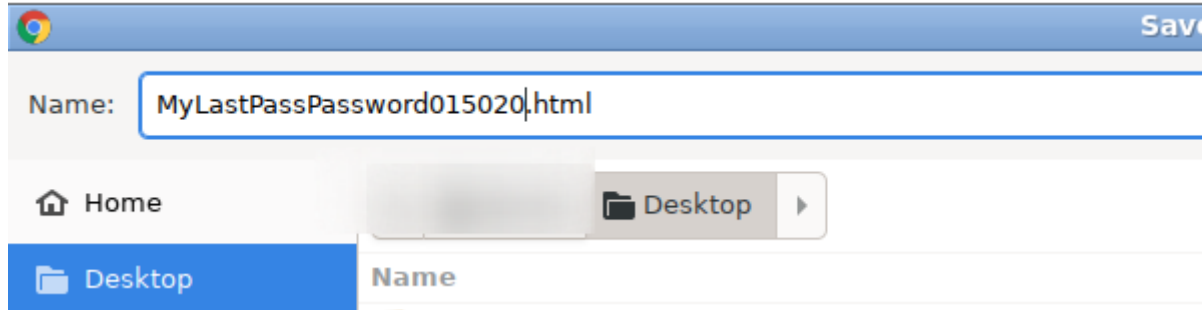
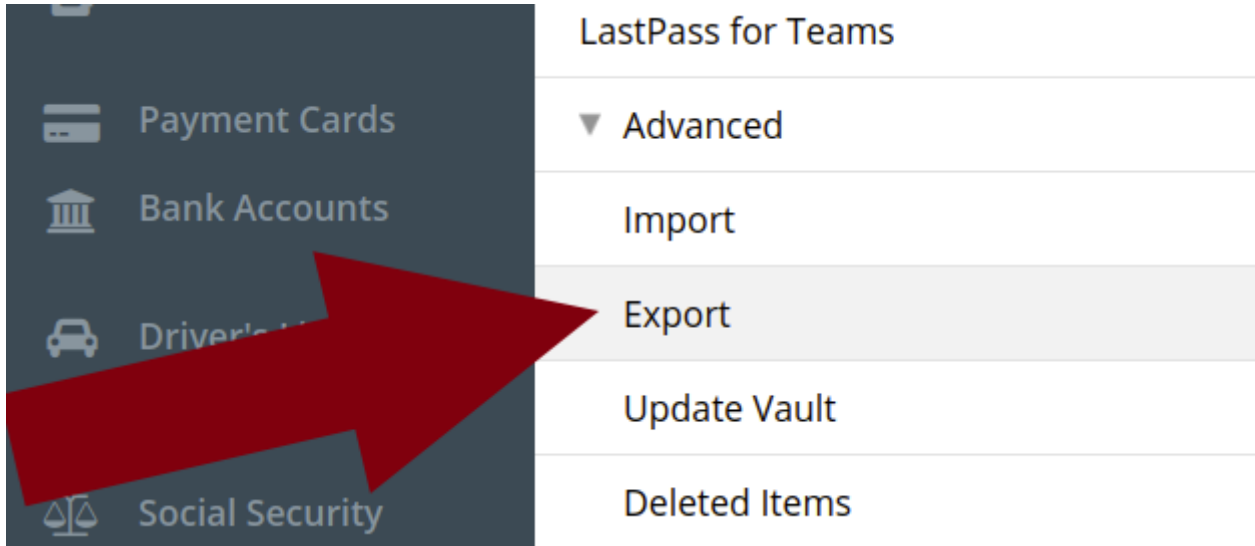
LastPass

Users can backup LastPass by navigating to the ‘Export’ functionality in advanced options.





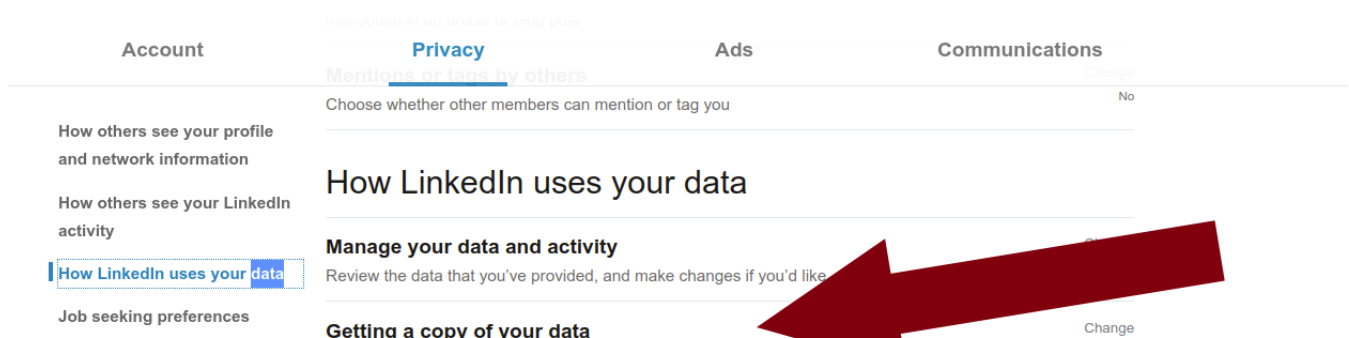
Then by clicking on 'Export'. And finally by saving the .php output and saving it to another cloud storage repository.



...

LinkedIn

LinkedIn has a decent user data export functionality nested under the privacy tab:



Blocking and hiding

See your options for accessing a copy of your account data, connections, and more

Manage who can discover your profile from your email address

Change Everyone

Choose who can discover your profile if they are not connected to you but have your email address

Manage who can discover your profile from your phone number

Change Nobody

Choose who can discover your profile if they have your phone number

After clicking ‘getting a copy of your data’ users can choose exactly what they would like to bundle in the archive:

Getting a copy of your data

Close

See your options for accessing a copy of your account data, connections, and more

Your LinkedIn data belongs to you, and you can download an archive any time or [view the rich media](#) you have uploaded.

☐ Download larger data archive, including connections, contacts, account history, and information we infer about you based on your profile and activity. [Learn more](#)

☐ Want something in particular? Select the data files you're most interested in.

☐ Articles

☐ Imported Contacts

☐ Invitations

☐ Recommendations

☐ Connections

☐ Messages


☐ Profile

☐ Registration


Request archive

Don't see what you want? Visit our [Help Center](#).

The final archive will be delivered via email:



Daniel Rosehill



Hi Daniel,

We've finished putting together the second part of your data **archive**. You can [download the second part of your data archive using this link](#). The **archive** is good for the next 72 hours. However, you can always request a new **archive** every 2 hours.

Because it may contain personal information, be sure to keep the link private and only download the **archive** to your own computer.

https://medium.com/daniels-tech-world/user-data-export-approaches-of-saas-cloud-service-social-media-providers-an-overvie... 8/20

link private and only download the **archive** to your own computer.
[Visit our Help Center for more information.](#)

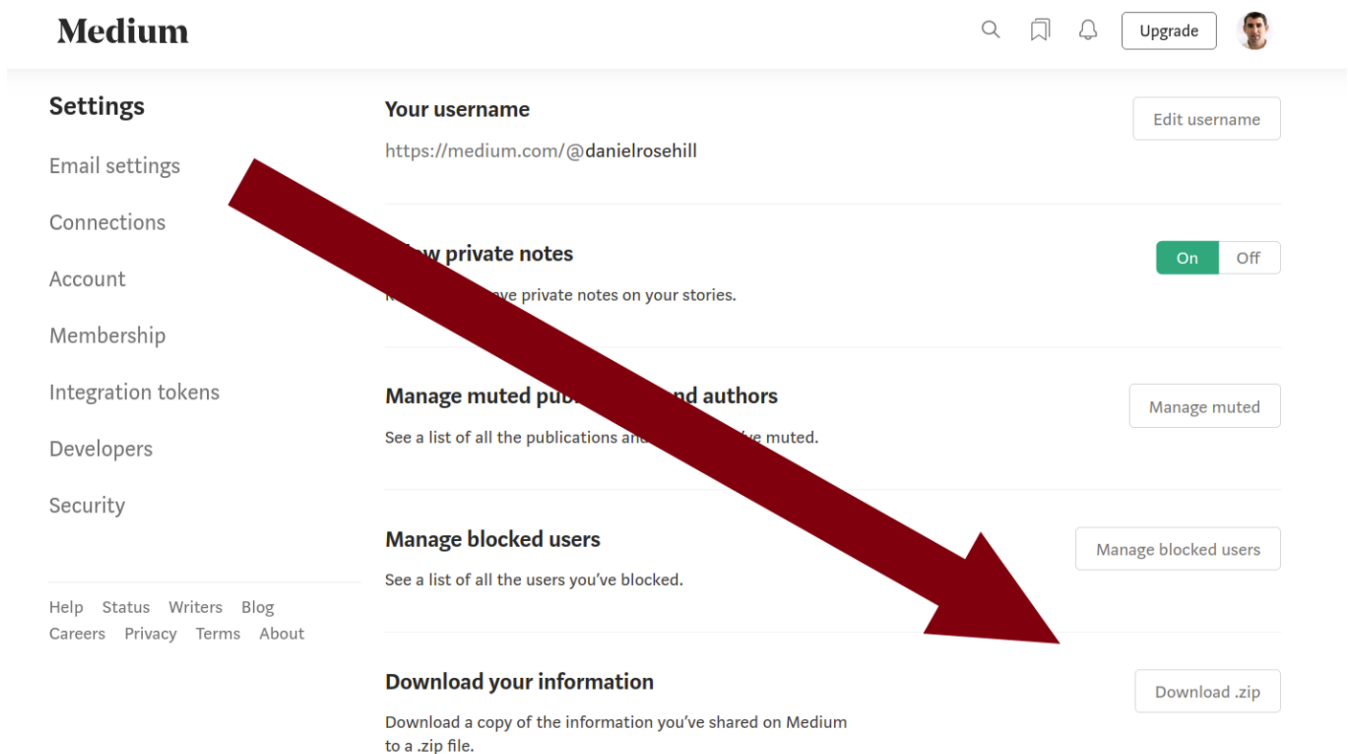
Thanks,
The **LinkedIn** Team

. . .

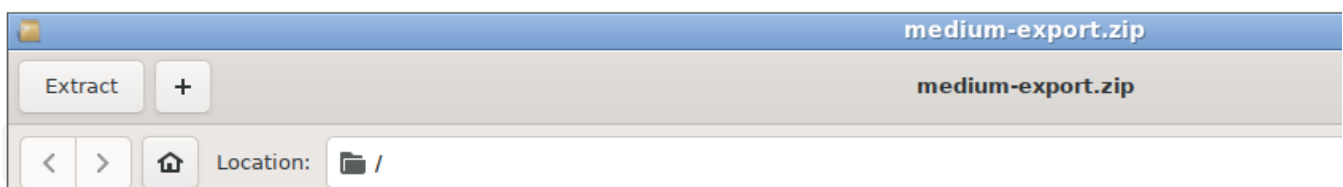
Medium

Dear, **Medium Staff**,

I love this platform and your clean-cut UI. I liked how easy it was to find the ‘Download your Information’ button:

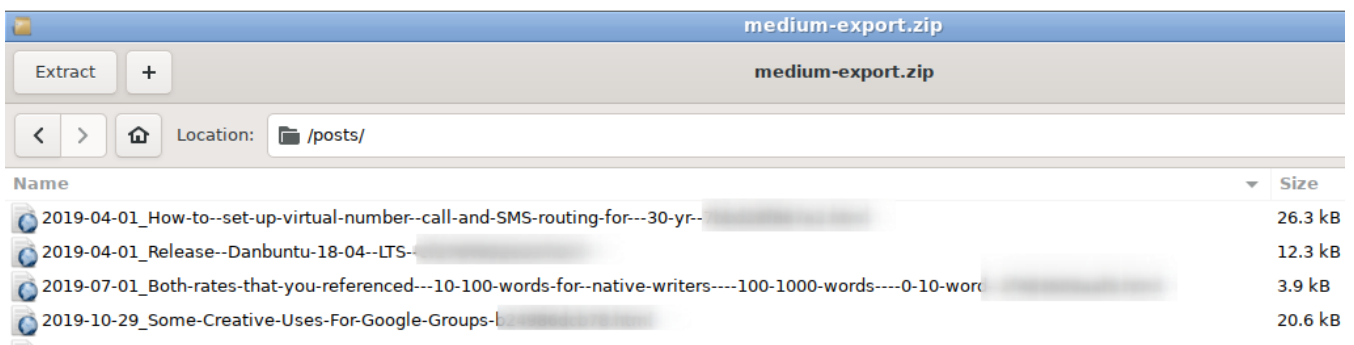


However, in the spirit of this article, I had to dig a little deeper and take a look at what you presented in the Zip archive:



Name	Size	Type	Modified
blocks	1.6 kB	Folder	
bookmarks	1.5 kB	Folder	
claps	7.1 kB	Folder	
highlights	4.7 kB	Folder	
interests	15.6 kB	Folder	
ips	1.7 kB	Folder	
posts	1.6 MB	Folder	
profile	5.9 kB	Folder	
pubs-following	15.3 kB	Folder	
sessions	8.0 kB	Folder	
topics-following	1.7 kB	Folder	
users-following	45.9 kB	Folder	

The articles are .HTML files. That's fine. But where are the images, I was wondered?



Answer: trapped in the [Medium](#) CDN!

```

t amount of legitimate newsletters as spam.</p><figure name= "4893" id= "4893" clas
div class="aspectRatioPlaceholder-fill" style="padding-bottom: 47.199999999999996
e" src="https://cdn-images-1.medium.com/max/800/0*XY1ngmJYFE1e2qnM.png"></div></f
tion setting edit:</p><figure name="4893" id="4893" class="graf graf--figure graf
ceholder-fill" style="padding-bottom: 51.7%;"></div><img class="graf-image" data-

```

[Medium](#), your CDN and progressive image loading technology is amazing.

But would you mind giving us access to our images, please?

...

Quora

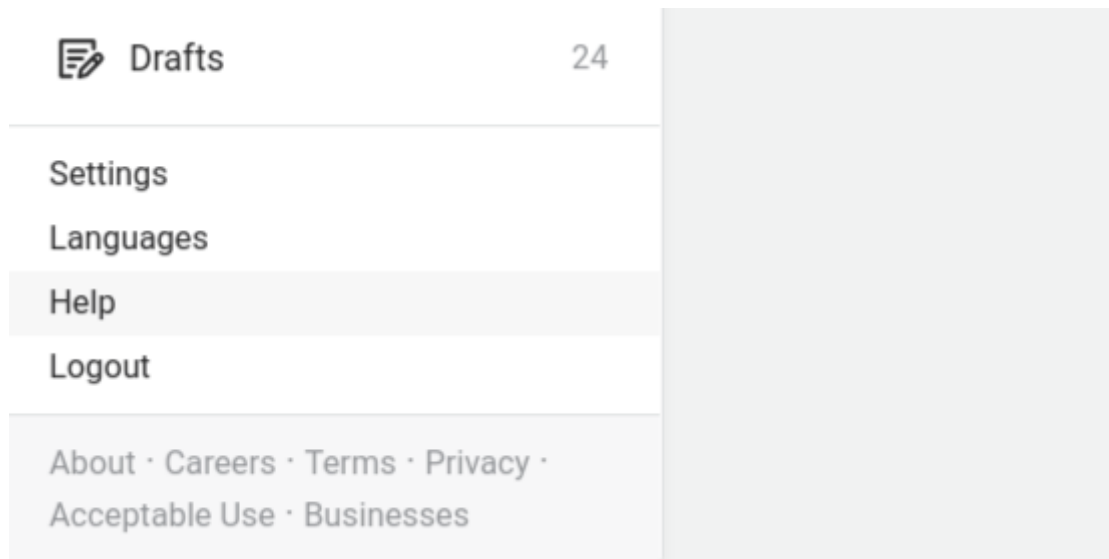
Better late than never, right?

Quora's lack of a built-in backup functionality has been something of a sticking point among the user community for many years.

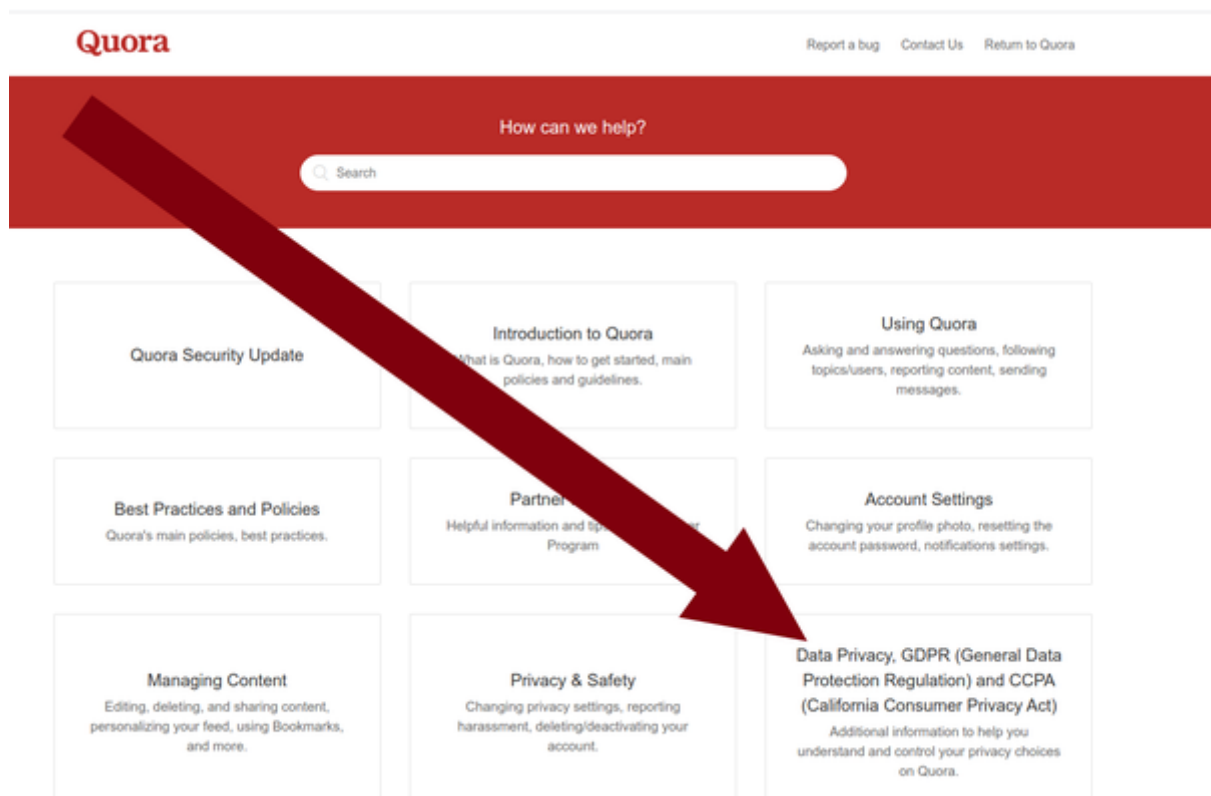
Thankfully, Quora seems to have finally seen the light.

Although it doesn't yet have a native backup export functionality, you can ask its support people to generate a download archive for you.

Proceed as follows:



Follow the link for data privacy:



You should eventually find this help center resource:

Can I get a copy of my data?

Yes. Quora will send an archive of your content and personal data to your account's primary email address on request...

help.quora.com

Follow the instructions and within 72 hours, as stated, you will get your data by email.

My archive came in a little over 24 hours.

Quora will give you two archives: data and content. Content contains your answers and images:

The HTML file which the process outputs looks like this in a web browser:

Answers

Answer

Question: How do I market myself as a Ghostwriter?
Content:

The only complication involved in marketing oneself as a ghostwriter (versus as a bylined author) is that quite often the majority (or entirety!) of one's portfolio is officially attributed to another author, i.e. the client.

There are various workarounds to this common problem, but at the very least the protocol is that clients' permission should always be sought before publicly claiming the work as that of the ghostwriter. This is vital because publicly stating that a written work is not the client's obviously negates the reason for which a ghostwriter is typically hired.

Other solutions include sharing samples privately on request (always check any contracts and NDAs signed before doing so) or forcing some kind of restriction on one's portfolio, such as password protection or token-based, time-controlled access. The objective being that it shouldn't be public-facing.

Other than that one does the typical things a writer does to get hired and win more clients. But in ghostwriting, more so than general content marketing, personal networks and referrals are key.

Creation time: Oct 15, 2019 09:11 PM PDT

Content language: English

Of course, if keep the images in the same folder you will get the images as well where you have them in your answers:

Answer

Question: How do I find out what LinkedIn skills are requested for jobs?

Content:

Currently, you will have to do as [Teddy Burriss](#) says and read the job specifications carefully and ask the recruiter if you are unsure.

It would be a cool addition to the current set of jobs filters, but specific skills are not currently supported on either side of the marketplace as far as I know (that is to say, those listing jobs cannot cite specific skills requirements and those applying for them cannot filter jobs based on what skills the recruiter has required)

All jobs filters

Clear

Cancel

Apply

Date Posted <input type="checkbox"/> Past 24 hours (2,000) <input type="checkbox"/> Past Week (8,734) <input type="checkbox"/> Past Month (22,495) <input checked="" type="checkbox"/> Any Time (22,495)	Content Location <input type="checkbox"/> Anywhere (22,495) <input type="checkbox"/> Under 10 Milestones (1,574) <input type="checkbox"/> In Your Network (875)	Job Type <input type="checkbox"/> Full-time (2,446) <input type="checkbox"/> Contract (626) <input type="checkbox"/> Temporary (595) <input type="checkbox"/> Internship (194) <input type="checkbox"/> Part-time (855) <input type="checkbox"/> Volunteer (3)	Company <input type="checkbox"/> Add a company <input type="checkbox"/> Specific Connections (86) <input type="checkbox"/> Language Networks (245) <input type="checkbox"/> Recruitment Sites (452) <input type="checkbox"/> No location (485) <input type="checkbox"/> Workplaces (1,334)
Industry <input type="checkbox"/> Add an industry <input type="checkbox"/> Computer Software (7,768) <input type="checkbox"/> Financial Services (887) <input type="checkbox"/> Marketing and Advertising (1,092) <input type="checkbox"/> Information (1,090)	Job Function <input type="checkbox"/> Add a job function <input type="checkbox"/> Information Technology (1,746) <input type="checkbox"/> Customer Service (822) <input type="checkbox"/> Sales (754) <input type="checkbox"/> Business (1,090)	Experience Level <input type="checkbox"/> Intermediate (2,624) <input type="checkbox"/> Entry-level (2,432) <input type="checkbox"/> Associate (1,436) <input type="checkbox"/> Mid-Senior level (2,022) <input type="checkbox"/> Director (264) <input type="checkbox"/> Executive (1,316)	Location <input type="checkbox"/> Portugal (2,624) <input type="checkbox"/> Belgium (486) <input type="checkbox"/> United Kingdom (817) <input type="checkbox"/> Netherlands (1,392) <input type="checkbox"/> Spain (2,088)

Creation time: Dec 02, 2019 05:29 AM PST

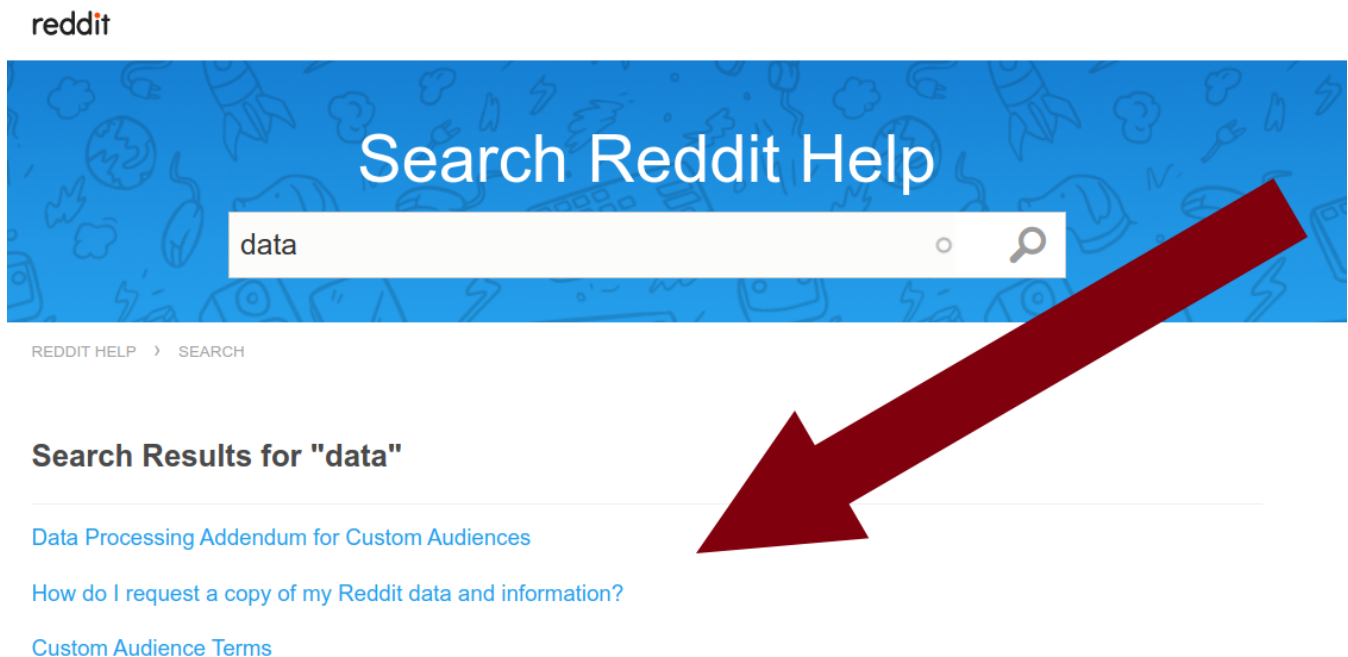
• • •

Reddit

Reddit has also never developed a self service backup functionality.

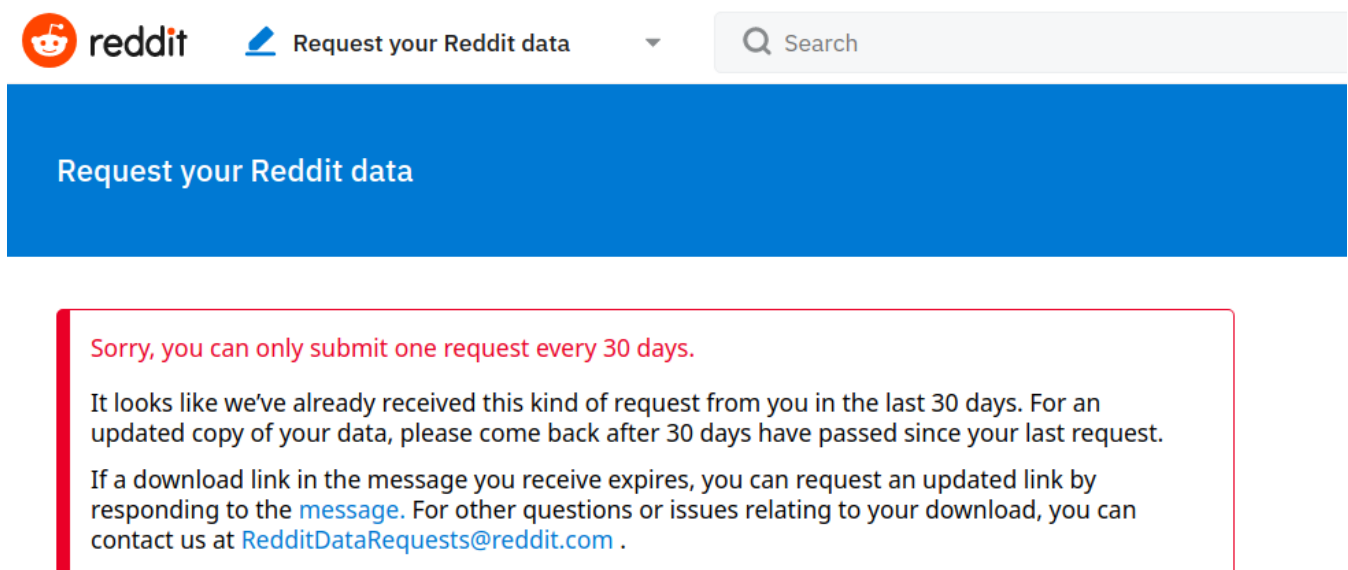
Like [Quora](#), you need to ask their support people in order to get an archive of your data from their storage.

Do a little bit of digging in the help center:



And ask them for an archive.

You can choose to download all your Reddit activity or just your activity over a specific date range:



Complete the form below to request a copy of data Reddit has about your Reddit account. Once submitted, please allow up to 30 days for us to process your request.

- ☒ I want data from my full time at Reddit
- ☐ I want data from a specific date range

A short while after you ask for your data, /u/RedditDataRequests will get back to you with your archive:

[/u/RedditDataRequests](#) **Download a copy of your data:**
[expand all](#) [collapse all](#)

[–] from [/u/RedditDataRequests](#) [A] sent 7 hours ago

Hello,

Your request has been processed, and you can download a copy of your data by clicking the link below.
















[Download here](#)

If you encounter an error trying to download your data, and this message is older than seven days, you can request an updated link by responding directly to this message. For other questions or issues relating to your download, you can also contact us at RedditDataRequests@reddit.com.

[Permalink](#) [Delete](#) [Mark Unread](#) [Reply](#)

All your account activity should be in the archive:

Name	Size	Type	Modified
approved_submitter_subreddits.csv	206 bytes	CSV docum...	30 April 2020, 16:12
chat_history.csv	64.8 kB	CSV docum...	30 April 2020, 16:12
checkfile.csv	2.4 kB	CSV docum...	30 April 2020, 16:12
comment_headers.csv	477.1 kB	CSV docum...	30 April 2020, 16:12
comments.csv	1.2 MB	CSV docum...	30 April 2020, 16:12
comment_votes.csv	281.2 kB	CSV docum...	30 April 2020, 16:12
drafts.csv	142 bytes	CSV docum...	30 April 2020, 16:12
friends.csv	15 bytes	CSV docum...	30 April 2020, 16:12
gilded_comments.csv	32 bytes	CSV docum...	30 April 2020, 16:12
gilded_posts.csv	32 bytes	CSV docum...	30 April 2020, 16:12

 grinded_posts.csv	32 bytes	CSV docum...	30 April 2020, 16:12
 hidden_posts.csv	549 bytes	CSV docum...	30 April 2020, 16:12
 ip_logs.csv	47.9 kB	CSV docum...	30 April 2020, 16:12
 message_headers.csv	19.5 kB	CSV docum...	30 April 2020, 16:12
 messages.csv	205.0 kB	CSV docum...	30 April 2020, 16:12
 moderated_subreddits.csv	239 bytes	CSV docum...	30 April 2020, 16:12
 multireddits.csv	92 bytes	CSV docum...	30 April 2020, 16:12
 poll_votes.csv	144 bytes	CSV docum...	30 April 2020, 16:12
 post_headers.csv	207.9 kB	CSV docum...	30 April 2020, 16:12
 posts.csv	704.8 kB	CSV docum...	30 April 2020, 16:12
 post_votes.csv	107.0 kB	CSV docum...	30 April 2020, 16:12
 reddit_gold_information.csv	48 bytes	CSV docum...	30 April 2020, 16:12
 saved_comments.csv	127 bytes	CSV docum...	30 April 2020, 16:12
 saved_posts.csv	582 bytes	CSV docum...	30 April 2020, 16:12
 statistics.csv	201 bytes	CSV docum...	30 April 2020, 16:12

This includes:

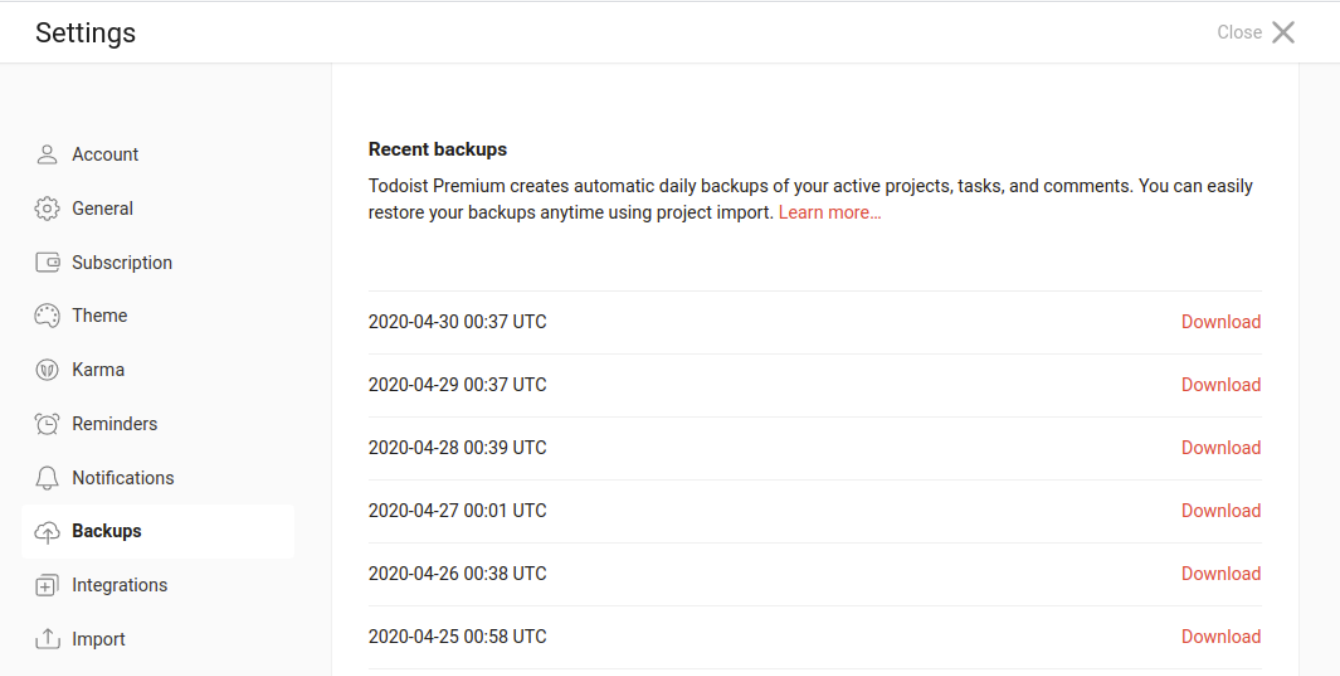
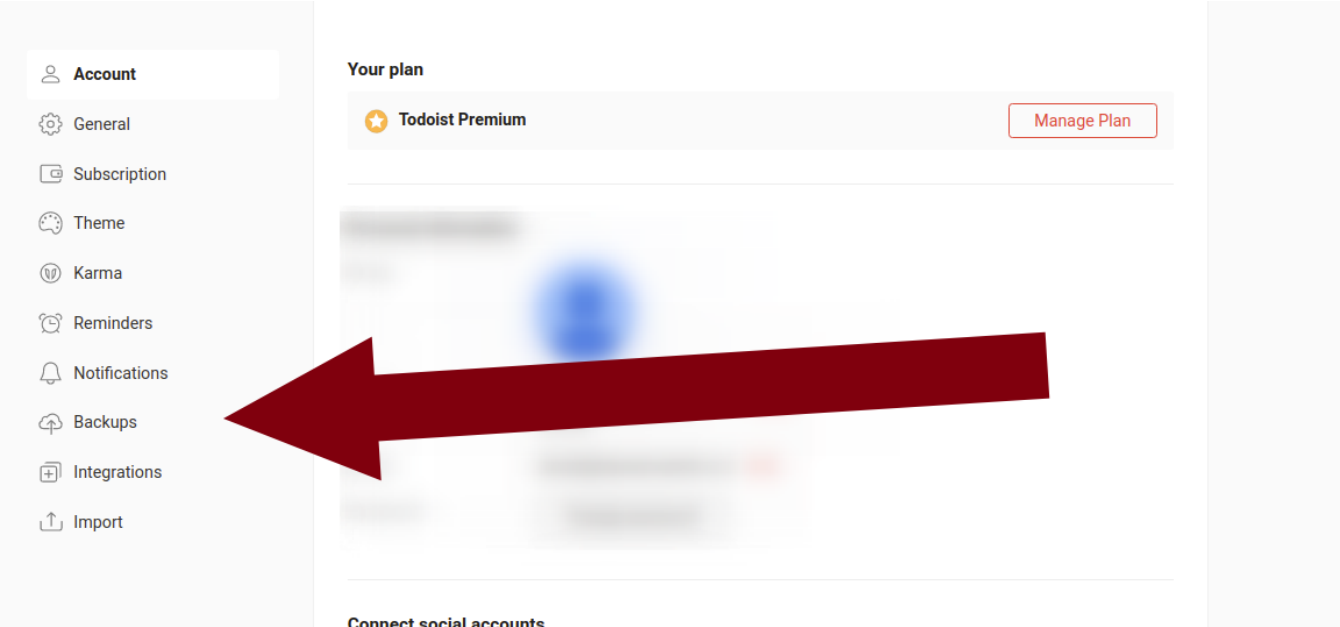
- Your posts
- Your comments
- Your IP log
- Your chat history
- Your hidden posts
- The subreddits

• • •

Todoist

I give my favorite task manager top marks for how easy it makes it to download backups of your own data.

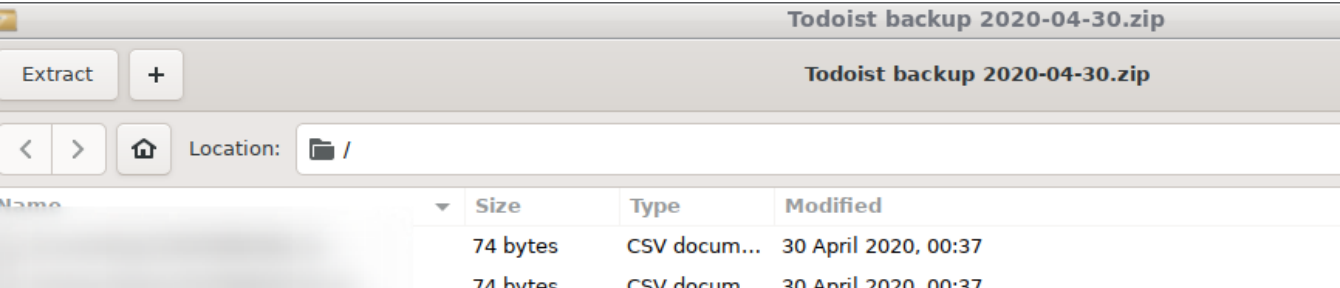
True, todo lists are basically just text files. But Todoist creates a bunch of snapshots and allows you to download whatever date points you like.



The data is presented as a .CSV file for each project.

However, I couldn't find task attachments — like images or audio clips — anywhere in the archive.

This could be improved upon.





74 bytes	CSV docum...	30 April 2020, 00:37
225 bytes	CSV docum...	30 April 2020, 00:37
6.5 kB	CSV docum...	30 April 2020, 00:37
74 bytes	CSV docum...	30 April 2020, 00:37

• • •

Twitter

Finally, the Twitter machine.

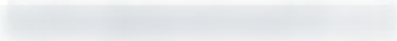
Twitter ranks towards the top of the table.

It doesn't take much digging to locate the data export functionality.

Data and permissions	
Display language English	>
Country Israel	>
Your Twitter data	>
Apps and sessions	>

Once you've done that it's a couple more clicks to download the archive. No data requests to support teams necessary.

Download an archive of your data

You can request a ZIP file with the information that we believe is most relevant and useful to you. You'll get an email sent to  when it's ready to be downloaded. Your download will include a file called "Your archive" that will allow you to see your data in a desktop web browser. [Learn more](#)

Twitter

Download archive

Periscope

Request archive

Deciphering your own tweets isn't too hard:

```
    },
    "display_text_range" : [ "0", "168" ],
    "favorite_count" : "0",
    "truncated" : false,
    "retweet_count" : "0",
    "possibly_sensitive" : false,
    "created_at" : "Wed Apr 22 17:39:49 +0000 2020",
    "favorited" : false,
    "full_text" : "Do you ever open your @LinkedIn feed and wonder when modesty disappeared as a human value? Good idea for a blog post: Is Modesty Dead? #linkedin https://t.co/X4fs8nnxUI",
    "lang" : "en",
    "extended_entities" : {
      "media" : [ {
        "expanded_url" : "https://twitter.com/danielsrosehill/status/1253015666447265793/photo/1"
```

And Twitter actually does a really good job at organizing these within the archive.

Your tweets are in /tweet.js and any media you include with your tweets — namely images — is at /data/tweet_media:

tweet_media		
Desktop/twitterarchive/data/tweet_media		
Name	Size	Modified
1235325370104270848-ESTBnGLWoAYeICc.jpg	73.0 KiB	22/04/20 19:15
1238177167693021190-ES7KjbX0AlaAt7.jpg	70.5 KiB	22/04/20 19:15
1239138082223513601-ETJNSccWoAAIXcc.jpg	161.8 KiB	22/04/20 19:15
1239212691706781698-mER9TRtDzf-2lthf.mp4	883.3 KiB	22/04/20 19:15
1239443914823733248-ETNjScAXkAEGpK7.png	28.1 KiB	22/04/20 19:15
1239450568722178049-ES5o9Q5XQAc_DAL.jpg	123.7 KiB	22/04/20 19:15
1240289630769070081-ETZksjjWkAEaEq.jpg	103.0 KiB	22/04/20 19:15

• • •

Concluding Thoughts

SaaS services and cloud providers vary enormously in how well they allow users to download a copy of their own data.

I would suggest various criteria to judge a provider by:

- Is the data export self service or do you need to contact support in order to get your data out? Self service is obviously optimal
- How is the data export packaged? Take a look inside an archive to find out
- Are your images included or are these trapped in the provider's CDN or simply not given out at all?

Hopefully, as awareness about the importance of data protection, governance, and portability continues to grow we can look forward to an era in which every internet user has on-demand access to a well-structured and complete archive of all the data they have created in every cloud or cloud service that they use.

. . .

Github Repository

<div>danielrosehilljlm/CloudBackupApproaches</div> <div>In this repository I will endeavor to maintain up to date documentation regarding how to backup your (user) data from...</div> <div>github.com</div>	
---	--

More About Backups

<div>How to Extract Your Data From Quora and Reddit</div> <div>Thank you, GDPR!</div> <div>medium.com</div>	
<div>3-2-1: A Common-Sense Approach For Backing Up Ubuntu - And Keeping It In Good Order</div> <div>If using Ubuntu, backups are a topic that you should give at least occasional thought to. I use Timeshift as my primary...</div> <div>linuxhint.com</div>	
<div>How to: Use Multcloud For Cross-Cloud Backups</div> <div>I've written a little bit about the importance of taking (3-2-1 compliant!) backups — whether you're using to back up...</div> <div>medium.com</div>	

<div><div><h3>How to: Backup Your Web Hosting (According to the 3–2–1 Principle)</h3><p>In my writings about backups I have emphasized the importance of adhering to the 3–2–1 principle across all your data...</p><p>medium.com</p></div></div>	
---	--

Technology

SaaS

Data

Backup

About

Help

Legal

Get the Medium app

