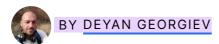




TOP ^

## **Snscrape: What it is, How** to Use it, & More!



LAST UPDATED ON FEBRUARY 25, 2025

#data scraping #proxy #social media scraper

If you've played around with social media data, you might have come across a tool called "snscrape."

As someone who frequently uses various software tools, I've found snscrape to be a great tool for some tasks, especially when combined with a proxy.

So let's break down what snscrape is, how it works, and why it can be a game-changer for your data collection workflows.

### **Table of Contents**

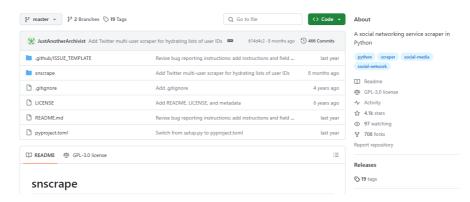
- 1. What is Snscrape?
- 2. How Does Snscrape Work?
- 3. How to Use Snscrape
- 4. What Data Can You Scrape from Twitter With Snscrape?
- 5. Snscrape Benefits





- 9. Alternatives to Snscrape
- 10. Final Words

## 1. What is Snscrape?



TOP ^

Screenshot via GitHub

Snscrape is a Python library that allows you to scrape data from social media platforms like X (Twitter), Facebook, and Instagram.

Unlike many other scraping tools, snscrape doesn't require an API key, which makes it usable by a wider range of users, including those who might not have the technical know-how to use APIs.

## 2. How Does Snscrape Work?

Using snscrape is relatively straightforward, especially if you have a basic understanding of Python.

The tool works by scraping social media websites for publicly available data. For example, you can use it to collect tweets containing specific keywords, hashtags, or from certain users.

Here's a simple example:





## 3. How to Use Snscrape

To get started, you'll need a basic setup that includes Python installed on your computer. Then, you can install snscrape using pip, Python's package installer. Once installed, you can begin writing scripts to scrape the data you need.



Here's a simple guide to get you started with this powerful tool.

#### a. Installation

First, you need to install snscrape. Open your command prompt or terminal and type the following command:

bashCopy code

```
pip install snscrape
```

This command uses Python's package manager, pip, to download and install the scraper.

#### b. Write a Basic Script

Once you install it, you can start writing a Python script to scrape data. Let's say you want to collect tweets with a specific hashtag. Here's a basic example:

pythonCopy code

```
import snscrape.modules.twitter as sntwitter #
Define the number of tweets to scrape max_tweets =
100 # Using TwitterSearchScraper to scrape data and
append tweets to list for i, tweet in
enumerate(sntwitter.TwitterSearchScraper('#technolo
gy').get_items()): if i > max_tweets: break
print(tweet.content)
```





printed out in your command prompt or terminal.

#### d. How to Customize Your Query

You can modify your search query easily. For example, if you want to scrape tweets from a specific user, you can change the query in the TwitterSearchScraper method:



pythonCopy code

```
sntwitter.TwitterSearchScraper('from:username')
```

Replace username with the Twitter handle of the user whose tweets you want to scrape.

#### e. Handling the Data

The data you scrape can be stored in various formats. For instance, you might want to save the tweets in a CSV file for easier analysis. You can modify your script to write the scraped data into a file:

pythonCopy code

```
import csv # ... [previous code] # Open/create a
file to append data to csvFile =
open('scraped_tweets.csv', 'a', newline='',
encoding='utf8') # Use csv writer csvWriter =
csv.writer(csvFile) csvWriter.writerow(['id',
'date', 'tweet']) for i, tweet in
enumerate(sntwitter.TwitterSearchScraper('#technolo
gy').get_items()): if i > max_tweets: break
csvWriter.writerow([tweet.id, tweet.date,
tweet.content]) csvFile.close()
```

This script will save the tweet ID, date, and content into a CSV file named scraped\_tweets.csv.





Enjoy fast, secure, and anonymous scraping, supported by exceptional performance and reliability.

Optimize your scraping with RapidSeedbox today!

TOP 🔨

# 4. What Data Can You Scrape from Twitter With Snscrape?

Snscrape can help you extract a wide range of data from X (Twitter). Here's a breakdown of the various data points you can scrape with snscrape:

#### a. Tweets

The primary use of snscrape is to collect tweets. This includes:

- **Tweet content**: The actual text of the tweet.
- Tweet ID: A unique identifier for each tweet.
- Date and time: When the tweet was posted.
- **URLs**: Any links included in the tweet.

#### b. User Information

Snscrape allows you to gather information about Twitter users, such as:

- Username: The Twitter handle of the user.
- **User ID**: A unique identifier for each user.





#### c. Engagement Metrics

While snscrape doesn't directly fetch engagement metrics like likes or retweets, you can still collect:

- Retweet count: The number of times a tweet has been retweeted (for some tweets).
- Reply count: The number of replies to a tweet (for some tweets).



#### d. Hashtags and Mentions

Snscrape can extract specific elements within tweets, including:

- Hashtags: Any hashtags used in the tweet.
- Mentions: Usernames of other X (Twitter) accounts mentioned in the tweet.

#### e. Media Content

If a tweet contains media, snscrape can help you identify:

 Media URLs: Links to images or videos attached to the tweet.

#### f. Advanced Search Queries

Snscrape is capable of handling advanced search queries, allowing you to scrape tweets based on:

- Keywords: Tweets containing specific words or phrases.
- Date Ranges: Tweets posted within a specified time frame.
- RapidSeedbox uses cookies to make the site simpler.

  Learn more ×





anouse, exampling

 Conversational tweets: Replies and quoted tweets, allowing you to track conversations.

## 5. Snscrape Benefits

Snscrape comes with several pros that are worth noting:



- No API Key required: This is a significant advantage.
   Most social media platforms require an API key for
   data scraping, which can be a problem for many
   users. This is one of the reasons many people prefer
   Snscrape.
- 2. **Flexibility**: You can tailor your search queries to be as broad or as specific as you need. This flexibility is crucial for research that requires nuanced data collection.
- Ease of use: For those familiar with Python, snscrape is user-friendly. Its straightforward commands and structure make it easy to integrate into your data collection workflow.

## 6. Snscrape Use Cases

- Market research: By scraping social media for mentions of specific products or brands, companies can gain valuable data for their customers and trends.
- Academic research: Researchers can use snscrape to collect data on public discourse around various topics.
- Personal projects: Even for personal projects, like tracking the popularity of a hobby or interest, snscrape can be a handy little tool.



 $\equiv$ 

Here's a step-by-step guide on how to integrate a proxy with snscrape:

#### a. Choose a Proxy Service

Select a <u>reliable proxy service</u>. There are various <u>types</u>
<u>of proxies</u> available, including free and paid services.

The latter generally offer better reliability and speed.



#### b. Get Your Proxy Information

Once you have chosen a proxy service, gather the necessary information: the proxy server address, port number, and, if applicable, the username and password.

#### c. Configure Your Python Script

Next, you'll need to modify your Python script to route snscrape requests through the proxy.

Here's an example of how to do this:

pythonCopy code

```
import snscrape.modules.twitter as sntwitter import
requests # Proxy configuration proxies = { 'http':
   'http://username:password@proxyserver:port',
   'https':
   'https://username:password@proxyserver:port', } #
Create a session and configure it to use the proxy
session = requests.Session()
session.proxies.update(proxies) # Use snscrape with
the session for tweet in
sntwitter.TwitterSearchScraper('keyword',
session=session).get_items(): print(tweet.content)
```

Replace username, password, proxyserver, and port with





- Bypass rate limits: Proxies can help avoid hitting X (Twitter)'s rate limits by distributing requests across different IP addresses.
- Avoid IP bans: Regular scraping from the same IP can lead to bans. Proxies mitigate this risk by <u>rotating</u> your IP address.



- 3. Access geographically restricted content: Proxies can provide IP addresses from different locations, allowing access to region-specific content.
- 4. **Anonymity and privacy**: Using a proxy hides your real IP address, enhancing your privacy and reducing the risk of being tracked.
- 5. **Improved performance**: Proxies can lead to faster data retrieval and reduce server overload risks by distributing the load.

## 9. Alternatives to Snscrape

While snscrape is a robust tool for social media data scraping, there are situations where you might need an alternative.

Whether it's due to different feature requirements, platform support, or ease of use, checking other tools might be helpful. Here are some notable alternatives to snscrape:

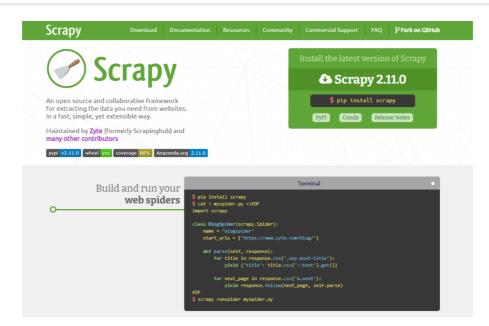
#### a. Twint

Twint is another popular Python library for scraping

Twitter data. It's known for its ability to scrape a large number of tweets without needing Twitter's API or any authentication.







TOP 🔨

Image credit: Scrapy

<u>Scrapy</u> is a more general <u>web scraping</u> framework in Python. While it's not specifically designed for social media, it's incredibly powerful for extracting data from any website.

Scrapy is suitable for complex scraping tasks, and it offers extensive customization and control over your scraping jobs.

It's ideal for users who have more advanced programming skills and need to scrape data from a variety of sources.

#### c. BeautifulSoup

BeautifulSoup is a Python library for parsing HTML and XML documents.

It's often used in combination with a request library to scrape data from web pages.

While it requires more setup compared to snscrape, BeautifulSoup offers great flexibility and is powerful in







Image credit: Octoparse

Octoparse is a user-friendly, point-and-click data extraction tool that doesn't require any coding skills.



It's suitable for non-programmers or those who prefer a graphical interface for scraping tasks.

Octoparse can handle both simple and complex data extraction from various types of web pages, including social media sites.

#### e. Data Miner

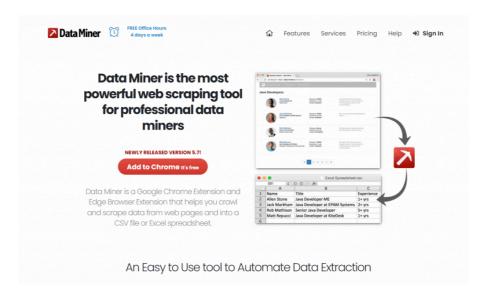


Image credit: Data Miner

<u>Data Miner</u> is a Chrome and Edge browser extension that allows you to scrape data from web pages and into a variety of file formats, including Excel and Google Sheets.

It's very user-friendly and suitable for those who need to scrape data quickly without writing any code.





Image credit: ParseHub

TOP 🔨

ParseHub is a visual data extraction tool that is equipped with machine learning technology to identify, extract, and transform data from web pages. Check our ParseHub review.

It's a powerful tool for scraping complex websites and can handle websites with JavaScript and AJAX.

Each of these alternatives has its strengths and use cases.

Depending on your specific needs, such as the type of data you need, its volume, your technical expertise, and the platforms you're targeting, one of these tools might be more suitable for your project than snscrape.

## 9. Final Words

Snscrape is a great tool for social media data extraction. With basic Python skills, you can customize your data scraping to suit a wide range of needs.

Just make sure you use snscrape properly and ethically to avoid any issues.

Looking to enhance your scraping with snscrape?

Boost your web scraping efficiency with



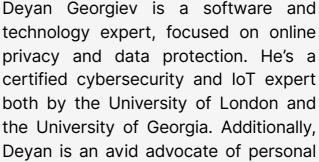


# Optimize your scraping with RapidSeedbox today!

TOP ^

#### **About author Deyan Georgiev**





specialization from Infosec.

data protection. He also holds a privacy

in



Search...

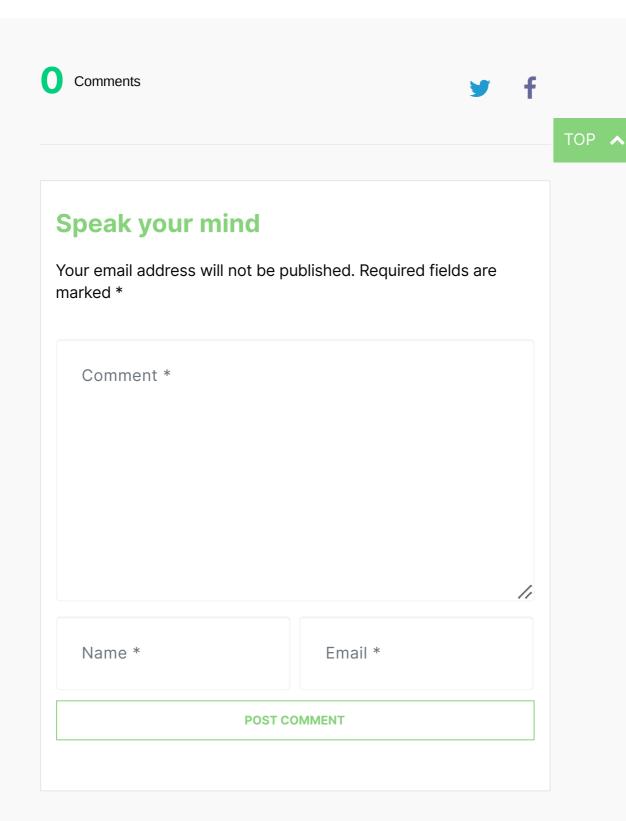
#### Join 40K+ Newsletter Subscribers

Get regular updates regarding Seedbox use-cases, technical guides, proxies as well as privacy/security tips.

Enter your email address











Quick Links	Company	
Knowledge Base	About Us	
Blog	Contact Us	
FAQ	Jobs	TOP ^
Become an affiliate	Facebook	
Dedicated Servers	Twitter	
Scholarship for CS Students	YouTube	
IP Tools	LinkedIn	