Accessing Twitter Data

Zachary Steinert-Threlkeld

Luskin School of Public Affairs, UCLA

09.08.2020

1. Introduction
2. Make Account
3. Create Application
4. Generate Token

Introduction

Me + Twitter + Research

Geolocated tweets to understand protests

- Streaming API for tweets in real time, multiple connections (apps) for multiple tasks.
- REST API for follower, friend relationships.
- REST API for tweet history.

!(Me + Twitter + Research)

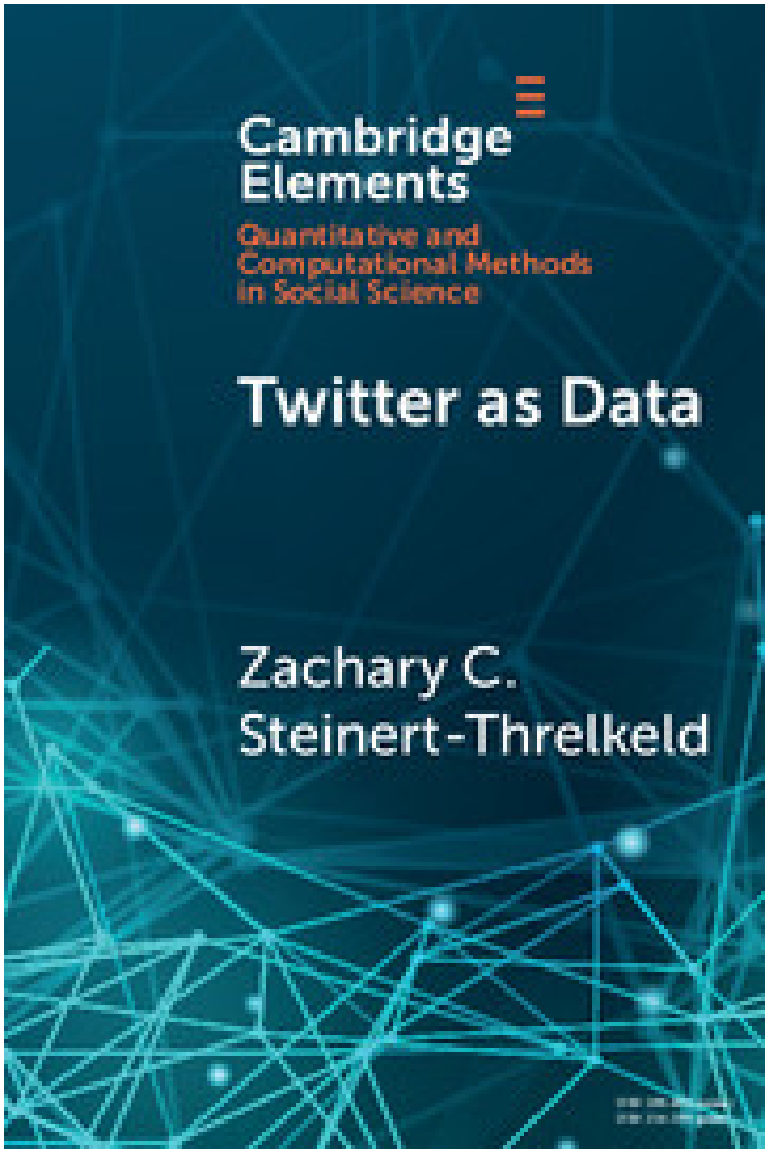
- Search API
- Experiments (bots)

!(Me + Twitter + Research)

- Search API
- Experiments (bots)

As of July 2020, Twitter is deploying a major update, v2, of its API. This course assumes v1, which is still the standard.

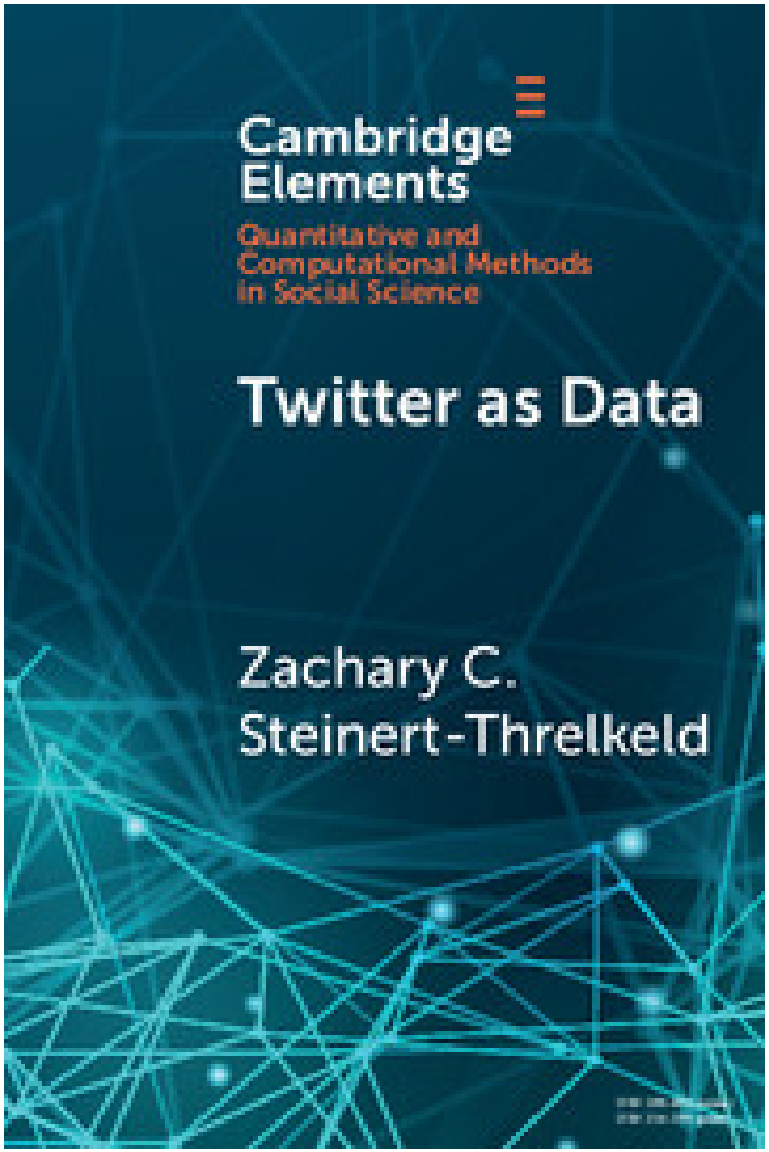
Self-Promotion



Twitter as Data

- 2018, largely developed 2014-2016.
- R + Python.

Self-Promotion



Twitter as Data

- 2018, largely developed 2014-2016.
- R + Python.
- Old: idiosyncratic technical features (rate limits, API endpoints), application creation was simpler.
- Today: Updated app creation; `rtweet`.

Pacing

If you want or need to create a Twitter developer account, please try to follow the steps as I present them. - Do not hesitate to ask a question or ask me to slow down.

If you are not able to create an app in time for the next class, do not worry.

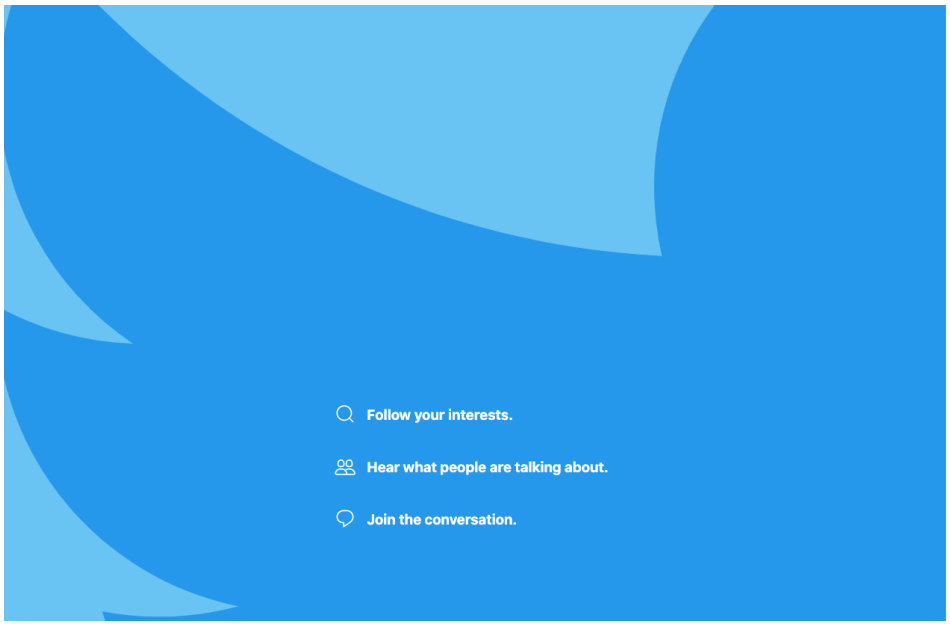
Pacing

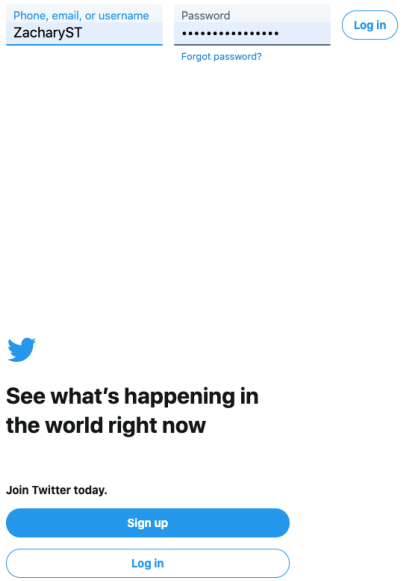
If you want or need to create a Twitter developer account, please try to follow the steps as I present them. - Do not hesitate to ask a question or ask me to slow down.

If you are not able to create an app in time for the next class, do not worry. - I expect you will not get the application for a developer account. - `rtweet` lets you authenticate as a user, meaning you can work on your own Twitter code in the next lecture.

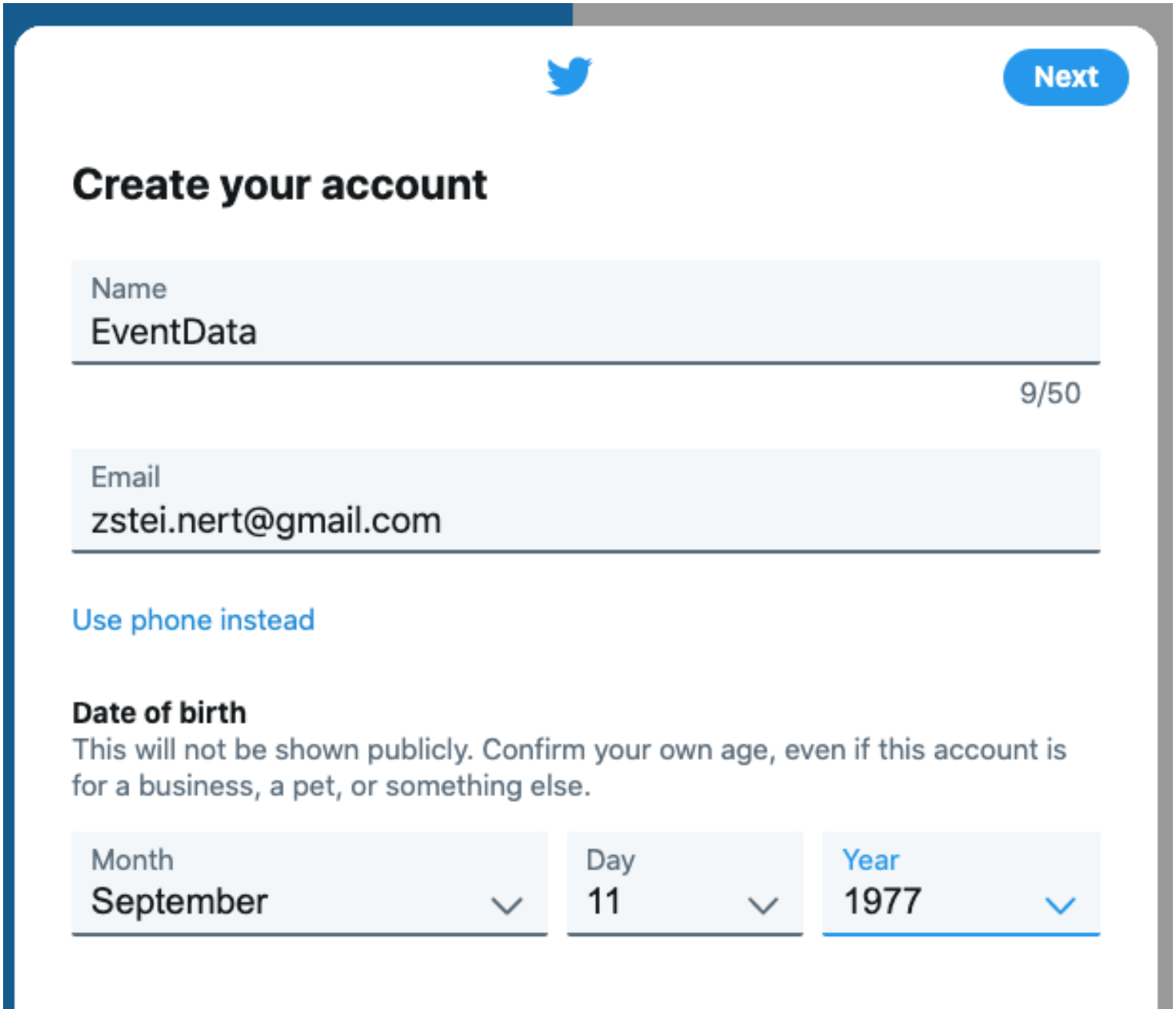
Make Account

Website

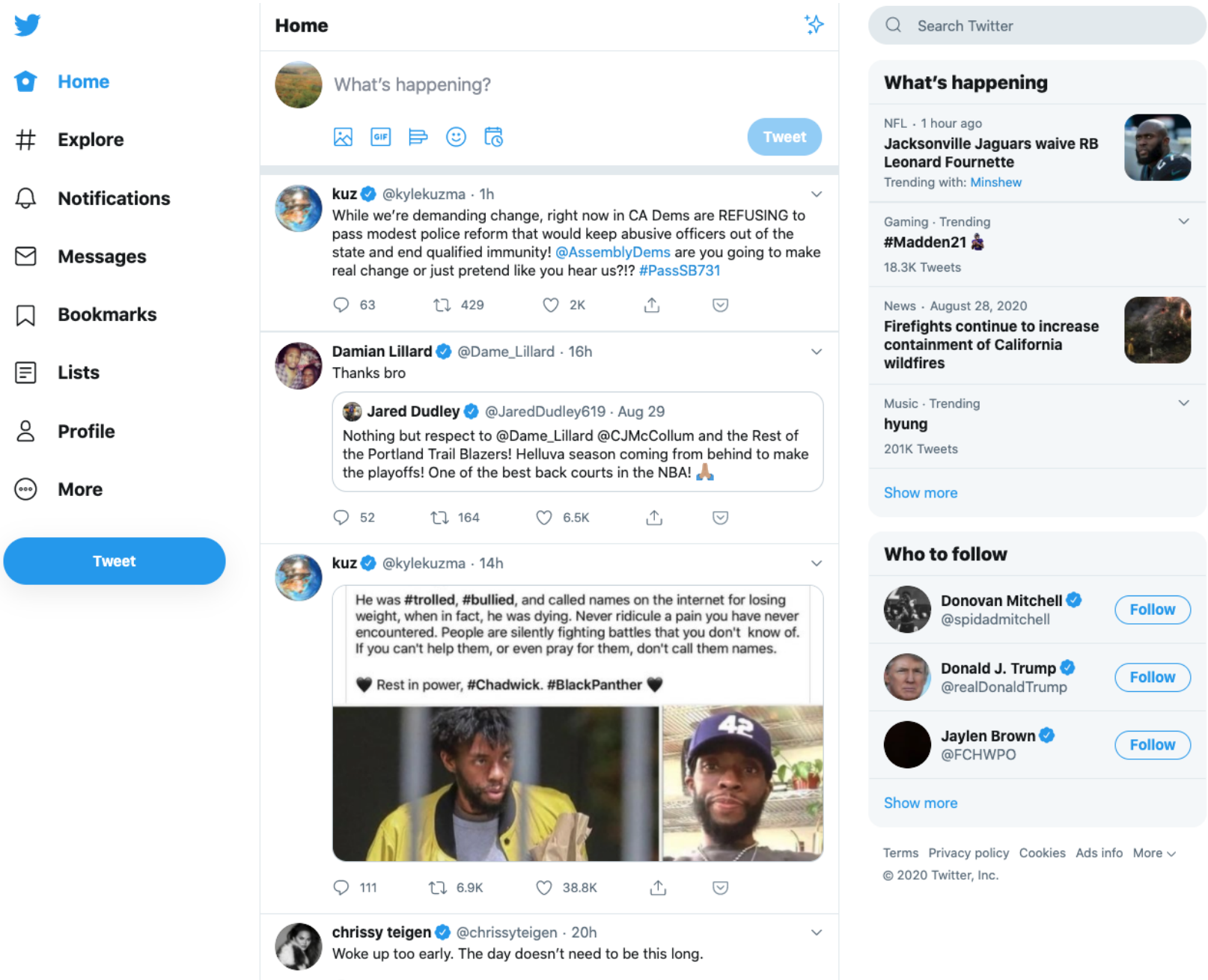

Follow your interests.
Hear what people are talking about.
Join the conversation.


Phone, email, or username
ZacharyST
Password
Log in
Forgot password?
See what's happening in the world right now
Join Twitter today.
Sign up
Log in

Create Account

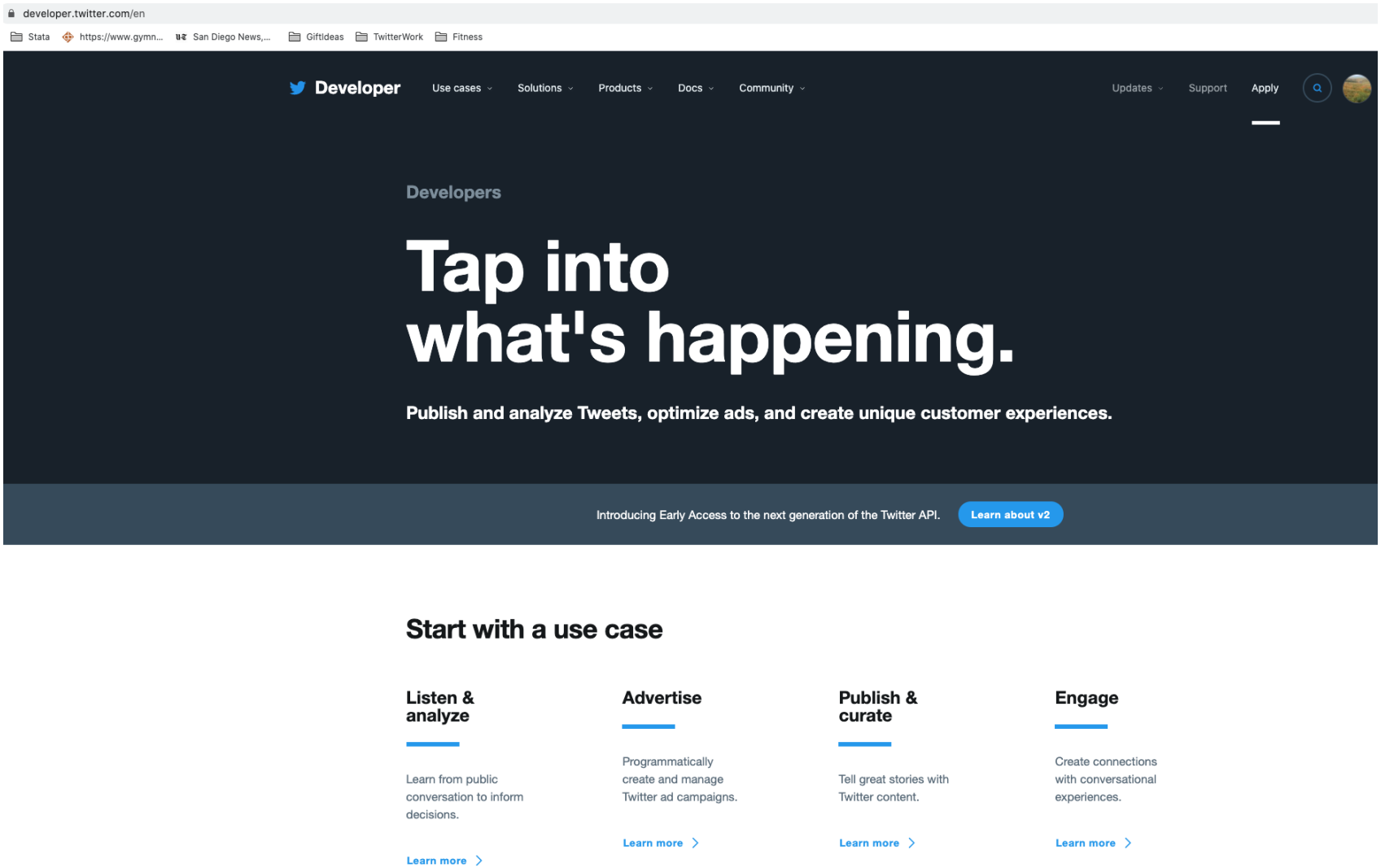


Create Account



Developer Screen

developer.twitter.com



Apply for Developer Account

First time ever, apply for developer account.

Get started with Twitter APIs and tools

Apply for access

All new developers must apply for a developer account to access Twitter APIs. Once approved, you can begin to use our standard APIs and our new premium APIs.

Apply for a developer account Restricted used cases >

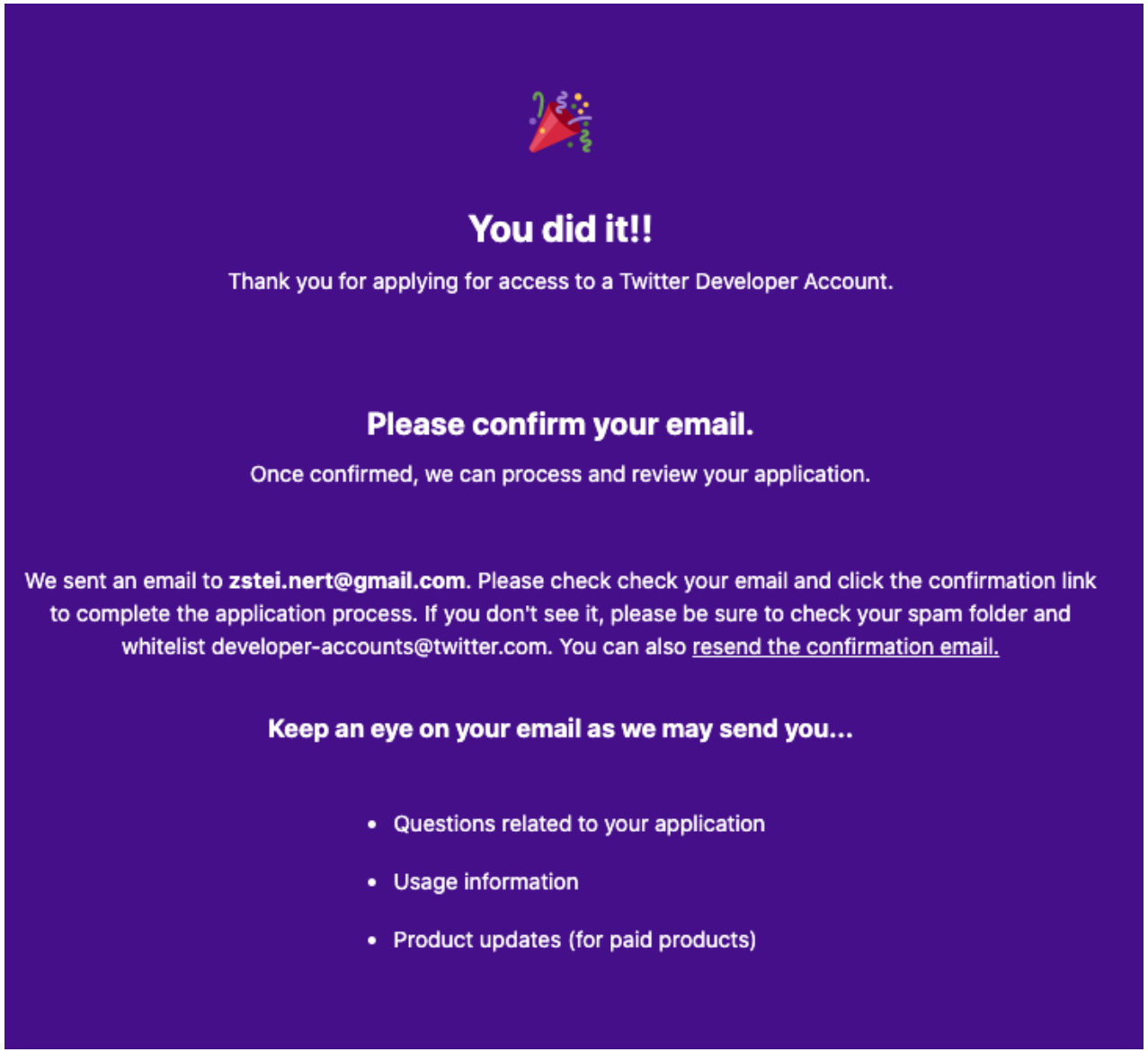
New premium APIs.

Our new premium APIs bring the reliability, stability, and access of our enterprise APIs to our broader developer ecosystem for the first time. The first premium offerings include the Search Tweets API, which includes a 30-day endpoint and a full-archive endpoint, as well as the premium Account Activity API. Learn more in the blog post and apply for a developer account to get started.

Apply for a developer account >

Apply for Developer Account

After many questions, you will see this screen.



Approval for a developer account takes a stochastic amount of days

Tips

- I prefer to have one account per e-mail address, separate browsers too if possible.

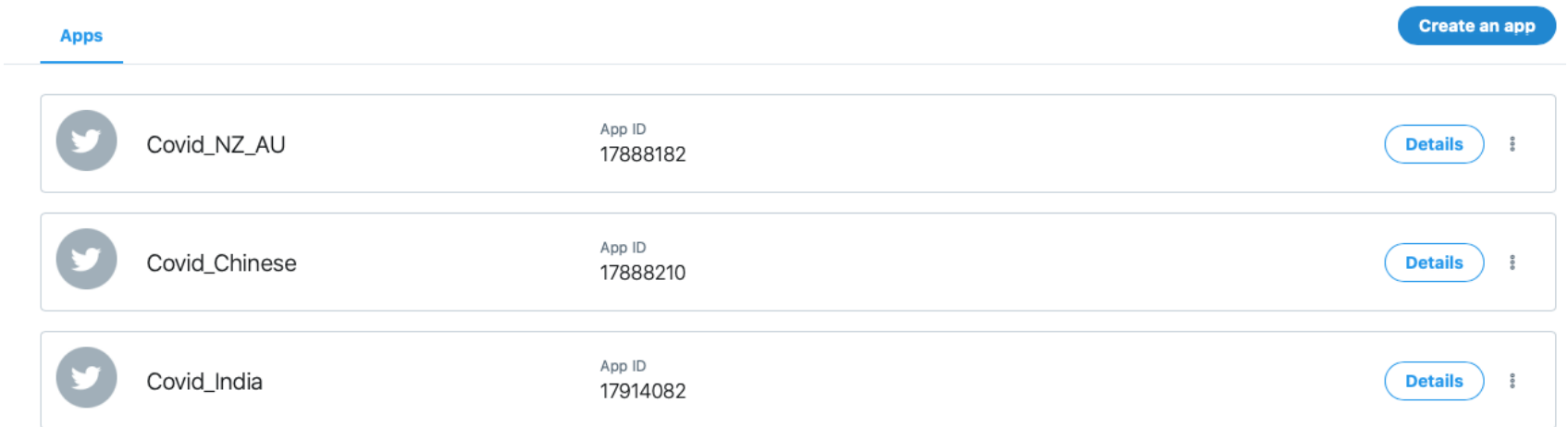
Tips

- I prefer to have one account per e-mail address, separate browsers too if possible. But with @gmail, add . within address and Google ignores the period but Twitter does not.
- If developer-only account, use semi-regularly to not look like a bot.

Create Application

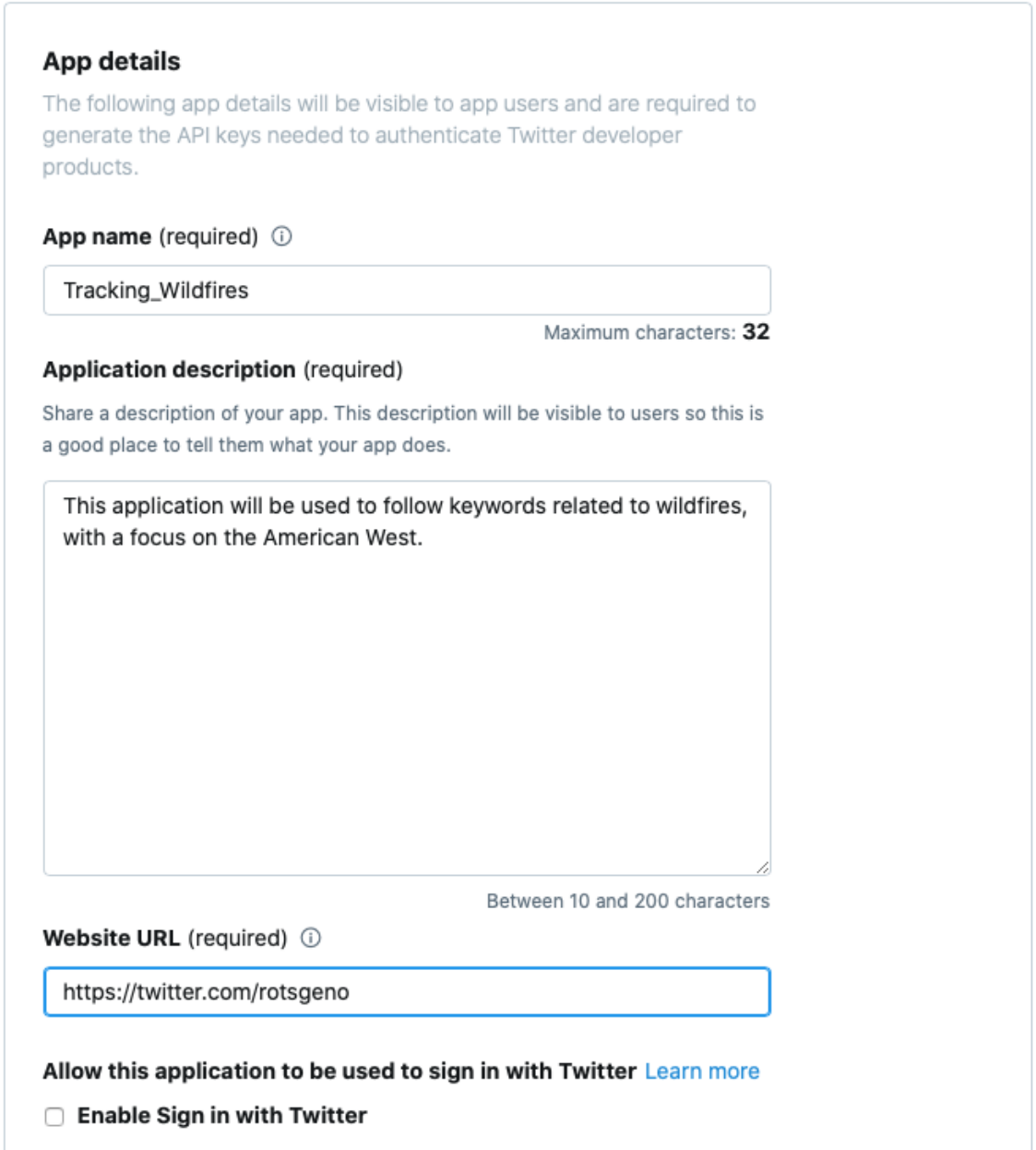
Apply

https://developer.twitter.com/en/apps/



Apply

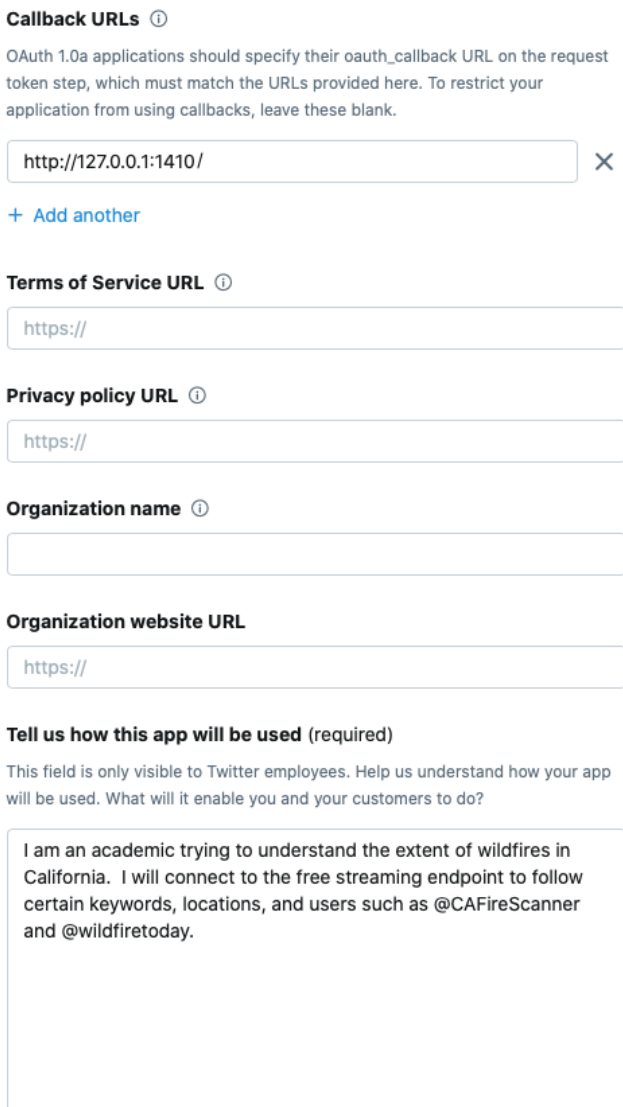
I prefer my account's Twitter homepage for the website URL.



Apply

You *must use* this callback URL for `rtweet`.

NB: `rtweet` is the first time I have seen someone require the Callback URL. No apps I have created otherwise have it, and they

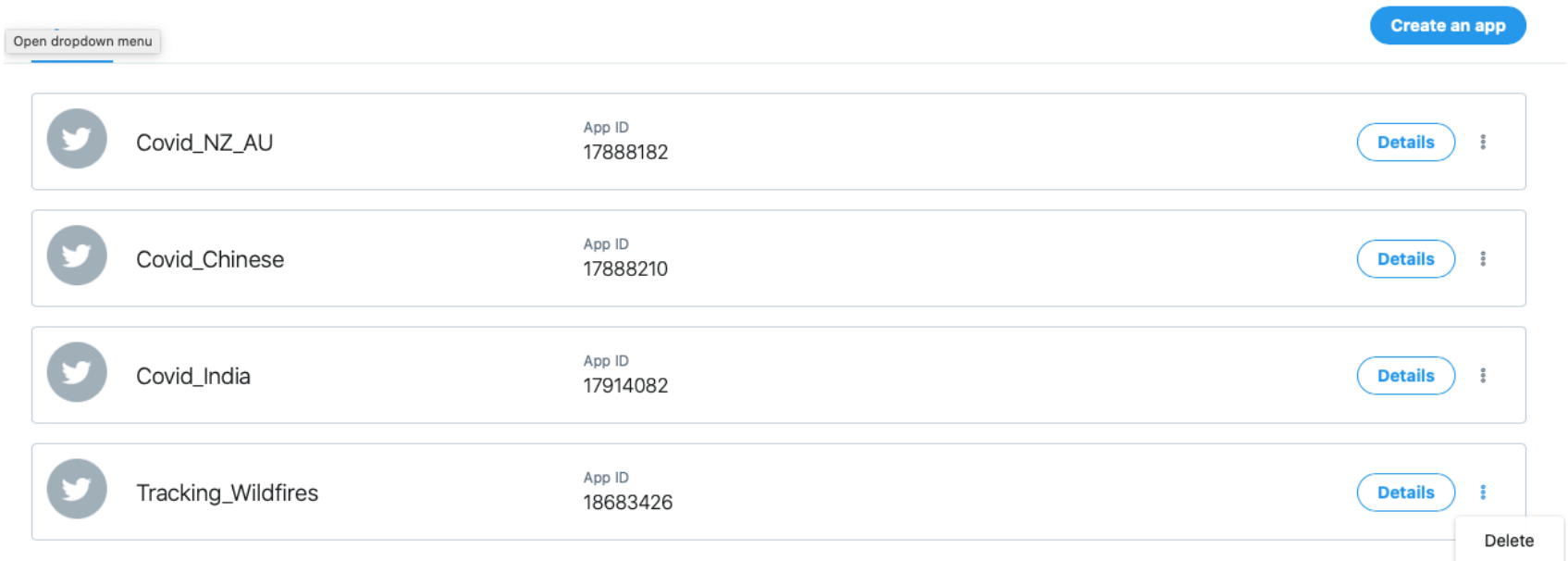


Apply

The application is now available for use.

Apply

The application is now available for use. *I am surprised it happened instantaneously.*



Generate Token

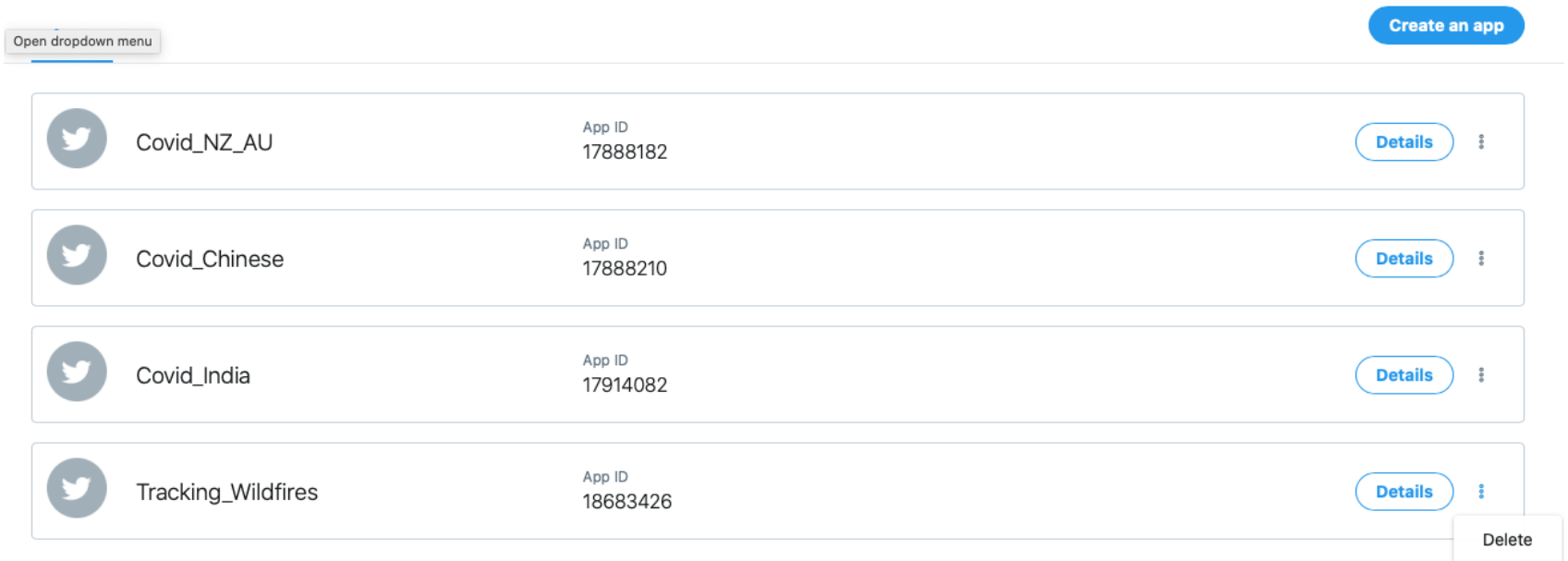
Overview

To use Twitter's API, you must create a token for the application. The token is a trusted certificate that you will later present to Twitter each time you connect to an API endpoint.
Two authentication methods.

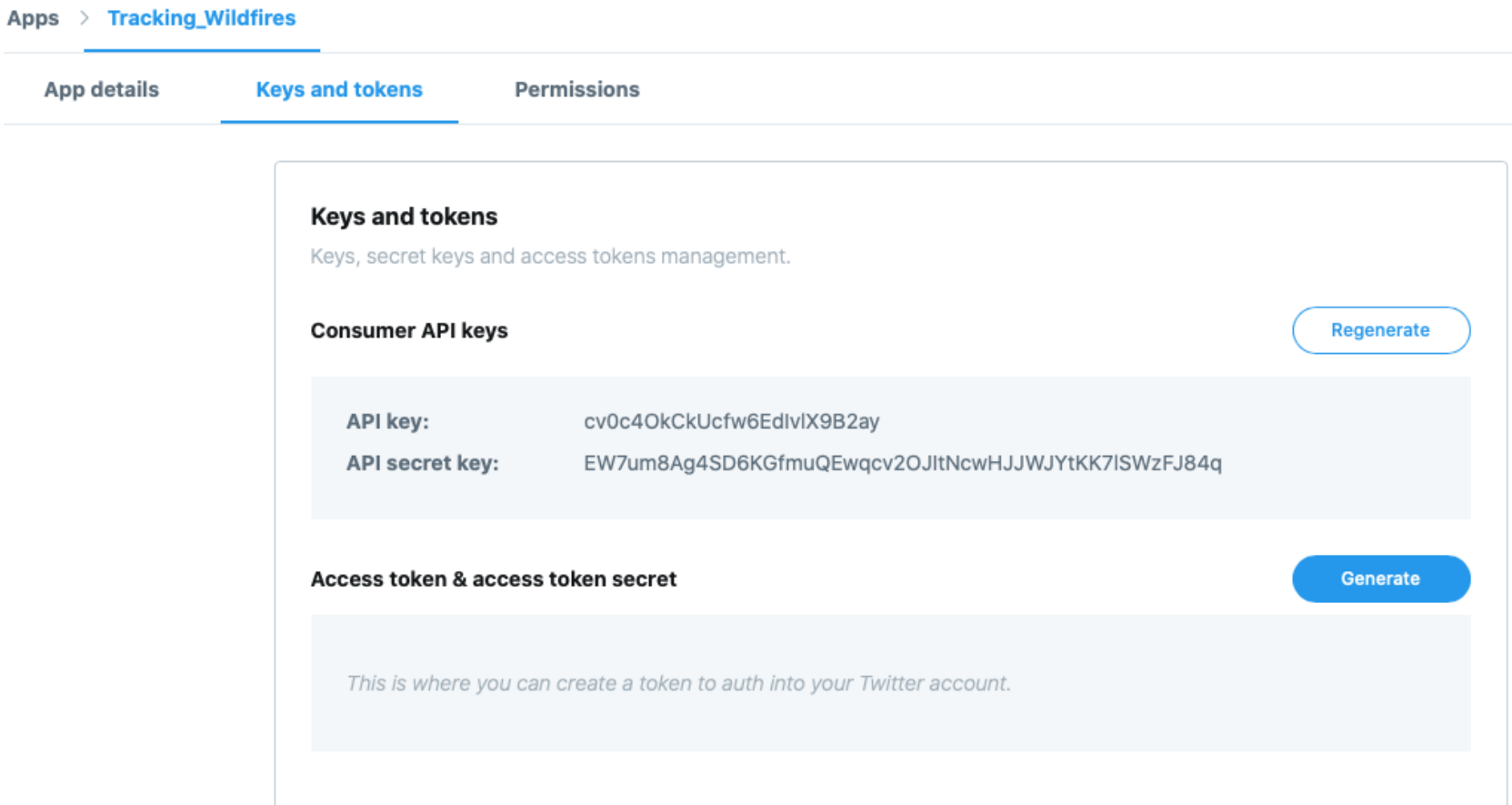
1. Browswer based from R
2. *Access token & access token secret*

I prefer #2, and it is what I show today. Both enable the same functionality, but #2 does not require an interactive session, which is important once you want to set up a persistent data collection pipeline.

Navigate to App



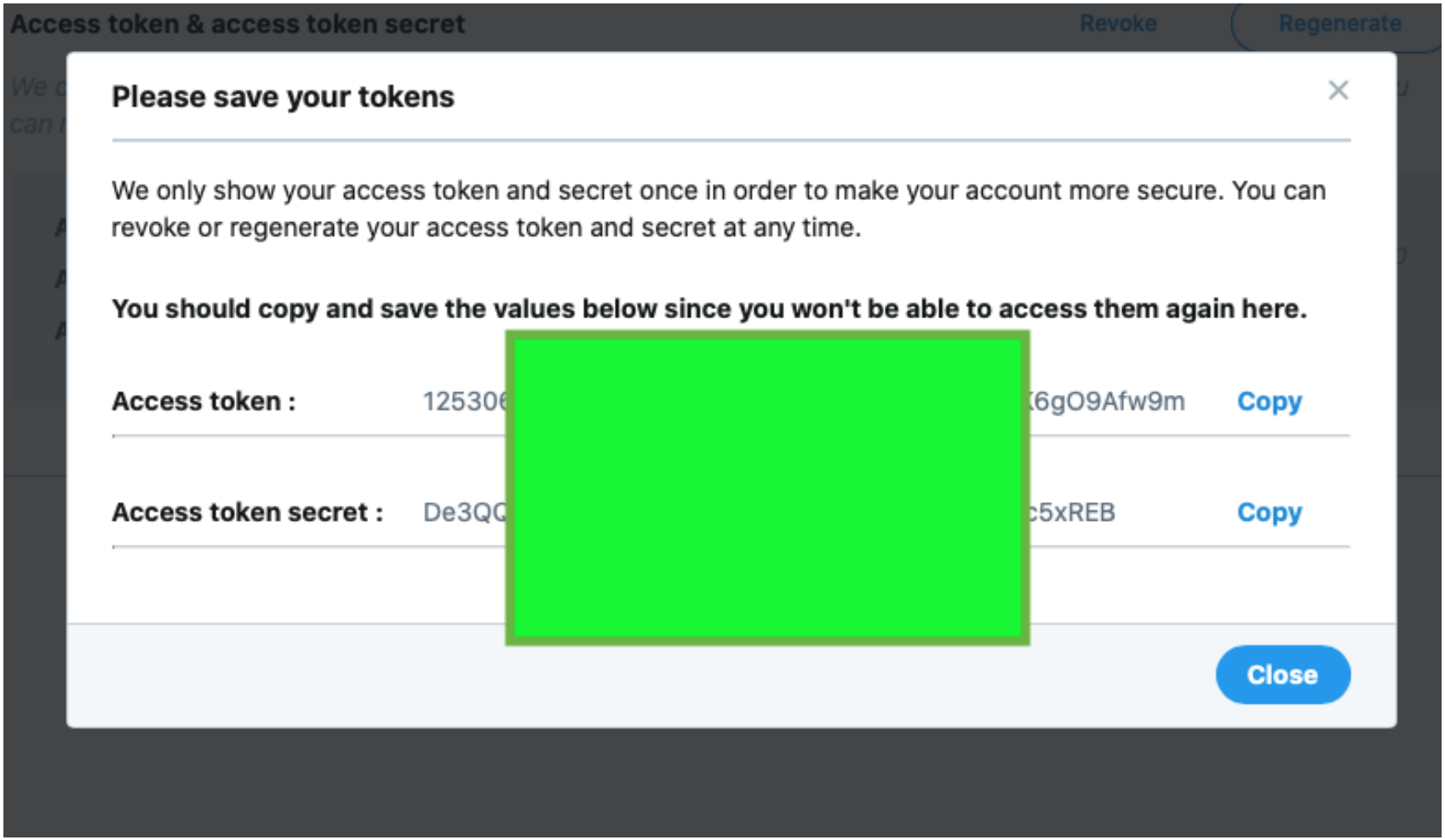
Note Consumer API Keys



Note Consumer API Keys

```
api_key <- 'CkUcfEdIvlX9c4OkBwcv062ay'
api_key_secret <- 'lS8AuQKK7g4SD6KWzFJ84qEW7Ewqcv2OJIt
NcwHJJWJYtumGfm'
```

Generate Access Token



Use Keys and Token String

```
api_key <- 'CkUcfEdIvlX9c4OkBwcv062ay'
api_key_secret <- 'lS8AuQKK7g4SD6KWzFJ84qEW7Ewqcv2OJI
tNcwHJJWJYtumGfm'
access_token <- '2K5LMiEwu2YXLK2389nslaqZoQA4Xc2uyI955
1451ZYIwaEmA94'
access_token_secret <- 'e5byw45zx7xc5d5EIPnEN1fasdf1mt
9vpi1fiFlZQpfbk2x62-w2kxKg'
```

Create Token

```
## install rtweet from CRAN
#install.packages("rtweet")

# load rtweet package
library(rtweet)

wildfire <- create_token(
  app = "Tracking_Wildfires",
  consumer_key = api_key,
  consumer_secret = api_secret_key,
  access_token = access_token,
  access_secret = access_token_secret
  )
```

`wildfire` is now the object containing the token, and rtweet has stored it as an environment variable. Future sessions will automatically load it.

Tips

- Twitter is nice to academic research, so be honest.
- Application approval process is opaque and often slow.

Tips

- Twitter is nice to academic research, so be honest.
- Application approval process is opaque and often slow. Perhaps a current favoring of COVID-19 research.
- Do not have apps duplicate work “or substantially similar or overlapping use cases".

Tips

- You may eventually have many apps and therefore many keys. Keep a document that tracks them. Do not share this document.

My Key Tracking

Date Checked	APP_KEY, APP_SECRET	OAUTH_TOKEN, OAUTH_TOKEN_SECRET	PURPOSE	IN USE?	AWS INSTANCE	Twitter App	Tested?	Works?	NOTES
04.24.2020	E11hDe69SaNaG9PRB1xAA ftvZkk8zGJTxSjKOULS5drwQ0KjqC9INQE5NcbvwkY	17367248-n8gCGbkFv7Glc5QCY9K1gqJPl4kKX2cpVx7WVjZbf AsoJhKHTD7LPm2EopuiRcVyC02kiD6DaRw0zLSFsc	Testing for HNG	N	N/A	Geocode_Stream; 4756711	Y	N	Used in Python script; revoked
04.24.2020	2hUJgevXlYdX3Ad3hWCw ZfBh7KQVGi6oK7uesCOJEZQSDZ6PJ4gWBETEkg9ww		"Troubleshooting"	N	N/A	Geocode_Stream2; 4989427	N	N	Revoked
04.24.2020	7icj78BaesOZgwz7eVQwe9f4O ScKTsl9dsGKdnXi0b5VmKcc34q7QQClflebqH5YTgrA7JQrTT0		An early grab of GPS tweets	N	N/A	HNG_AWS; 5592934	N	N	Revoked

App vs. User Auth

- You will notice, especially when checking rate limits, a distinction between user and app authentication.
- User authentication is for when a user gives your app permission to perform certain functions for it, and these functions have higher rate limits. See the process here.
- Since my research, and creating event data more broadly, does not involve users of an application, I always use app authorization only.

If Time

If we have reached this slide and time remains in the class, I will create a new application from scratch.