

# **PROGRAMA DE CIENCIAS DE LOS DATOS**

**Curso: Big Data**

**PROYECTO FINAL:**

**Streaming de tendencias con #hashtags en Twitter.**

**Parte 1: Conexión con API TWITTER + APACHE SPARK.**

**Profesor: MSc. Felipe Meza Obando**

**Alumnos:**

**Lester Salazar Viales.**

**Randal Salazar Viales.**

## Arquitectura

Para este proyecto final se va a utilizar una arquitectura Cliente – Servidor la cual consiste en:

- Se crea un socket el cual va a estar conectado al api de Twitter y que además estará extrayendo los mensajes de Twitter.
- El socket también creará un canal para que se conecte la aplicación de spark que se encargará de hacer los filtros del twitter.
- Se crea una aplicación cliente para que consulte al socket que extrae los twitter, esto con el fin de realizar el filtrado de los hashtag y revisar cuales son las tendencias en dicha red social.

### Librerías para Twitter que vamos a utilizar con Conda

Debido a la utilización de Conda para el desarrollo de la aplicación para realizar streaming se decidió que se iba a requerir de la librería Tweepy, por lo que se tuvo que instalar en el ambiente de trabajo.

### Instalación de Tweepy en Conda

Para instalar Tweepy debemos de ir a la siguiente dirección web:

<https://anaconda.org/conda-forge/tweepy>

Donde se nos mostrará una página similar a la siguiente:

The screenshot shows the Conda Forge package page for 'tweepy' version 3.8.0. The page has a light blue header with the text 'conda-forge / packages / tweepy 3.8.0' and a navigation bar with icons for home, code, files, star, and a count of 18. Below the header, there's a description: 'An easy-to-use Python library for accessing the Twitter API'. A green tab labeled 'Conda' is selected, showing details like 'License: MIT', 'Home: http://www.tweepy.org', 'Development: https://github.com/tweepy/tweepy', 'Documentation: http://tweepy.readthedocs.io', '84775 total downloads', and 'Last upload: 6 months and 24 days ago'. Other tabs for 'Files', 'Labels', and 'Badges' are visible but not selected.

En dicha página buscamos el apartado donde se encuentran los pasos para instalar la librería tweepy.

Ahí se nos indican los comandos que se pueden utilizar para instalar tweepy en Conda.

## Installers

Info: This package contains files in non-standard labels.

### conda install ?

linux-64	v3.5.0
win-32	v3.6.0
noarch	v3.8.0
win-64	v3.6.0
osx-64	v3.6.0

To install this package with conda run one of the following:

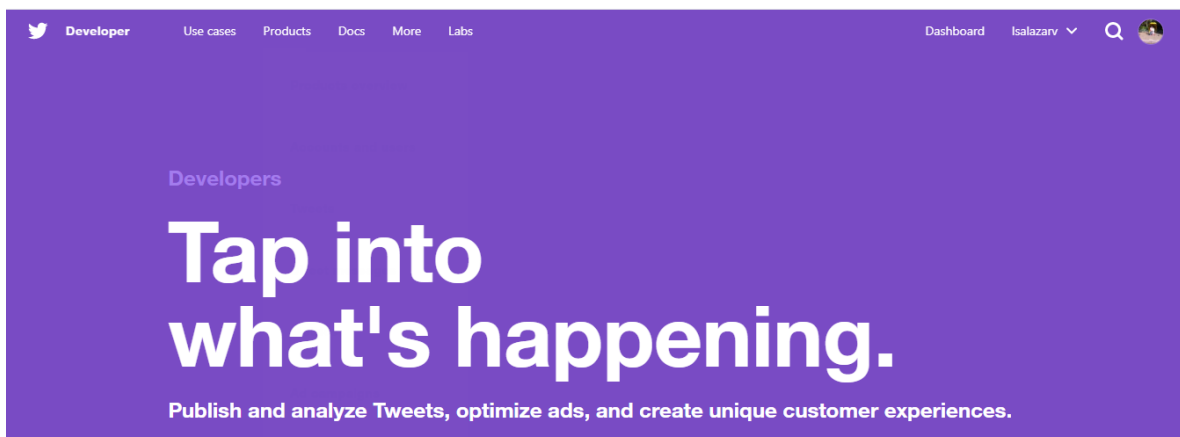
```
conda install -c conda-forge tweepy
conda install -c conda-forge/label/cf201901 tweepy
```

## Api Twitter

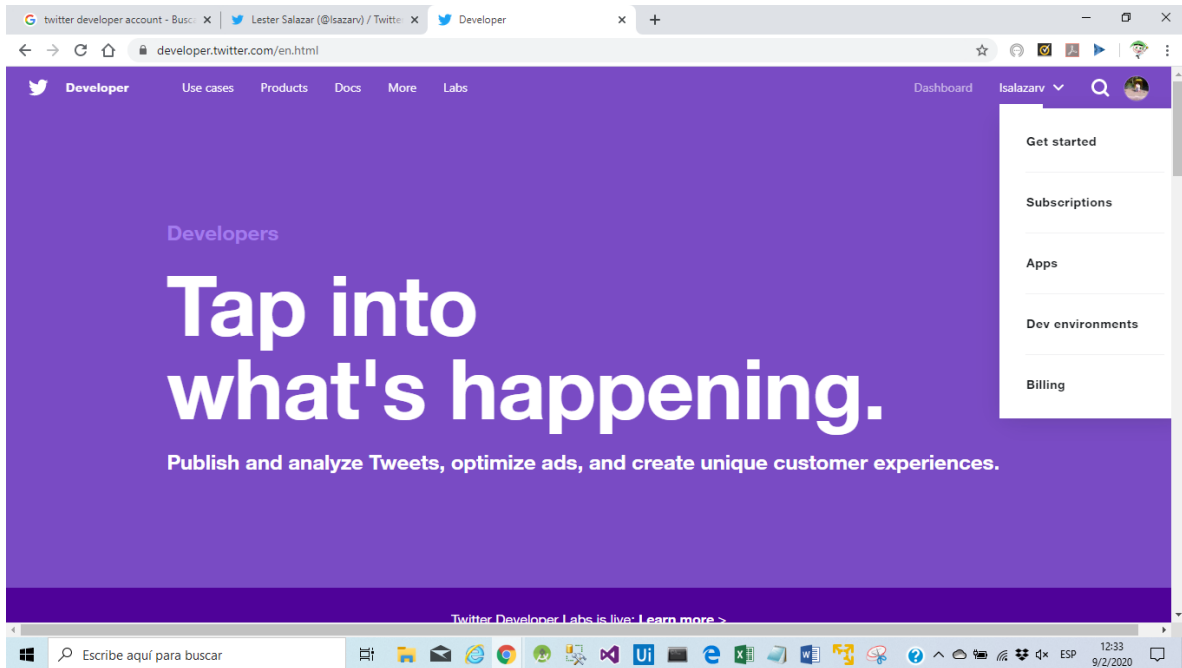
Para el caso de nuestro proyecto decidimos utilizar el API de Twitter. Por lo que decimos ingresar al sitio para developers de Twitter en la siguiente dirección web.

<https://developer.twitter.com/>

Lo que nos lleva a una página similar a la siguiente.



Una vez en dicha pantalla hacemos click sobre el nombre de usuario, lo que nos despliega un menú como el que se ve en la imagen, en dicho menú seleccionamos la opción de App para crear la app que vamos a utilizar del API.



Una vez presionado la opción de App nos aparecerá una pantalla similar a la siguiente donde se nos da la opción para crear la app que vamos a utilizar.



Se presiona el botón que dice “Create an app” que nos lleva a la pantalla donde procederemos a crear nuestra app de twitter.

**Organization website URL**

**Tell us how this app will be used** (required)

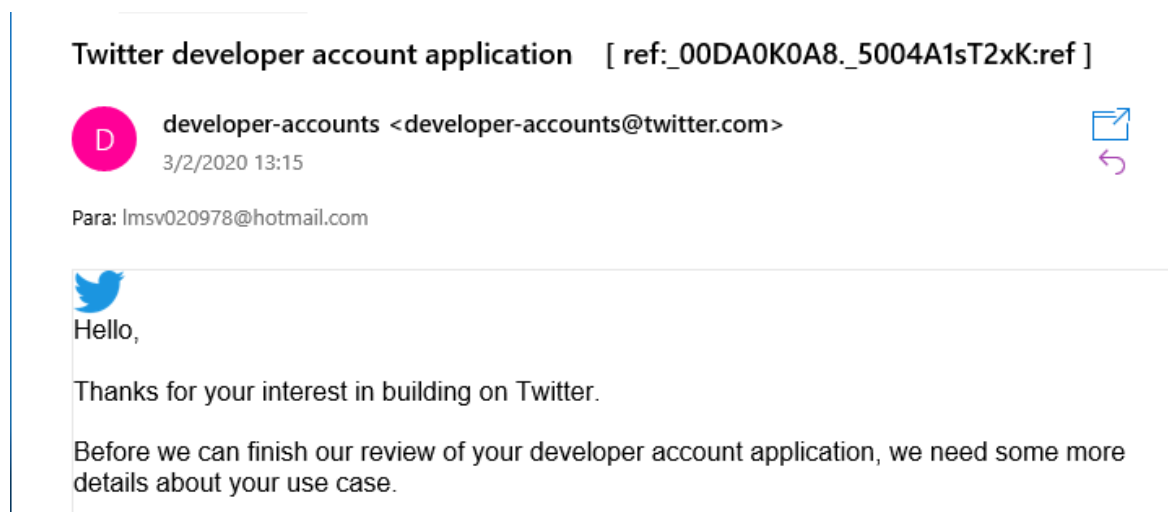
This field is only visible to Twitter employees. Help us understand how your app will be used. What will it enable you and your customers to do?

Please be detailed.

Minimum characters: **100**

En dicha pantalla se llenan los datos tales como el nombre de la app y el uso que se le va a dar. Una vez que los datos han sido ingresados se procede a presionar el botón que dice “Create”, lo que envía un correo a la cuenta que se utilizó para crear la cuenta de Twitter.

El correo es muy similar al de la imagen siguiente, en caso de que requieran más explicación del uso de la cuenta de desarrollo, se deben de responder unas preguntas que envían de la empresa.



Una vez que las preguntas son respondidas y validadas por el personal de twitter nos envían un correo donde nos indican que la cuenta fue aprobada y que ya puede ser utilizada.

## Account Application Approved



Twitter Developer Accounts <developer-accounts@twitter.com>

3/2/2020 16:26



Para: Lester Salazar



Developer

Account Application Approved

## Your Twitter developer account application has been approved!

Thanks for applying for access. We've completed our review of your application, and are excited to share that your request has been approved.

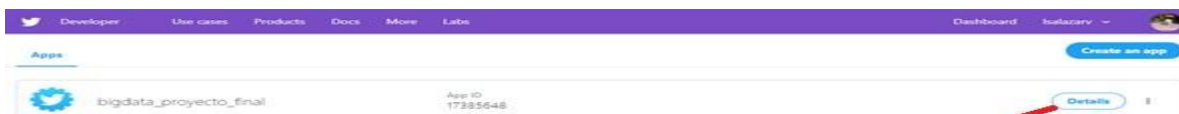
Sign in to your [developer account](#) to get started.

Thanks for building on Twitter!

### Generación Token/Keys

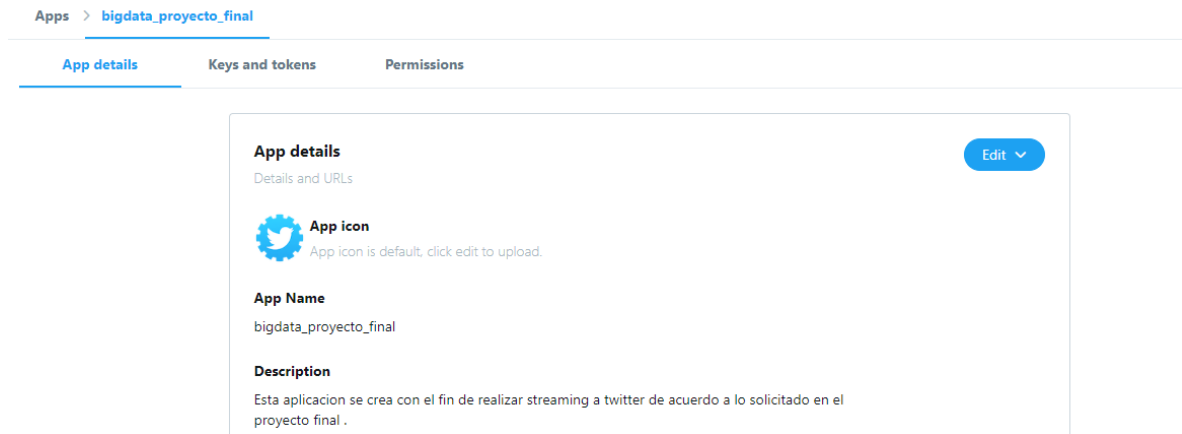
Una vez creada la aplicación se procede a crear los tokens y los keys que se van a utilizar en la aplicación básica que va a realizar streaming de twitter.

Para generar los tokens presionamos el botón que dice "Details" en el app que creamos recientemente

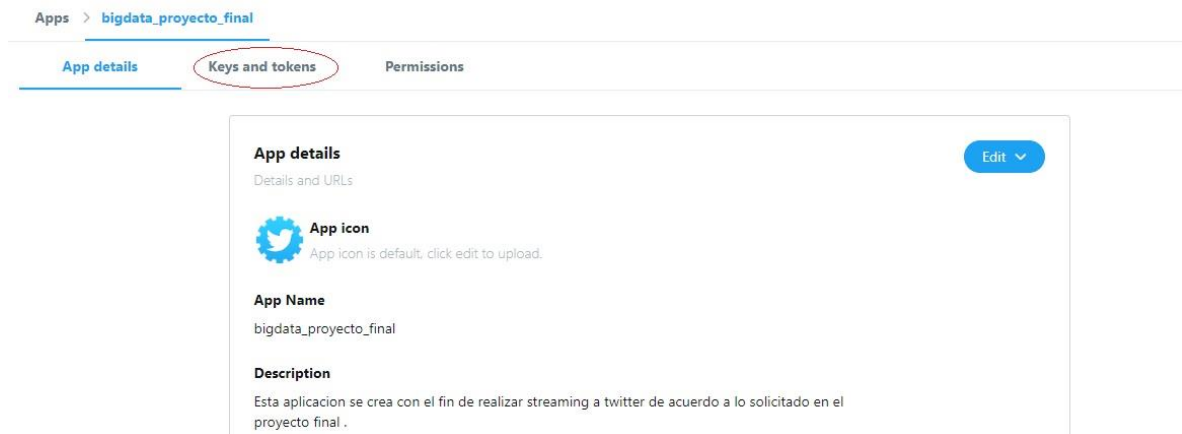


*Botón para ingresar al detalle del App para generar los token y keys*

Cuando se presiona dicho botón nos lleva a la siguiente pantalla similar a la que se muestra en la siguiente imagen



En la pantalla de detalle de la app seleccionamos el tab que dice “Keys and tokens” para generar los tokens y los keys que necesita nuestra aplicación para poder conectarse al api de twitter.



Una vez en la pantalla de Keys and tokens procedemos a generar los consumer API keys y los Access token que va a ocupar nuestra aplicación que va a estar realizando streaming a Twitter por medio de la Api.

Looking for your secret token? For security, API tokens are only displayed once. You will need to regenerate access tokens for previously authenticated apps. To learn more,

Keys, secret keys and access tokens management.

## Regenerate

**API secret key:** wBiDmKBKAd9ZTiE7P4PdMF49r0yoGIsFuXRJIW69OGAuZMyJKC

Revoke

*We only show your access token and secret when you first generate it in order to make your account more secure. You can revoke or regenerate them at any time, which will invalidate your existing tokens.*

## Consumer API keys

Regenerate

**API secret key:** wBiDmKBKAd9ZTiE7P4PdMF49r0yoGlsFuXRJIW69OGAuZMyJKC

### Access token & access token secret

Revoke

Regenerate