



# Nội dung

---

1. Giới thiệu Spark Streaming
2. Spark Streaming Fundamentals
3. Làm việc với PySpark Streaming





# Làm việc với PySpark Streaming

---

## □ Cài Tweepy

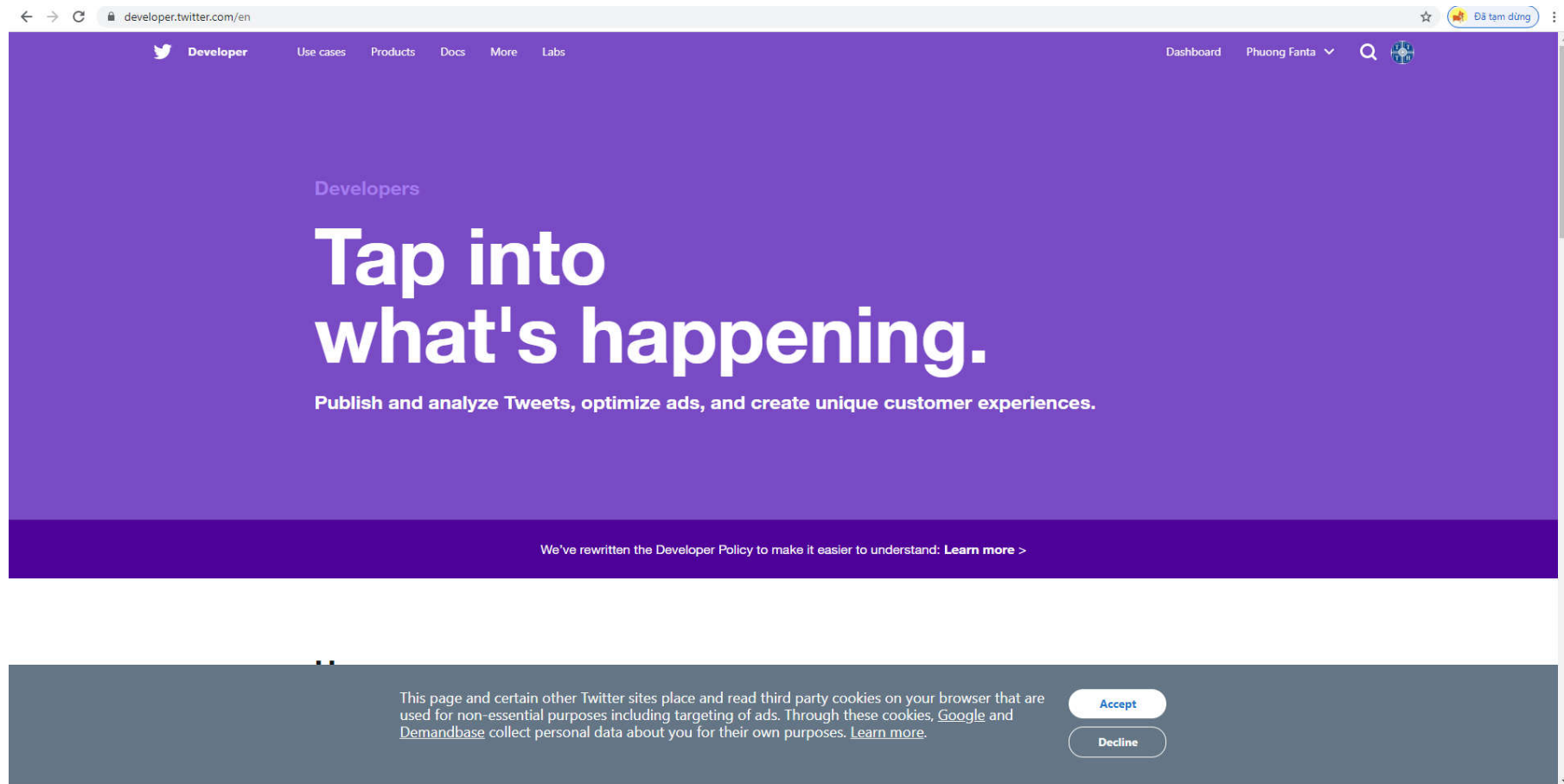
- Cài đặt: `pip install tweepy`



# Làm việc với PySpark Streaming

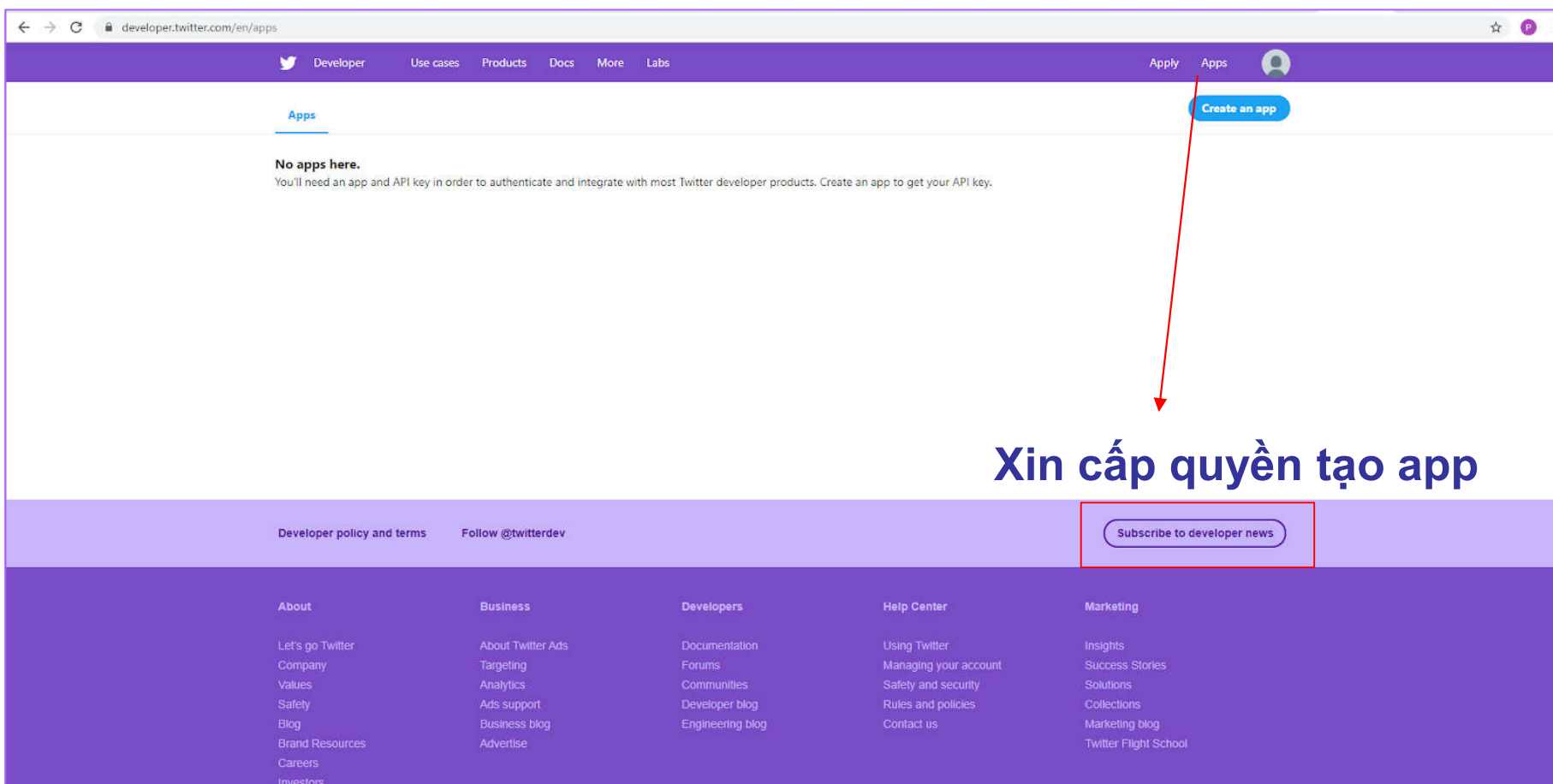


❑ Vào <https://developer.twitter.com/en>



# Làm việc với PySpark Streaming

## ❑ Tạo Twitter Developer Account



The screenshot shows the Twitter Developer portal at [developer.twitter.com/en/apps](https://developer.twitter.com/en/apps). The page has a purple header with navigation links: Developer, Use cases, Products, Docs, More, Labs, Apply, Apps, and a user profile icon. A blue button labeled "Create an app" is visible in the top right. The main content area has a sub-header "Apps" and a message: "No apps here. You'll need an app and API key in order to authenticate and integrate with most Twitter developer products. Create an app to get your API key." A red arrow points from the "Create an app" button to the text "Xin cấp quyền tạo app". At the bottom, there is a footer with links for Developer policy and terms, Follow @twitterdev, and a "Subscribe to developer news" button. The footer also contains a grid of links categorized under About, Business, Developers, Help Center, and Marketing.

**Xin cấp quyền tạo app**

# Làm việc với PySpark Streaming

## ☐ Chọn/điền các thông tin

developer.twitter.com/en/application/use-case

Developer Use cases Products Docs More Labs Apply Apps

### Get access to the Twitter API

**#welcome**

We're excited you want to use Twitter APIs and data!

As a developer platform, our first responsibility is to our users: to provide a place that supports the health of conversation on Twitter.

**This application process helps us to:**

1. Prevent abuse of the Twitter platform.
2. Better understand and serve our developer community.

**Thank you for your time and thoughtful responses.**  
Applications are final once submitted and can't be edited.

Professional ...for commercial uses	Hobbyist ...for a personal project	Academic ...for education or research	Other I don't fit any of those
Building B2B products	Making a bot	Doing academic research <input checked="" type="radio"/>	Embedding Tweets on a website
Building consumer products	Building tools for Twitter users	Teaching	Doing something else
Build customized solutions in-house	Exploring the API	Student	
Publishing ads programmatically			

© 2019 TWITTER, INC. PRIVACY COOKIES TERMS OF SERVICE DEVELOPER POLICY & TERMS

Next



# Làm việc với PySpark Streaming

## ❑ Chọn/điền các thông tin

developer.twitter.com/en/application/organization

Developer Use cases Products Docs More Labs Apply Apps

### Get access to the Twitter API

Twitter @username > **Organization** > Intended use > Review > Terms

#### Team developer account

You are signing up for a team developer account.

These are typically used for:

- companies
- organizations
- educators
- group collaboration

If you do not think you will need to invite other people to your account in the future to share API access or apps, you can [create an individual developer account](#) instead.

**Team name**  
This will be the name of your account  
Phuong Khuat

**Legal entity name**  
Of company, institution or parent organization  
Computer Science Center

**Organization Twitter @username**  
@ PhuongKhuat2

**Website URL (optional)**  
csc.edu.vn

**Organization primary country of operation**  
Vietnam

**How do you categorize your organization?**  
Academic

**What industries do you / will you serve?**  
Select all that apply  
Academic Add...

Do you currently have customers? Yes No

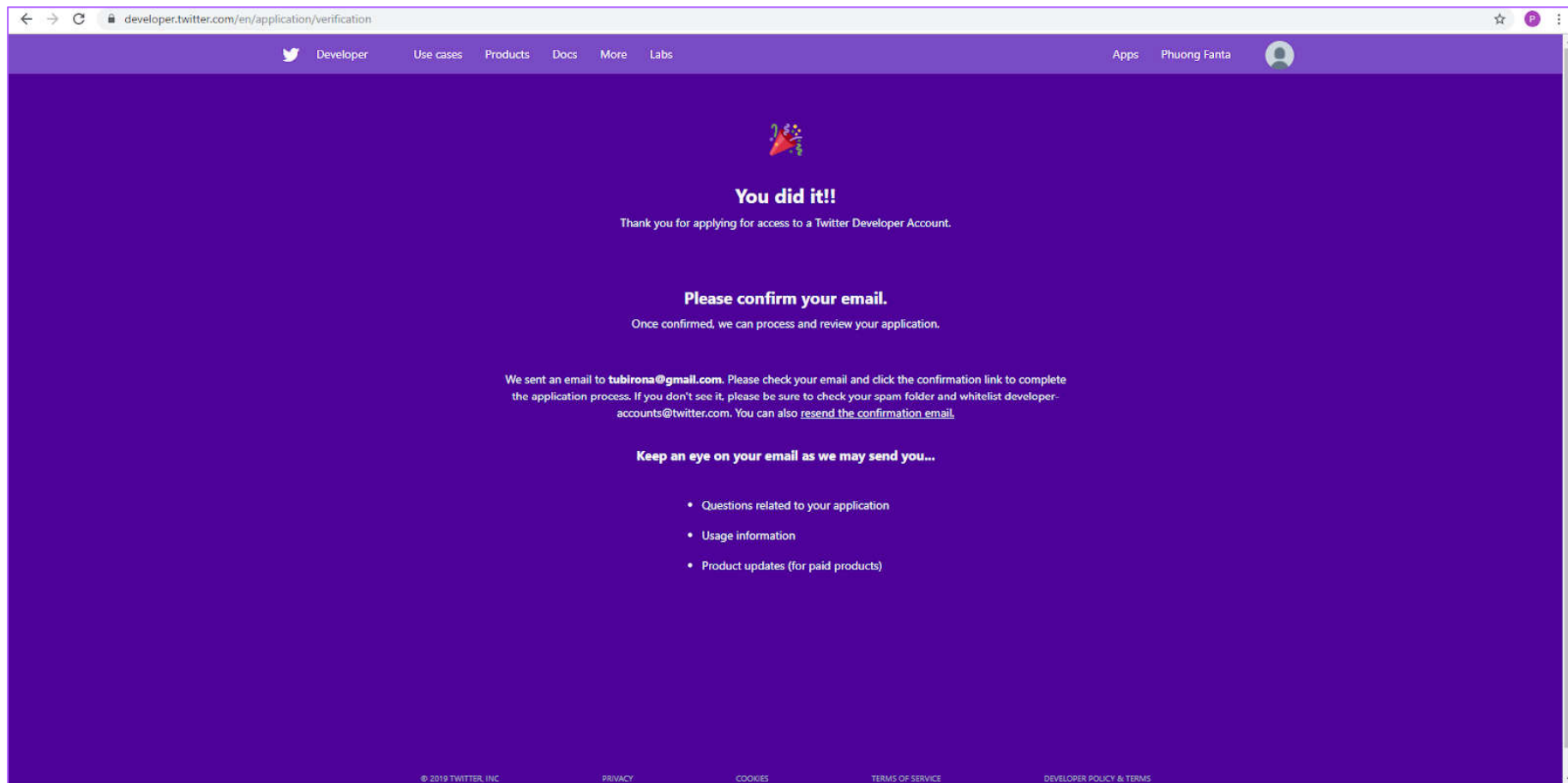
Back Next



# Làm việc với PySpark Streaming



## ❑ Confirm your email



# Làm việc với PySpark Streaming



## □ Tạo app

Developer

Use cases

Products

Docs

More

Labs

Dashboard

Phuong Fanta

We've launched Twitter Developer Labs! Test new Twitter API previews, provide feedback, and get the latest news on what we're building. [Learn more](#) or [sign up!](#)

Apps > [Create an app](#)

Understanding apps

What is an app?

Why register an app?

Which products require an API key?

**App details**

The following app details will be visible to app users and are required to generate the API keys needed to authenticate Twitter developer products.

**App name (required)**

ktphuong\_first\_stream

Maximum characters: 32

**Application description (required)**

Share a description of your app. This description will be visible to users so this is a good place to tell them what your app does.

My first Spark app - streaming Twitter data.

Between 10 and 200 characters

**Website URL (required)**

http://phuong13021982.

**Allow this application to be used to sign in with Twitter**

[Learn more](#)

☒ Enable Sign in with Twitter

**Callback URLs (required)**

OAuth 1.0a applications should specify their oauth\_callback URL on the request token step, which must match the URLs provided here. To restrict your application from using callbacks, leave these blank.

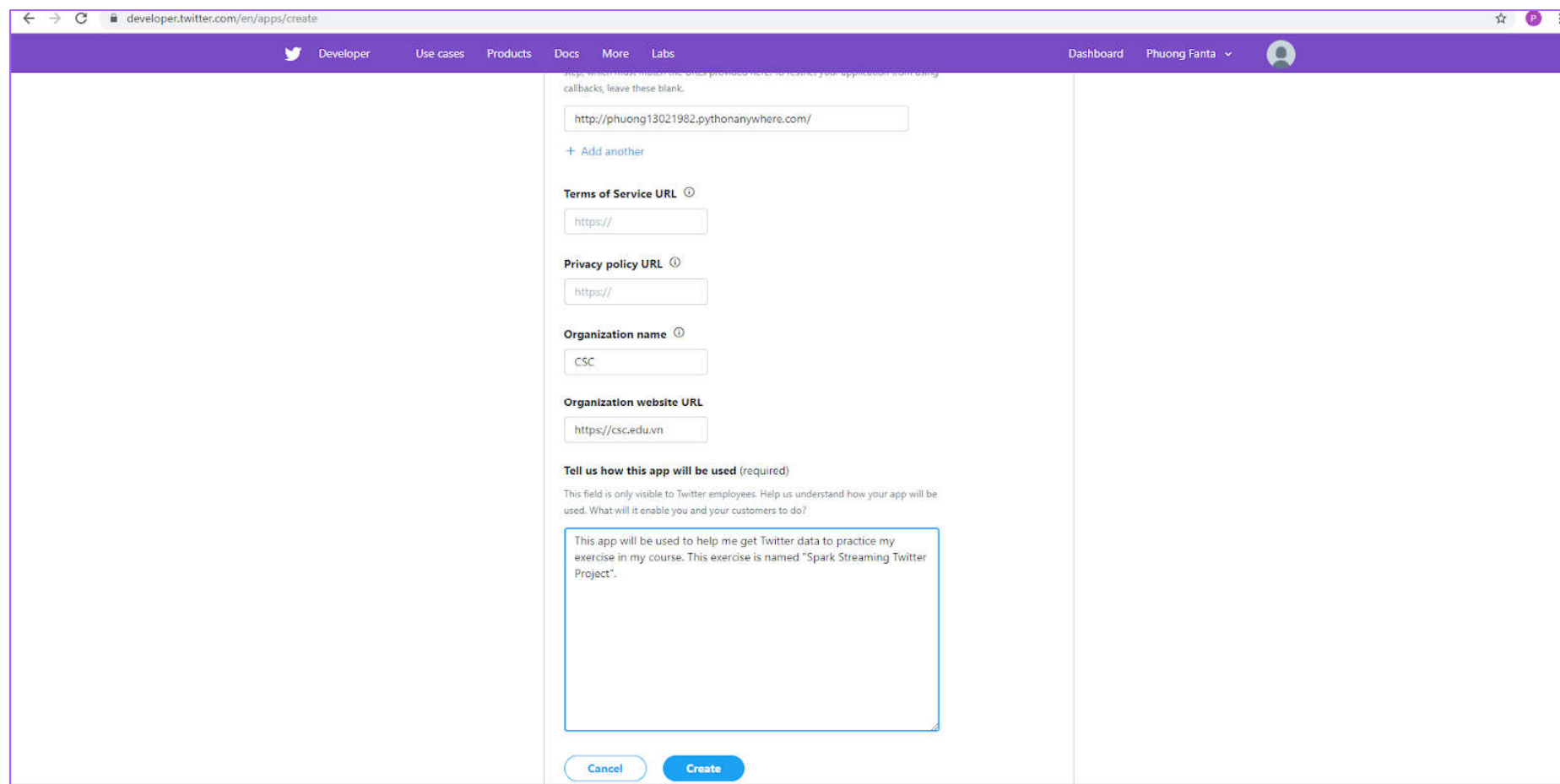
http://phuong13021982.pythonanywhere.com/





# Làm việc với PySpark Streaming

## □ Tạo app

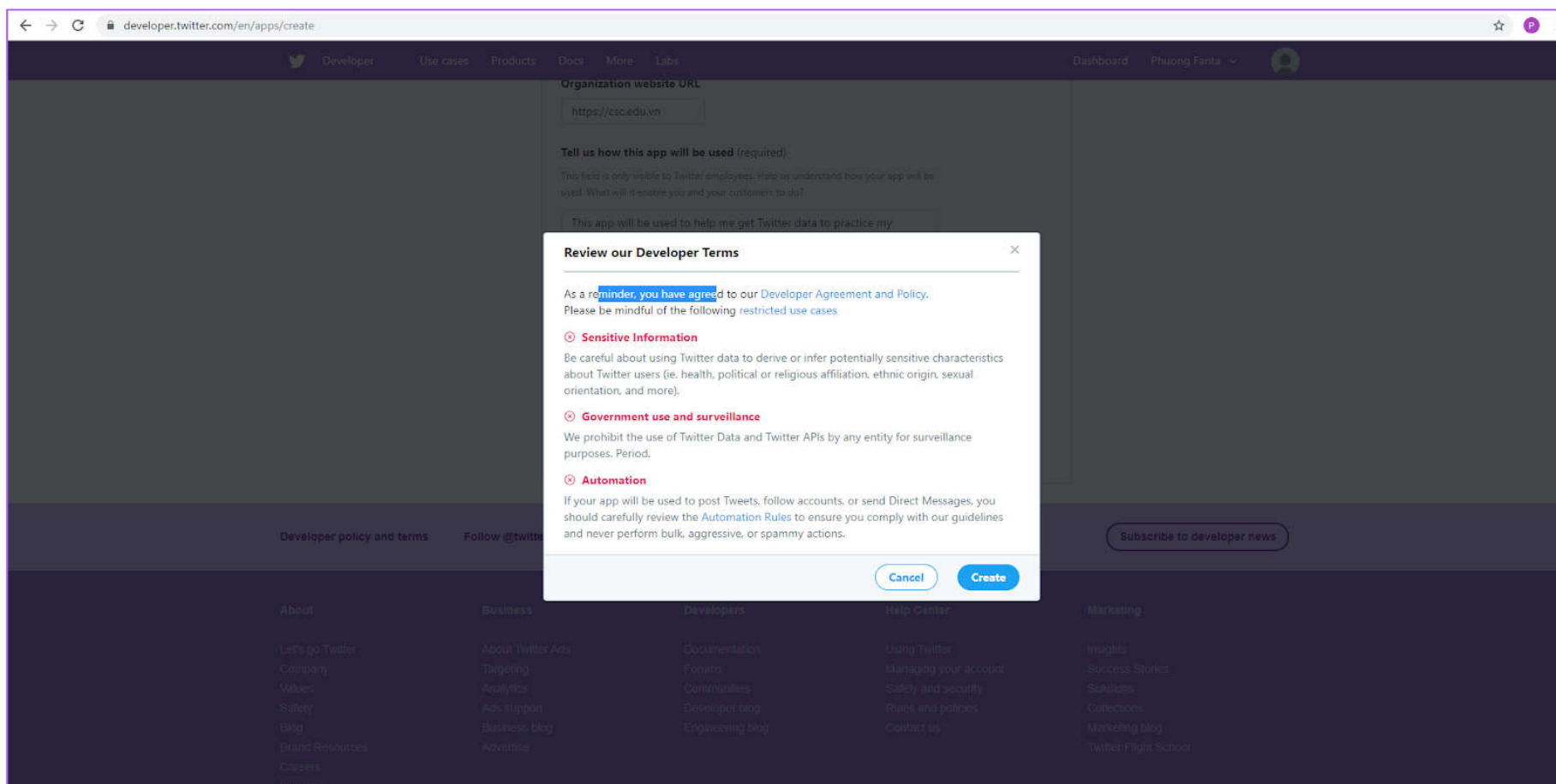


The screenshot shows the Twitter Developer app creation page. The browser address bar displays `developer.twitter.com/en/apps/create`. The page has a purple header with navigation links: Developer, Use cases, Products, Docs, More, Labs, Dashboard, and a user profile for 'Phuong Fanta'. The main content area is divided into two columns. The left column contains a large empty space for a profile picture. The right column contains the app creation form. At the top of the form, it says 'Step 1: Which most closely matches the type of application you're building?'. Below this is a text input field containing `http://phuong13021982.pythonanywhere.com/` and a '+ Add another' link. The form then has several sections: 'Terms of Service URL' with a text input field containing `https://`; 'Privacy policy URL' with a text input field containing `https://`; 'Organization name' with a text input field containing `CSC`; and 'Organization website URL' with a text input field containing `https://csc.edu.vn`. Below these is a section titled 'Tell us how this app will be used (required)' with a subtext: '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?'. There is a text area containing the text: 'This app will be used to help me get Twitter data to practice my exercise in my course. This exercise is named "Spark Streaming Twitter Project"'. At the bottom of the form are two buttons: 'Cancel' and 'Create'.



# Làm việc với PySpark Streaming

## □ Tạo app





# Làm việc với PySpark Streaming

## □ Tạo app

The screenshot shows the Twitter Developer App details page for an app named 'ktphuong\_first\_stream'. The page is divided into three tabs: 'App details', 'Keys and tokens', and 'Permissions'. The 'App details' tab is active, showing the following information:

- App details** (Details and URLs): Includes an 'App icon' (Twitter logo) and a description: 'My first Spark app - streaming Twitter data.'.
- Website URL**: <http://phuong13021982.pythonanywhere.com/>
- Sign in with Twitter**: Enabled
- Callback URL**: <http://phuong13021982.pythonanywhere.com/>
- Terms of service URL**: None
- Privacy policy URL**: None
- Organization name**: CSC
- Organization website URL**: <https://csc.edu.vn>
- App usage**: This app will be used to help me get Twitter data to practice my exercise in my course. This exercise is named "Spark Streaming Twitter Project".

A tooltip message is displayed over the app icon: "You are using the default icon now, change it in app editing mode."

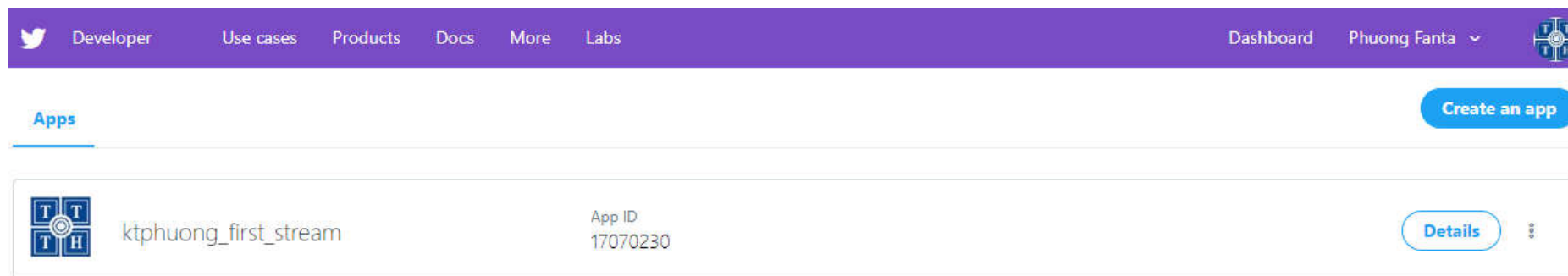




# Làm việc với PySpark Streaming

## □ Kết quả

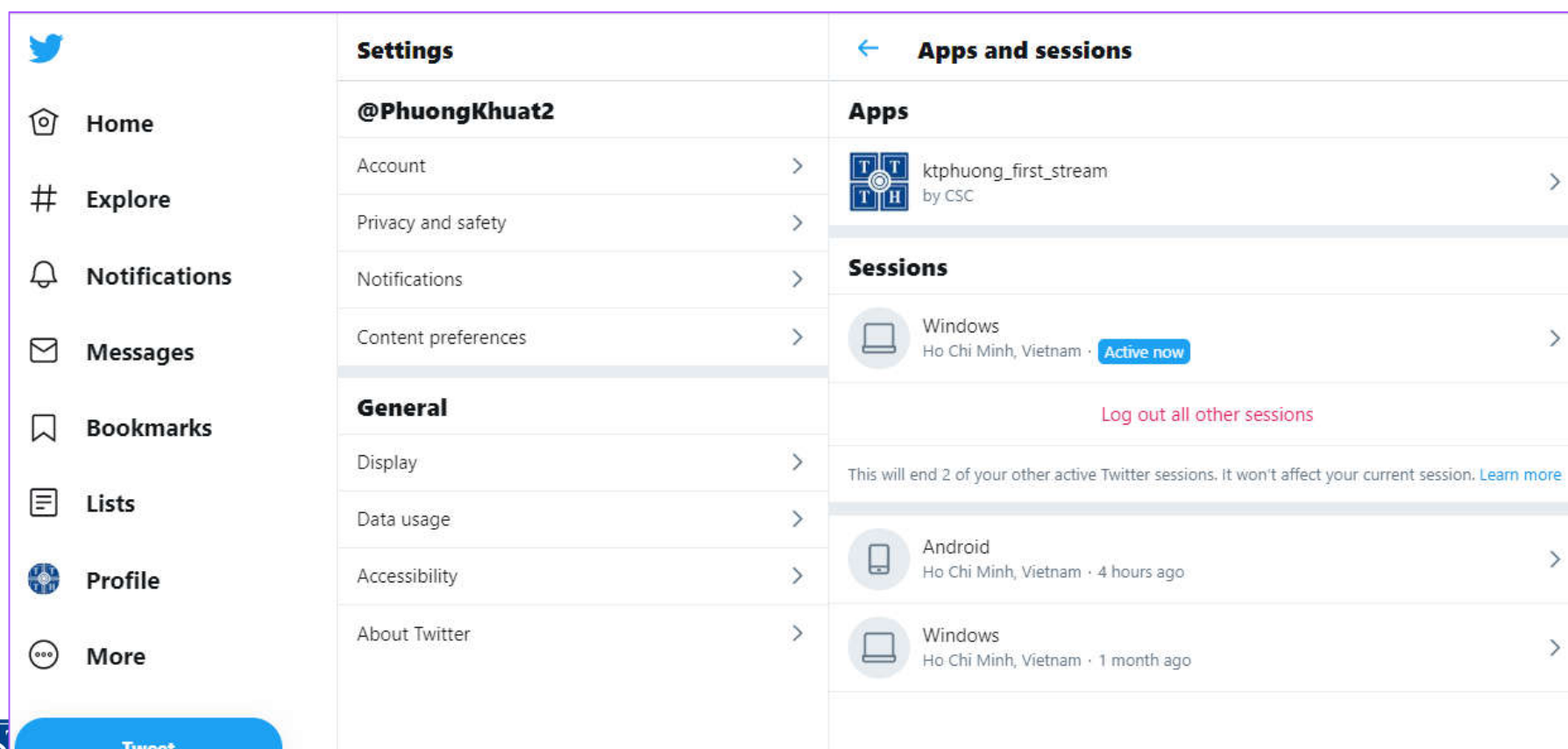
- Khi đăng nhập vào Twitter Developer



# Làm việc với PySpark Streaming

## ❑ Kết quả

- Khi đăng nhập vào Twitter



The screenshot shows the Twitter mobile app interface. On the left is a navigation menu with options: Home, Explore, Notifications, Messages, Bookmarks, Lists, Profile, and More. The main content area is divided into two panels. The left panel shows the 'Settings' for user '@PhuongKhuat2', with categories like Account, Privacy and safety, Notifications, Content preferences, and General. The right panel, titled 'Apps and sessions', shows a list of active sessions. Under the 'Apps' section, there is an entry for 'ktphuong\_first\_stream by CSC'. Under the 'Sessions' section, there are three entries: 'Windows' (Ho Chi Minh, Vietnam, Active now), 'Android' (Ho Chi Minh, Vietnam, 4 hours ago), and 'Windows' (Ho Chi Minh, Vietnam, 1 month ago). A red link 'Log out all other sessions' is visible between the first and second session entries, with a warning message below it: 'This will end 2 of your other active Twitter sessions. It won't affect your current session. Learn more'.



# Làm việc với PySpark Streaming

---

## Demo



