An ornate doorway with a dark door and decorative panels. The door has a semi-circular window with a colorful design. The surrounding walls are highly decorative with intricate carvings and patterns.

Guidelines for API Access to Social Media

This document was prepared by Thomas Treml, Rania Wazir, and Armin Rabitsch as part of the “Social Media Data Access” project, within the framework of the Spotlight on Artificial Intelligence & Freedom of Expression (SAIFE) Project of the Organization for Security and Co-operation in Europe - Office of the Representative on Freedom of the Media (OSCE RFoM).

January 2021

This document is released under a Creative Commons [CC BY-SA 4.0](https://creativecommons.org/licenses/by-sa/4.0/) License

FRONT PAGE IMAGE BY [piper60 from Pixabay](https://pixabay.com/users/piper60/)

API GUIDELINES

Content

1.	Introduction.....	4
2.	Obtaining Twitter API access:.....	5
2.1.	Creating a Twitter Developer Account.....	5
3.	Obtaining Facebook API access:.....	11
3.1.	Creating a Facebook Developer Account:	11
3.2.	Creating an App:.....	12
3.3.	App Review:	13
3.4.	Verifying your Business/Organization:.....	19
3.5.	After Public Pages Access has been granted:	21
3.6.	Obtaining access to the Facebook Ad Library API:	21
3.7.	Access Ad Library using Graph API:.....	22
3.8.	Reports:.....	24
3.9.	CrowdTangle.....	27

1. Introduction

Over the past decade, Social Media platforms have gained prominence as much of our communication and interaction has moved online. This has happened not just in the private, but also the public sphere. Accordingly, in order to understand how societies are communicating, in order to monitor political processes, but also in order to safeguard citizen and consumer rights vis-à-vis online providers, access to Social Media data for researchers, journalists, and CSOs alike has become a necessity.

In order to obtain data from Social Media platforms¹, there are a few options available. When the data volume is low, it is of course possible to manually collect the data (screenshots, tweet URLs, copy/paste, etc.) but this is an approach that does not work well for larger amounts of data. In such cases, the data either needs to be scraped², or API³ access needs to be obtained directly from the platform providers. Scraping data does not entail the tiresome process of requesting API access from the provider. But this approach lives in a grey zone of legality⁴, and requires much greater technical skills, because the platform providers often change their web layouts and put in other obstacles in order to *actively* prevent scraping from their websites. Obtaining API access to a social media platform is akin to knocking at the door, and being allowed to enter a huge villa as a guest: it may be difficult to gain admittance, and might come together with restrictive house rules - but it includes a guided tour of the premises, and what they have to offer. Scraping is more like sneaking in through the window.

In the sections below, we describe the procedures for obtaining access to the APIs of Twitter and Facebook. We focus on these two platforms because they are currently the largest source of textual content, which has been the focus of our own research. Google does provide API access to YouTube meta-content, but this is outside the scope of this document. Other social media platforms, such as TikTok and Instagram, do not have API access that permits the downloading of data from public pages. Such platforms instead require having page administrator rights in order to download data from a given page, or else they require that an application⁵ be created that platform users will want to add to their accounts (such as a gaming, or a lifestyle app), and then such applications can obtain access to the data that the user consents to. While user apps may turn out to be a viable option for research on filter bubbles and micro-targeting, they suffer from selection bias (data is only obtained on such people who decide to install the app), and do not lend themselves to wide-scale research, particularly in the context of monitoring political campaigns, and observing human rights in the public arena.

We therefore restrict our attention here to API access required in order to obtain data from public pages/public accounts - i.e., from pages and content that are visible to anyone on the internet, without having to be a follower or friend of that particular account. While this is the default mode on the Twitter API, on Facebook such access is available only through the Pages Public Content Access API.

A pre-requisite for requesting API access on both Twitter and Facebook is the possession of a regular Twitter/Facebook user account. If this is not the case, a user account will need to be created prior to making the request for API access.

1 Twitter, Facebook, Instagram, YouTube, TikTok are examples of social media platforms. We will also often refer to the companies that run these platforms as “providers”

2 scraping is the process of collecting data from a website using an automated script (i.e. program). In order for the scraped data to be useful, it is necessary for the website layout to be well understood, so that the scraping program can differentiate between actual website content, and formatting elements

3 API means Application Programming Interface. It is a set of documented rules that permits automated access to the features or data of a particular program.

4 Consider the fate of NYUs Ad Observer - adobserver.org; which was recently hit with a Cease and Desist Order by Facebook: <https://www.wsj.com/articles/facebook-seeks-shutdown-of-nyu-research-project-into-political-ad-targeting-11603488533>

5 Often referred to as a “user app”

Also, a note of warning: the APIs are not stable - this means, they are subject to frequent change. In drawing up this document, the authors discovered that the procedures for requesting API access from Twitter have changed (we describe the updated procedure below, as of 18.12.2020). However, many changes will also impact the API itself (how to access it, what can or cannot be downloaded, how much can be downloaded over a given period of time, which commands to use, and the type of authentication required, are just some examples). This means that, once API access is obtained, much work needs to be invested in maintaining the code for accessing the data, and keeping it up-to-date with the latest changes.

The data4good GitHub [repository](#) contains examples of code for downloading data from Twitter and Facebook. This code was functional until July 2020. It may need some adaptation in order to work with recent API changes in both Twitter and Facebook. While we continue to work with Social Media data, we will do our best to keep the code up to date with regular frequency, or to indicate where modifications are necessary.

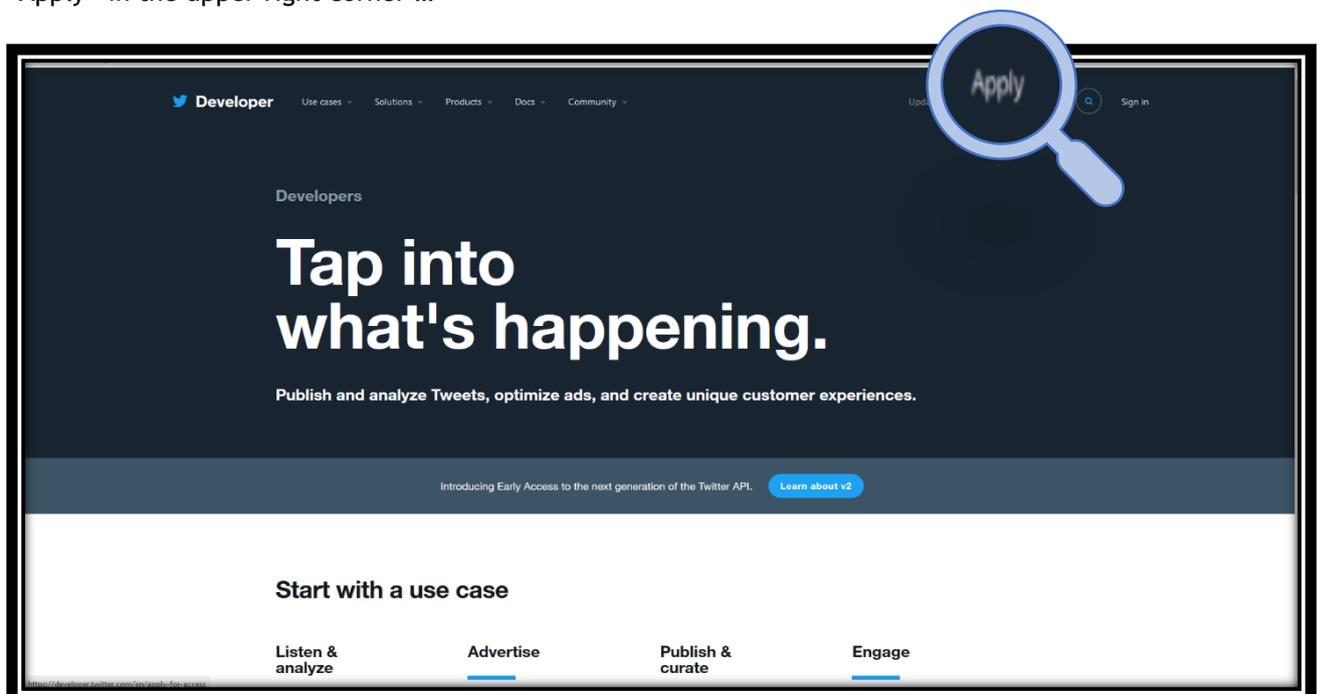
2. Obtaining Twitter API access:

In this section, we describe the steps necessary to obtain Twitter API access. This will give access to the data (such as posts, comments, replies, likes, shares) from public Twitter accounts (this is the default on Twitter) - that is, pages that can be seen online without any kind of connection to the page owner/account holder.

This requires creating a Twitter Developer Account⁶, if one does not already exist; the Twitter Developer Account creation entails an application process, because the developer account is automatically granted Twitter API access.

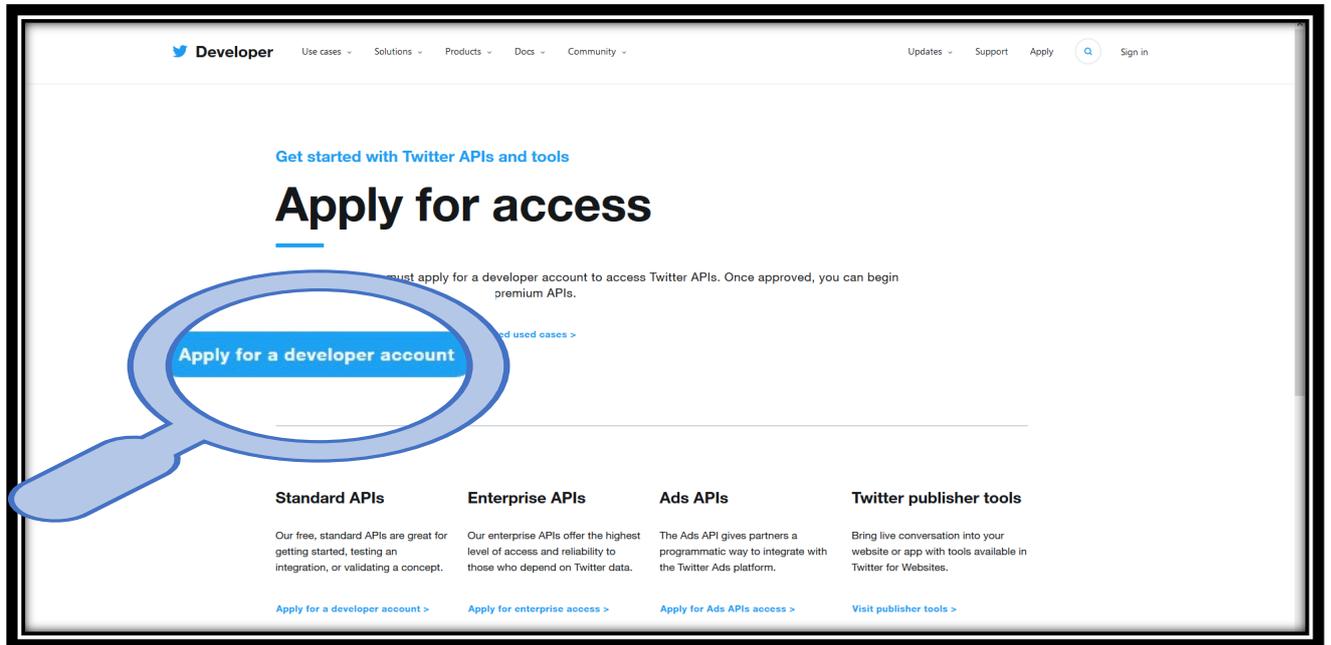
2.1. Creating a Twitter Developer Account

In order to set up a Twitter Developer Account, go to <https://developer.twitter.com/en>, select “Apply” in the upper right corner ...

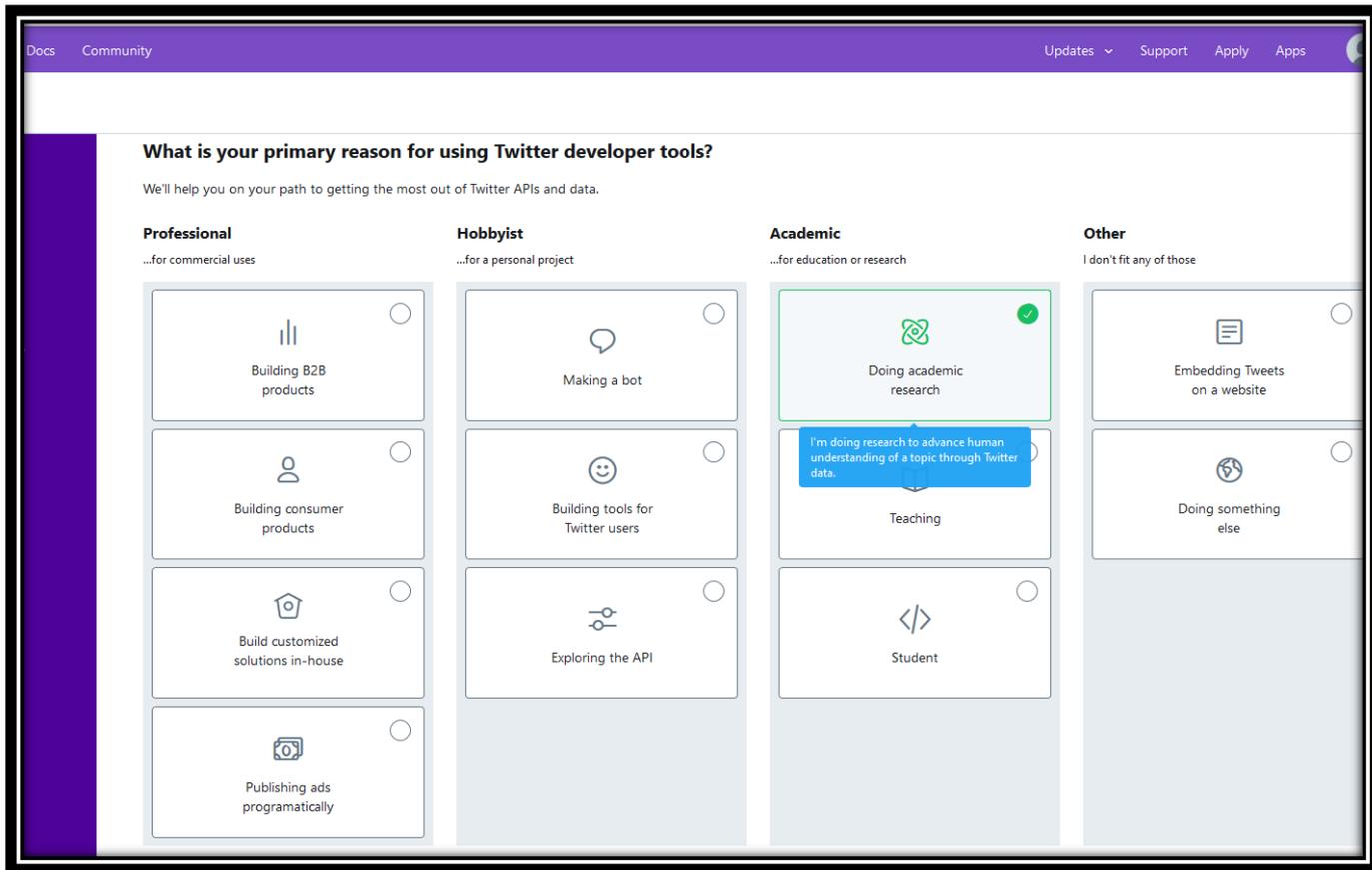


⁶ A Twitter developer account is additional to a regular Twitter account, and allows you to create Twitter apps (programs), and write code that interacts with the Twitter platform and content

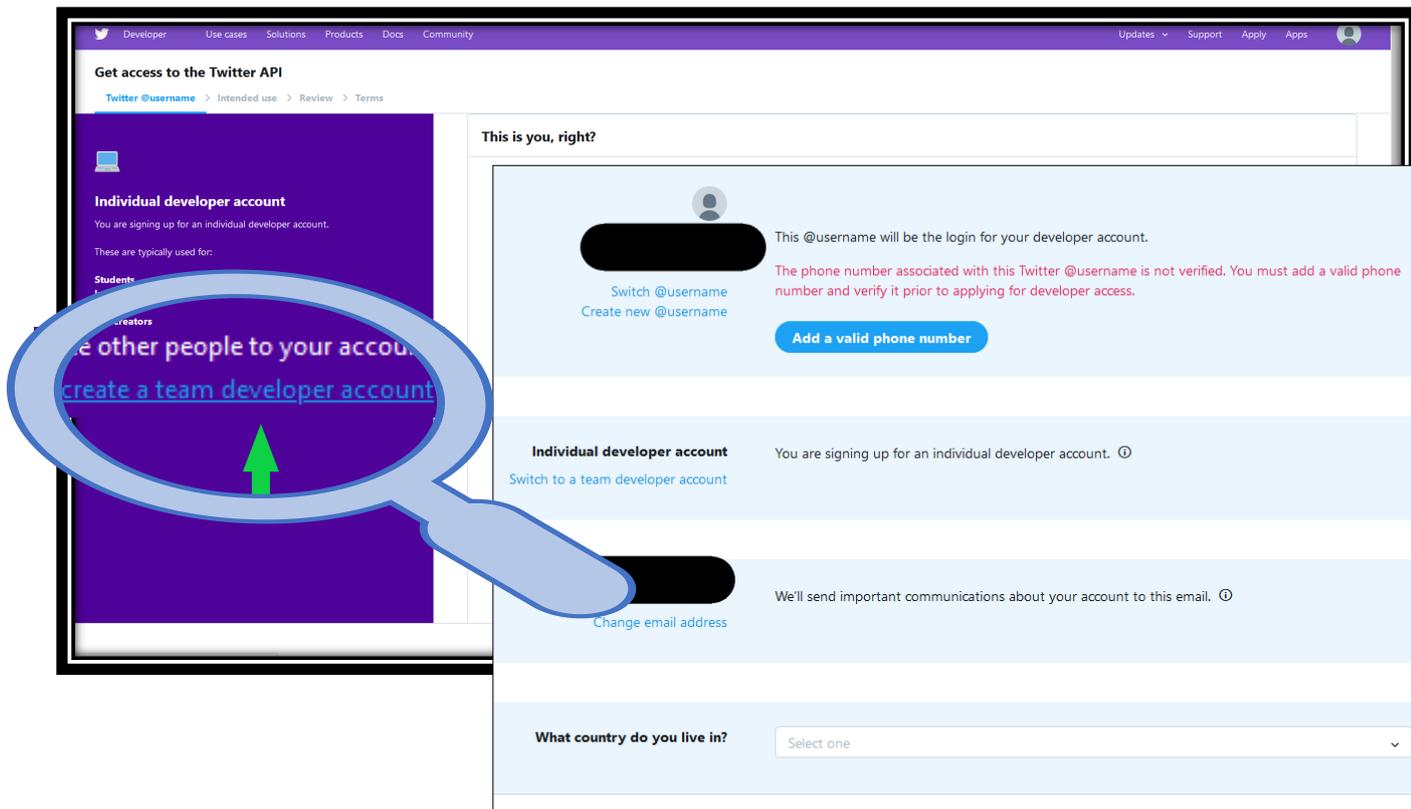
... and then click on “Apply for a developer account”.



Twitter will ask a few questions, such as the main reason for getting a developer account:

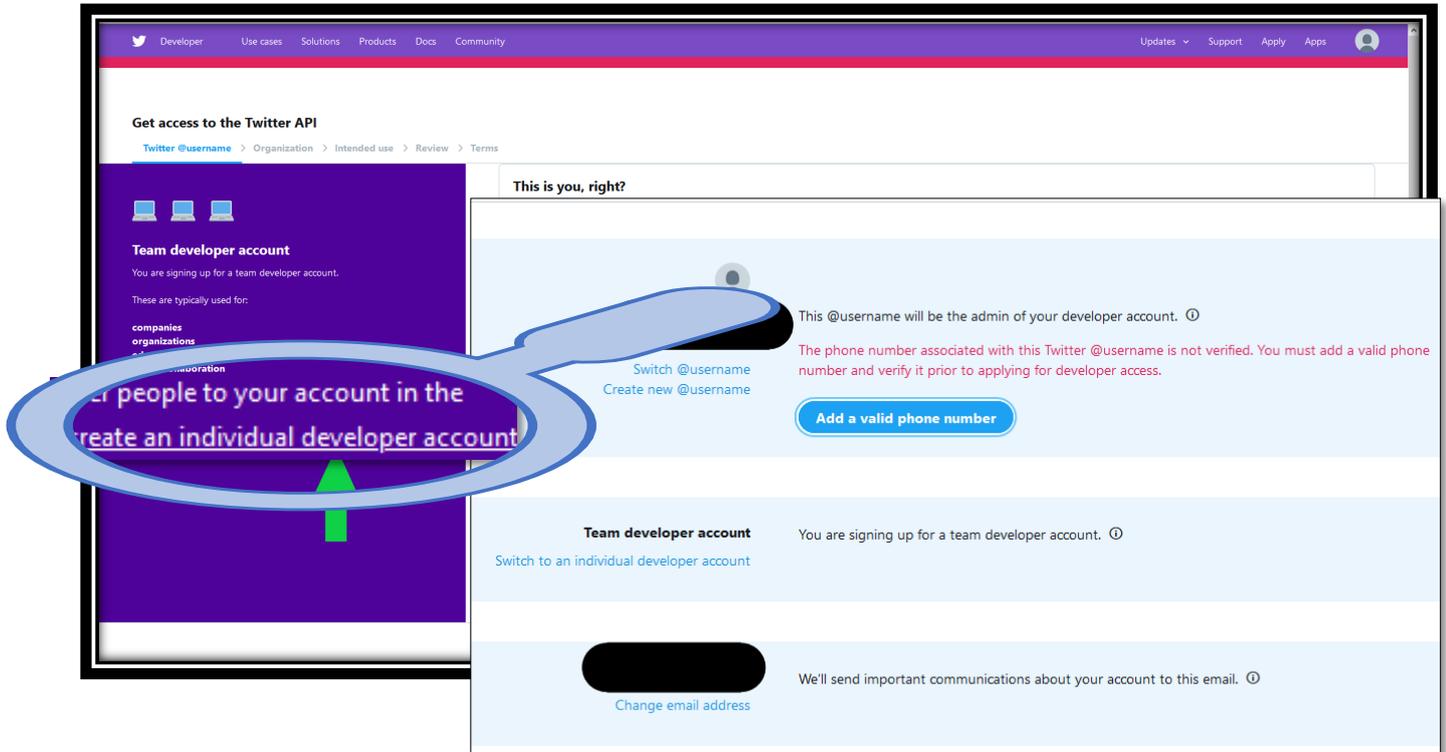


Should the account be an individual developer account (i.e. student, individual researcher, maker, etc.) ...



API Guidelines

... or a team developer account (for example, an enterprise or organization, or a research group, that will need to share access to the account).



For the team account, a few details about the organization must be shared with Twitter:

The screenshot shows the organization details form for a team developer account. The form includes the following fields and options:

- Team name**: This will be the name of your account. **Required**
- Legal entity name**: Of company, institution or parent organization. **Required**
- Organization Twitter @username**: @ **Required**
- Website URL (optional)**: www
- Organization primary country of operation**: Select one **Required**
- How do you categorize your organization?**: Nonprofit
- What industries do you / will you serve?**: Select all that apply. Nonprofit Academic Add...

Afterwards, describe what the API access will be used for, what kind of data will be needed, and how the data will be analyzed.

The screenshot displays the Twitter Developer application review interface. On the left, a purple sidebar titled "Get access to the Twitter API" contains sections: "Key things to keep in mind", "Restricted uses", "Automation", and "Be thorough". The main content area is titled "How will you use the Twitter API or Twitter data?". It includes a text input field under "In your words" with a character count of 200. Below this is the "The specifics" section, which contains a toggle switch for "Are you planning to analyze Twitter data?" (set to "Yes") and a prompt: "Please describe how you will analyze Twitter data including any analysis of Tweets or Twitter users."

Once the description is complete, there will be an opportunity to review the application, and Twitter's Terms of Service must be accepted before proceeding. These include the following:

- If you didn't do so when setting up your Twitter account, now is the time to check out the Twitter Rules: <https://help.twitter.com/en/rules-and-policies/twitter-rules>
- Developer Agreement and Policy: <https://developer.twitter.com/en/developer-terms/agreement-and-policy>
- Restricted Use (note, this contains a ban on anything that could lead to government surveillance): <https://developer.twitter.com/en/developer-terms/more-on-restricted-use-cases>
- Display Requirements: <https://developer.twitter.com/en/developer-terms/display-requirements>
- Automation Rules (note, automated scripting - i.e., scraping - could lead to suspension of your account): <https://help.twitter.com/en/rules-and-policies/twitter-automation>

Finally, the application can be submitted for review. In the authors' experience, a response *normally* arrives in less than 48 hours.

Once the Twitter API access has been obtained, some time must be spent gaining familiarity with the [Twitter documentation](#), which explains how to make Twitter searches and retrieve data.

In the older version of the Twitter API (v. 1.1), the so-called “firehose” API was available for free. BUT, in [Twitter’s own words](#):

Before digging in, it’s important to know that the standard search API is focused on relevance and not completeness. This means that some Tweets and users may be missing from search results. If you want to match for completeness you should consider the [premium](#) or [enterprise](#) search APIs.

It is unclear how much of the data might be missing, and what *exactly* are the criteria for selection, if retrieved through the standard search API. This has implications for data quality, as it is not possible to determine if there was some selection criterion for the missing tweets, or if these were random⁷. As indicated above, the premium API does not suffer from this disadvantage, but larger volumes of data must be purchased, and quickly become prohibitively expensive. No corresponding warnings have been found for the new Twitter API (v. 2). Additionally, Twitter has recently announced unrestricted API access for academic (but not yet CSO or journalist) researchers⁸.

A great open source library for working with Twitter data in the programming language R is [rTweet](#) - as of this writing, it is also maintained and kept up to date by the authors.

Data obtained from the Twitter API will include information such as user name and ID connected to all tweets downloaded. In the interest of preserving privacy (and GDPR compliance), we strongly advise that as a first processing step, user names be removed. If not necessary for your research, remove also the user ID - otherwise, hash the ID, or use some other strong form of pseudonymization. In order to comply with GDPR requirements, keep the data stored securely, with restricted access, on servers located within the EU and run by EU-based firms⁹. Several cloud service providers have popped up within the EU that provide this possibility. We are aware of, for example, <https://www.fairkom.eu/> or <https://thegood.cloud/en>, but there are many more.

⁷ To make the problem more tangible: if the objective of the research is to determine how much hate speech is produced on twitter over a given period of time, and the firehose API, instead of randomly selecting the tweets to leave out, actively filters out possibly offensive tweets, then the sample collected from the firehose API will under-represent the percentage of hate speech comments. Since, however, the researcher is unaware of the criteria used for filtering out tweets, they cannot correct for this factor.

⁸ https://www.theverge.com/2021/1/26/22250203/twitter-academic-research-public-tweet-archive-free-access?utm_campaign=theverge&utm_content=entry&utm_medium=social&utm_source=linkedin

⁹ pursuant to the ECJ ruling on the US-EU Privacy Shield, data stored in the US is no longer considered secure, and some additional contractual agreements with the storage provider would be needed

3. Obtaining Facebook API access:

In this section, we describe the steps necessary to obtain Facebook Public Pages API access. As mentioned before, this will give access to the data (such as posts, comments, replies, likes, shares) from public pages - that is, pages that you can see online without any kind of connection to the page owner.

This requires three steps:

1. Creating a Facebook Developer Account¹⁰, if you do not already have one;
2. Creating an app¹¹, and applying for the Public Pages Access permission
3. Verifying your business or organization

Once you have created an app, verifying your business or requesting public pages access can be done in any order, or even concurrently (the preferred mode, as both procedures can take a long time - while officially the review is said to take 5 working days, you should be prepared to have the application rejected; in our experience, the first time Pages Public Content Access is requested, it can take 1-2 months to obtain a positive review, if at all).

3.1. Creating a Facebook Developer Account:

Go to <https://developers.facebook.com/> In the top right corner, click on “Log In” or “Get Started” (be aware that not everyone gets the same screen, and the layout changes frequently!)

You will then have to log in to your regular Facebook account, before being led to the next step, that asks you to verify yourself by providing a telephone number to which a verification code can be sent (or else eventually a credit card). After the verification, you will be asked to tell them what category of user you belong to:

- Developer
- Marketer
- Analyst
- Product Manager
- Student
- Owner/founder
- Other

That completes the Developer Account set up - you will next be led to create your first app.

Before proceeding though, make sure to read through the various Terms and Conditions:

- Facebook Platform Terms: <https://developers.facebook.com/terms>
- Developer Policies: <https://developers.facebook.com/devpolicy/>

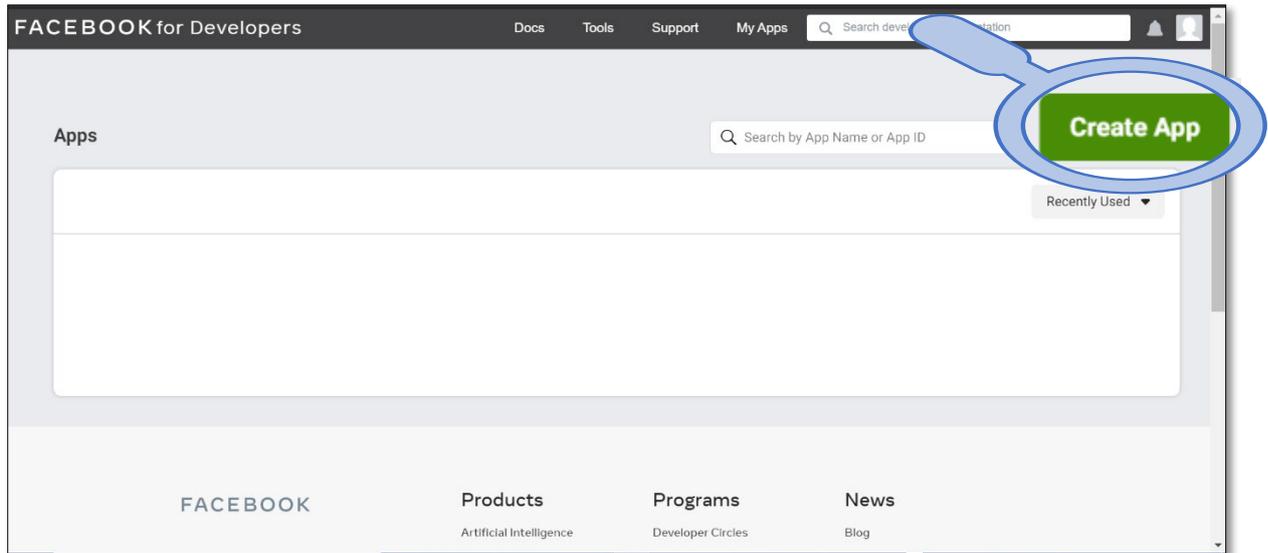
Each of those pages contains links to other pages, so be prepared to spend some time going through the various policies.

¹⁰ A Facebook developer account is additional to a regular Facebook account, and allows you to create Facebook apps (programs), and write code that interacts with the Facebook platform and content

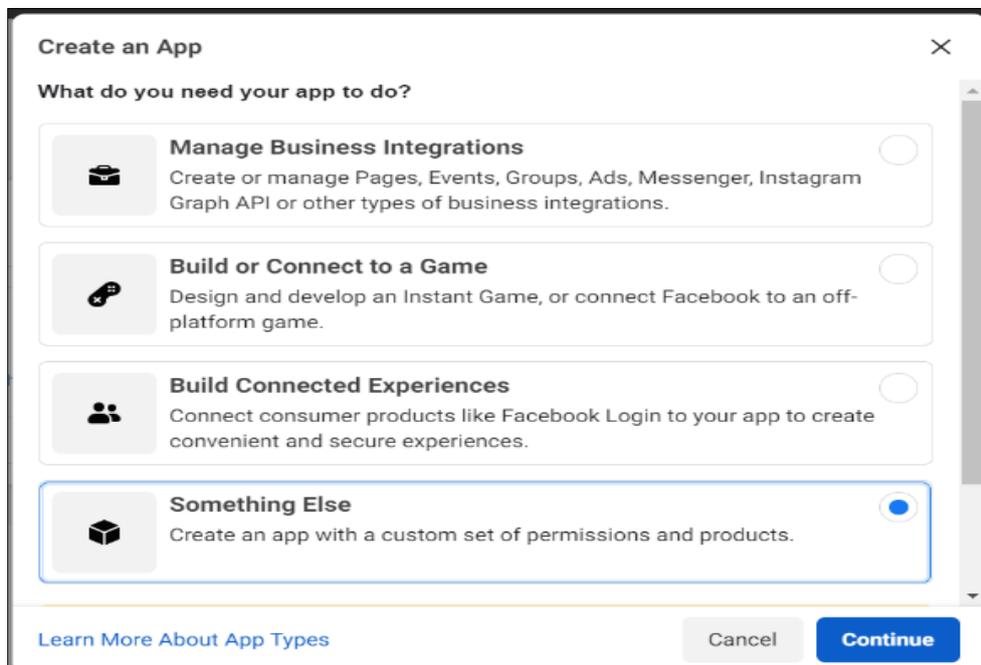
¹¹ An app is just a program - if you have a Facebook account, you’ve probably seen games or lifestyle apps

3.2. Creating an App:

Click on the “Create App” button in the upper right side of the screen.



In a few steps, describe what the app is for (for readers of this document, the answer is likely “something else”), give it a name (for the purpose of this guide, we called ours “Trial2”), and provide a contact email. If you/your organization already have a Business Manager account, you can connect the app to it already. If not, the instructions for doing so are below.



After clicking on “Create App” in the bottom right corner, the Facebook app is created. But right now, it is just a place holder, and does nothing.

The screenshot shows a mobile-style form titled "Create an App" with a close button (X) in the top right corner. The form contains the following sections:

- App Display Name:** A text input field with a placeholder character "|". Below it is the text: "This is the app name associated with your app ID."
- App Contact Email:** A text input field. Below it is the text: "This email address is used to contact you about potential policy violations, app restrictions or steps to recover the app if it's been deleted or compromised."
- Do you have a Business Manager account? - Optional:** A dropdown menu currently showing "No Business Manager Account selected". Below it is the text: "In order to access certain aspects of the Facebook platform, apps may need to be connected to a verified Business Manager account. If you haven't yet set up an account, you can create one now or later in the process."

At the bottom of the form, there is a line of text: "By proceeding, you agree to the [Facebook Platform Terms](#) and [Developer Policies](#)". Below this are two buttons: a "Back" button with a left arrow and a blue "Create App" button.

In order for the app to be useful, it must have certain access rights and properties. If, for example, the app needs to be able to download data from Facebook’s public pages, it will need to have the so-called “Page Public Content access” - and this can only be obtained via an application process that will be described below. But before proceeding, it is important to understand the two types of apps that Facebook offers.

The first, and most supported kind, is the “user app.” This is the kind of app that users add to their Facebook accounts (for example games, or lifestyle apps) and which are supposed to enhance the user experience on Facebook. Such apps only interact with the data of users that have added them to their accounts. This is the kind of app necessary in order to capture individual users’ experiences on Facebook - for example, to try to determine what kind of content is delivered to them in their feeds. However, this kind of data collection is invasive, and depends on the users’ consent.

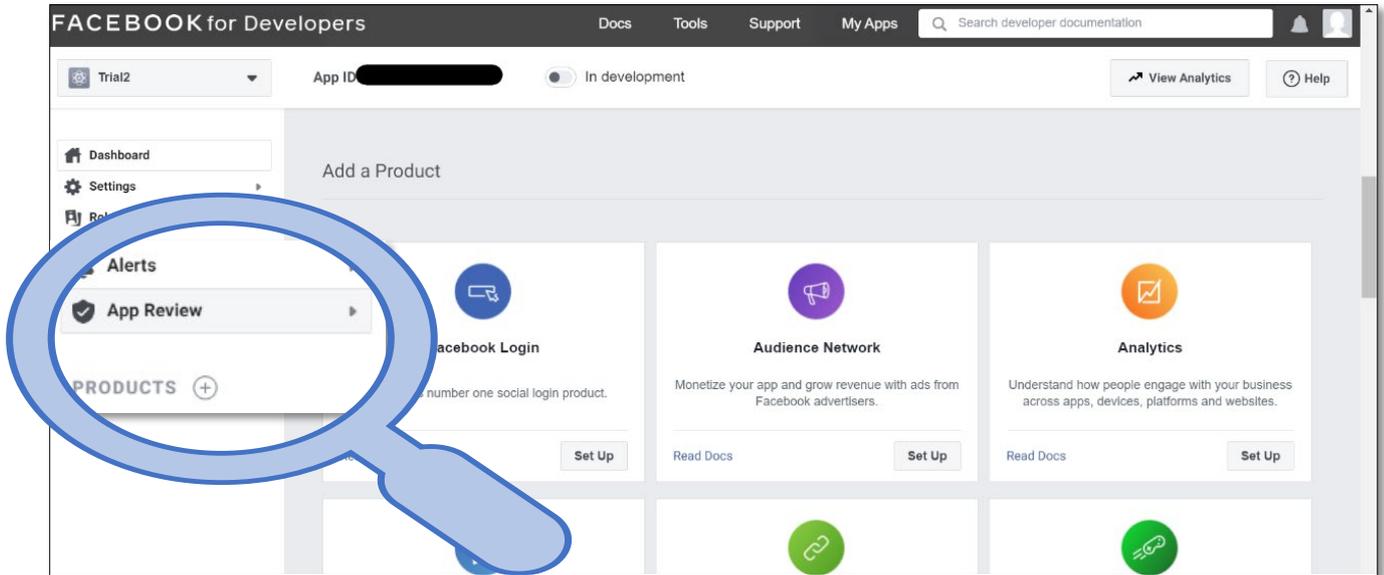
The second kind of app is a server-to-server app. Such an app is intended for exchanging data directly with Facebook - without any Facebook user interaction. This is the kind of app most likely required in order to conduct research on Facebook public pages content, and the steps below are intended to explain how to request Page Public Content access for server-to-server apps.

3.3. App Review:

In order to obtain the “Page Public Content Access” permission, the app must go through the Facebook App Review procedure. This review process has been designed by Facebook with *user apps* in mind - it can be rather trying to go through this procedure as an NGO with a server-to-server app, and is one of the main motivations for providing this guidance. (Facebook does have some instructions on how to go through the App Review with a server-to-server app - they can be found [here](#)).

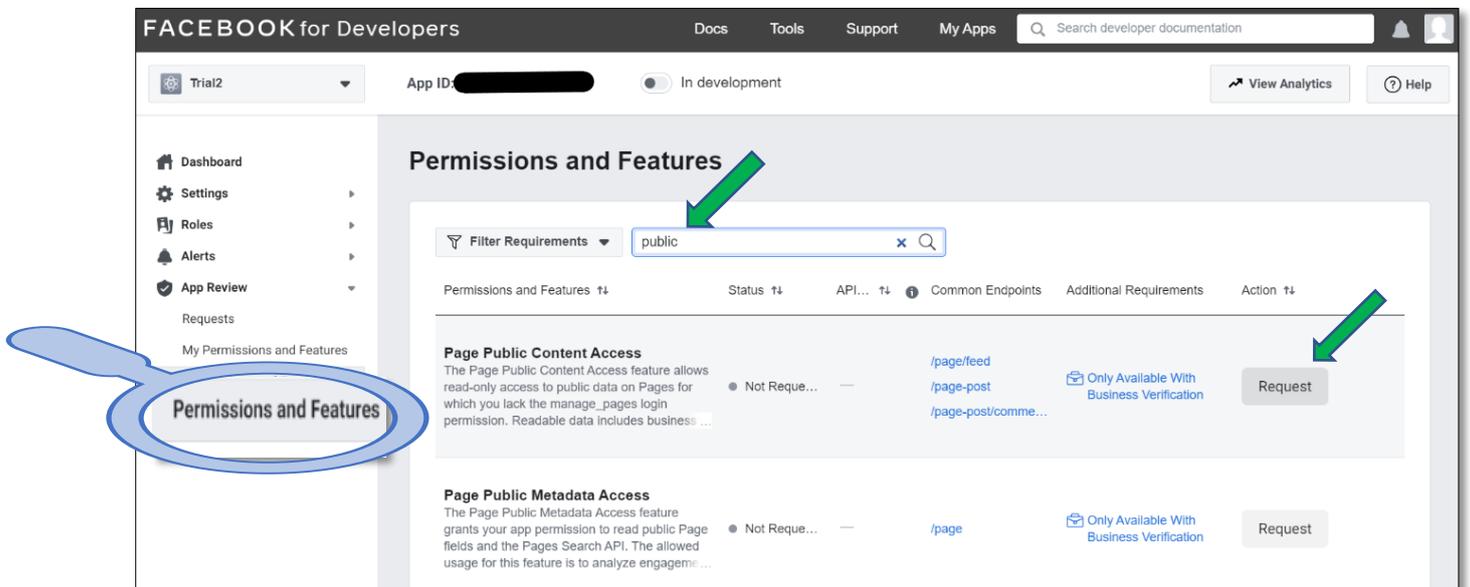
API Guidelines

The “Add a Product” page may not be the first page that appears upon creation of an app - but this is irrelevant¹², as the important part for the app review process is in the menu on the left - the “App Review” menu item.



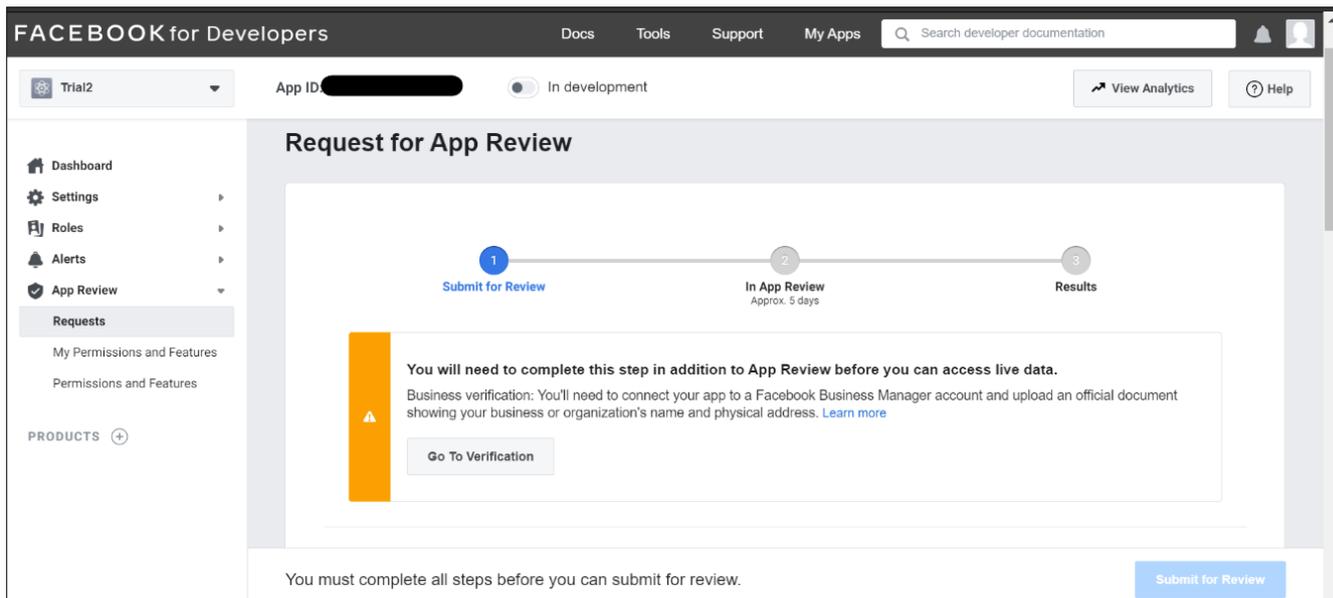
Select “Permissions and Features” from the drop-down menu. Since there are many Permissions and Features to choose from in the page that appears on the right, save some time by searching for the word “public”.

In the field “Page Public Content Access”, click the button “Request”.

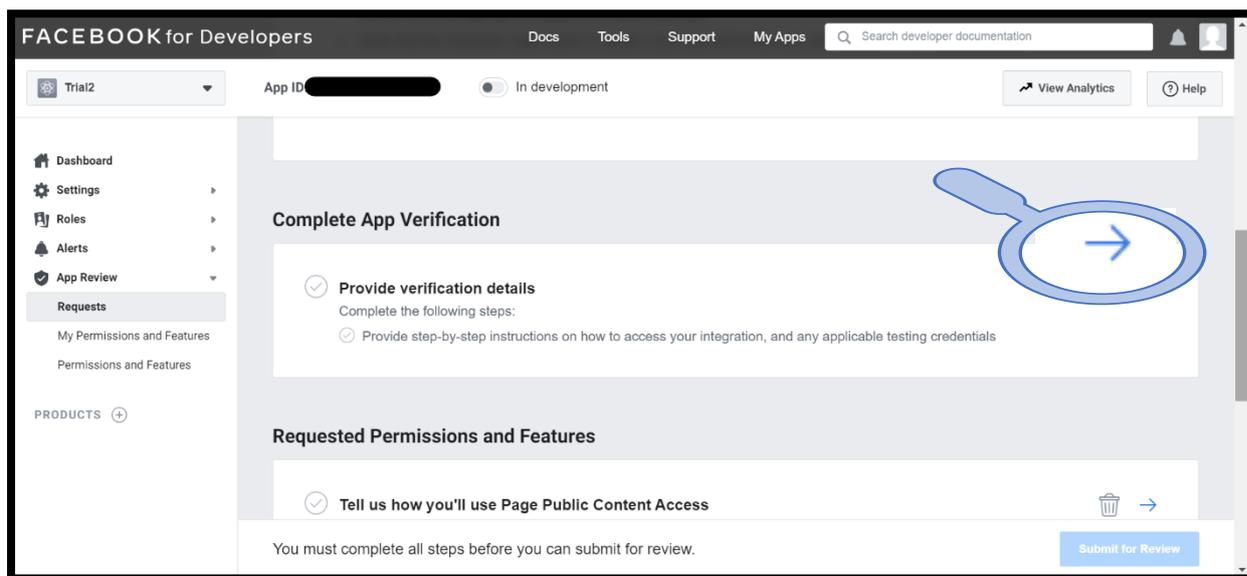


At the top of the page that appears now is the Business Verification. Skip this step for now and scroll down to the next item.

¹² However, for the sake of completeness: should the “Add a Product” page be required - look in the menu on the left, and click on the last item - “PRODUCTS”.



Under “Complete App Verification”, a description of how Facebook’s review team can log in to the app is requested. Click on the arrow on the right.



In the pop-up window, click on “No”: There is no authentication solution that the Facebook review team can use to log in to the app, because the app does not have a user interface - it is a server-to-server app.

In the field that asks for an explanation of why the reviewer isn’t able to log in to the app, make sure to state very clearly that the app is a server-to-server app. Give also a detailed description of what kind of analyses will be done with the data, and why the public pages access is necessary (this description can be re-used in the “Tell us how you’ll use Page Public Content Access” window below).

API Guidelines

Does your app use a Facebook authentication solution, like Facebook Login, that the review team can use to login and verify your app?

Yes

No

Some apps don't have a clear user interface, or use Facebook Login, because they are [server-to-server](#) apps and exchange data directly with Facebook APIs. Please select the option(s) that best describe why a Facebook reviewer can't access or install your app.

My app doesn't use the Facebook Login dialog.

My app posts my own articles, or blog posts, from my website to my own Facebook Page.

My app or integration sits behind a local intranet.

My app transfers leads from my company's ad account to a database or CRM.

Other

Describe why a reviewer isn't able to log in to your app to see how you are using the requested permissions or features.
[View Best Practices.](#)

The app is a server-to-server app.

Take screen shots of all the information you are providing in the pop-ups. They may come in handy when having to respond to the line-up of reviewers who will likely reject the request, or will ask for information that was already provided in a previous iteration.

When done, click on “Save” in the bottom right corner. The pop-up will disappear, and the App Review page will re-appear. Scroll down to the “Requested Permissions and Features” item, and click on the little arrow on the right to complete this section.

FACEBOOK for Developers

Docs Tools Support My Apps Search developer documentation

Trial2 App ID [redacted] In development View Analytics Help

Dashboard Settings Roles Alerts App Review Requests My Permissions and Features Permissions and Features PRODUCTS

Requested Permissions and Features

Tell us how you'll use Page Public Content Access 

Review the policies for Page Public Content Access and tell us how you intend to use it

- Describe how your app uses this permission or feature
- Upload screencast showing the end-to-end user experience
- Agree that you will comply with allowed usage

[+ Add additional permissions and features](#)

You must complete all steps before you can submit for review.

This pop-up requests information that is quite difficult to provide when you are making a server-to-server app. Therefore, you can reuse the description from before, making sure to state prominently that this is a server-to-server app. Describe what analyses you plan to run on the collected data. Note that, when you collect data from public pages using the Page Public Content Access, Facebook does not, in general, provide user or user ID of users who post, comment or reply on the page, unless they are the page owner. (This point bears some clarification: even if, on the Facebook Page itself, the user names of users who comment are clearly visible, this information is *not* delivered via API.) However, it is good to reiterate here how you will protect the data and the privacy of users. Use this section to indicate also that the screencast¹³ you will be providing below is based on simulated (that means synthetic, generated only for the purpose of demonstration) data.

The next item requests a screencast, describing how the public pages data will be used. Go ahead and view Facebook's [sample server-to-server submission](#). On the [server-to-server documentation page](#), Facebook does now (this passage did not exist in 2019) say *Unlike other app types, screencasts are not required for App Review*.

However, you *may* find that the system will not accept your application as complete unless you submit a screencast anyway. In that case, go ahead and submit a screencast as described in the next paragraph¹⁴, and refer reviewers who dislike the quality of your screencast to the server-to-server documentation page.

The Facebook sample may be illuminating for a *business* requesting the public pages access. If this is not your situation, your best strategy is to generate a simple visualization based on simulated data, of the kinds of aggregate analyses that you intend to perform on the data collected from the public pages. In the description field above, you should have made sure to indicate that the screencast is based on simulated data, because until the app passes the App Review, you cannot obtain any live Page Public Content data (some reviewers seem to be unaware of this).

Tell Us Why You're Requesting Page Public Content Access ✕

The Page Public Content Access feature allows read-only access to public data on Pages for which you lack the `manage_pages` login permission. Readable data includes business metadata, public comments and posts. The allowed usage for this feature is to provide aggregated, anonymized public content for competitive analysis and benchmarking. You may also use this feature to request analytics insights to improve your app and for marketing or advertising purposes, through the use of aggregated and de-identified or anonymized information (provided such data cannot be re-identified).

Please provide a detailed description of how your app uses the permission or feature requested, how it adds value for a person using your app, and why it's necessary for app functionality. ^[?]

Provide a screencast that shows how data exchanged with our APIs is ultimately used by your app. For example, if your app pulls data from our APIs and feeds it into a CRM, show us how the CRM uses that data. Read our documentation on how to describe data usage and create a screencast for your server-to-server app. [View A Sample Submission](#)

Drag and Drop Your File

📁 Upload File

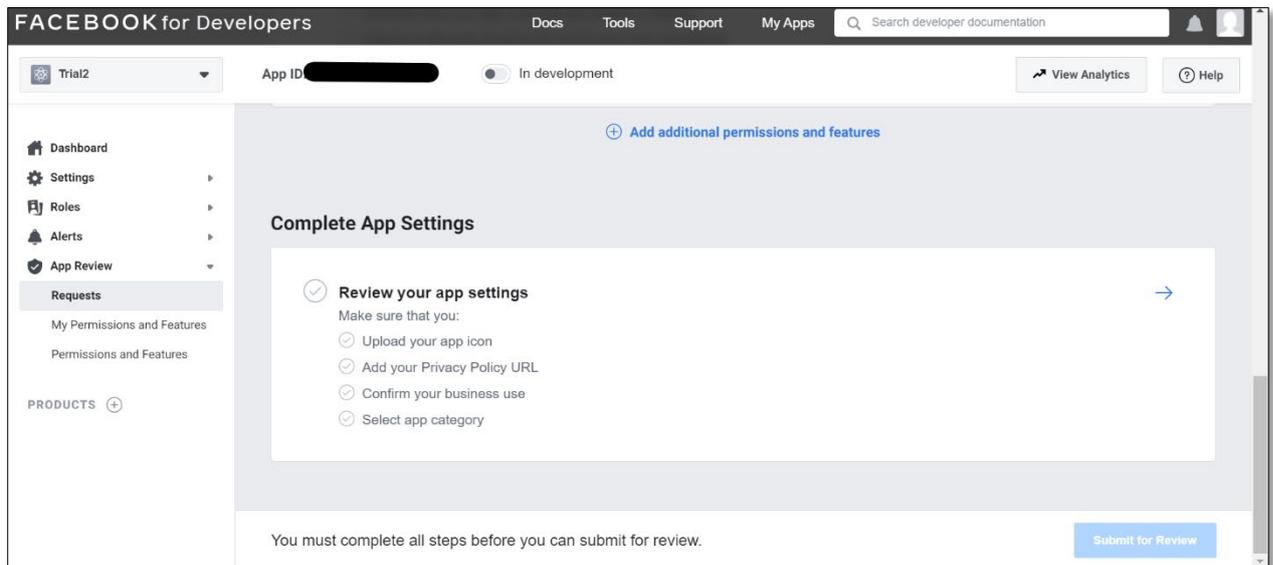
If approved, I agree that any data I receive through **Page Public Content Access** will be used in accordance with the allowed usage.

¹³ A screencast is a video version of a screenshot - it shows a video of your computer screen

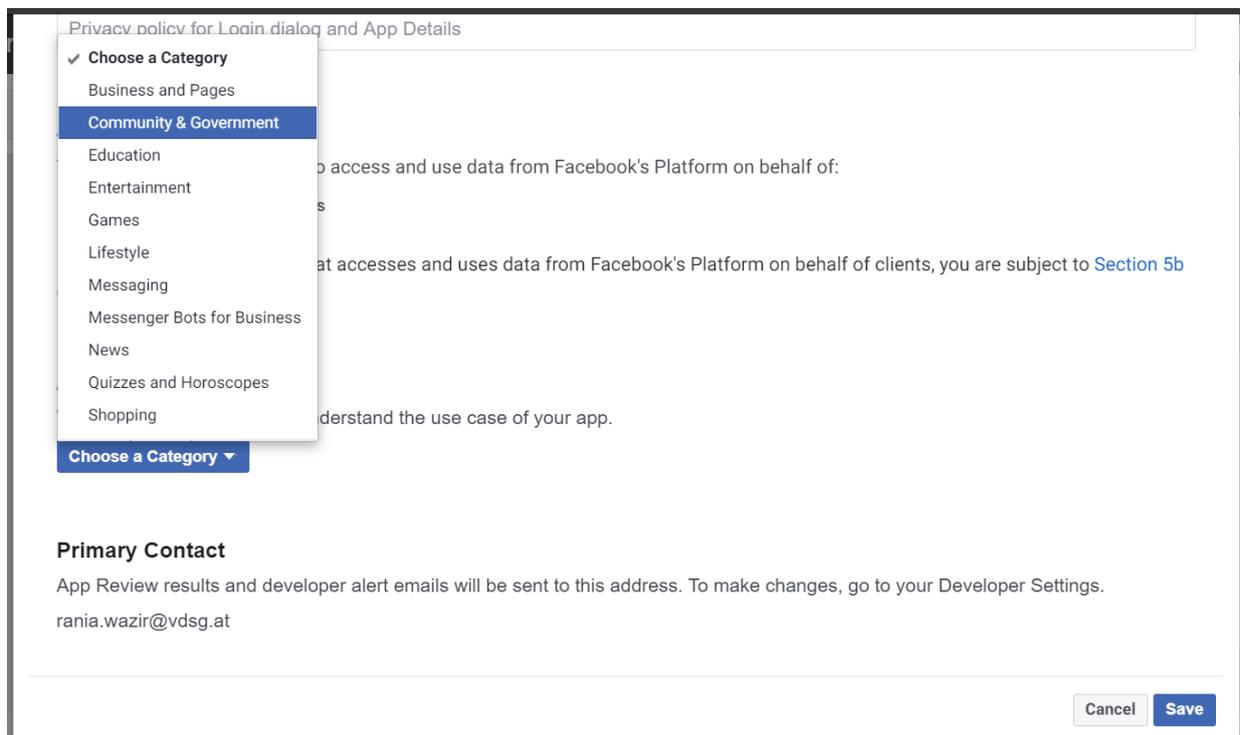
¹⁴ It might also work to submit a screencast of the Facebook [website](#) saying server-to-server apps do not need to submit a screencast - but as this particular piece of Facebook documentation is new, we have not had the opportunity to test this procedure.

API Guidelines

When you're done (don't forget the screenshots!) and have saved your changes, you will return to the main app review page, and scroll down to the bottom of the page, where you can complete the field "Complete App Settings".



Here you will need to upload your organization's logo as the app logo, and your organization's Privacy Policy URL, confirm whether you are developing the app on behalf of yourself or an organization, and finally, choose the app category: if you're an NGO or developing on behalf of an NGO, your choice would most likely be Community & Government.



After you save, you are done, and on the main app review page, you can click the "Submit for Review" button in the bottom right corner. At the top of the App Review page, there is a timeline that indicates it takes around 5 working days to go through the review and get a response. However, this has never been our experience. The app review takes longer (based on experience, the estimate is about 1-2 months), and the first submission usually comes back with a rejection.

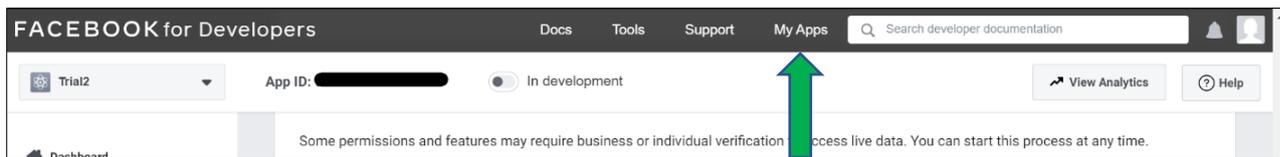
Although it may not be immediately obvious, the rejection can be appealed. Through the Business Account, it is possible to contact, via messaging, Facebook app review support. Be patient and persistent, document all your interactions, and if at all possible, try to find a contact at Facebook that you can actually talk to on the phone, and who can accompany your application (note that no telephone numbers or contact persons are provided on the Facebook page, so this usually has to be a separate initiative).

3.4. Verifying your Business/Organization:

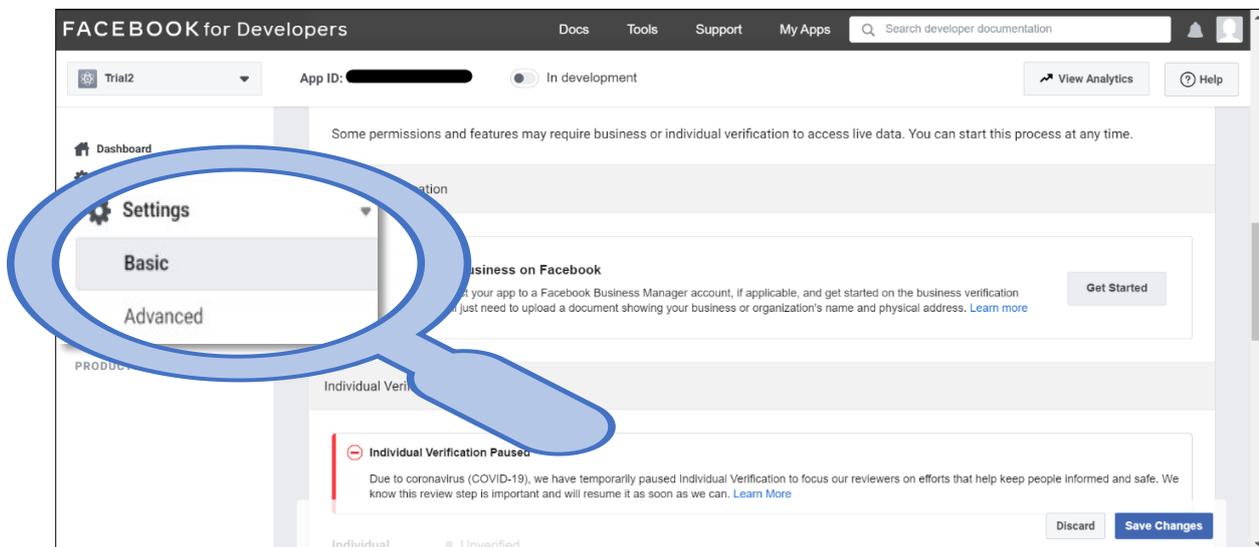
You can check out what Facebook has to say about Business Verification here: <https://developers.facebook.com/docs/apps/business-verification/>

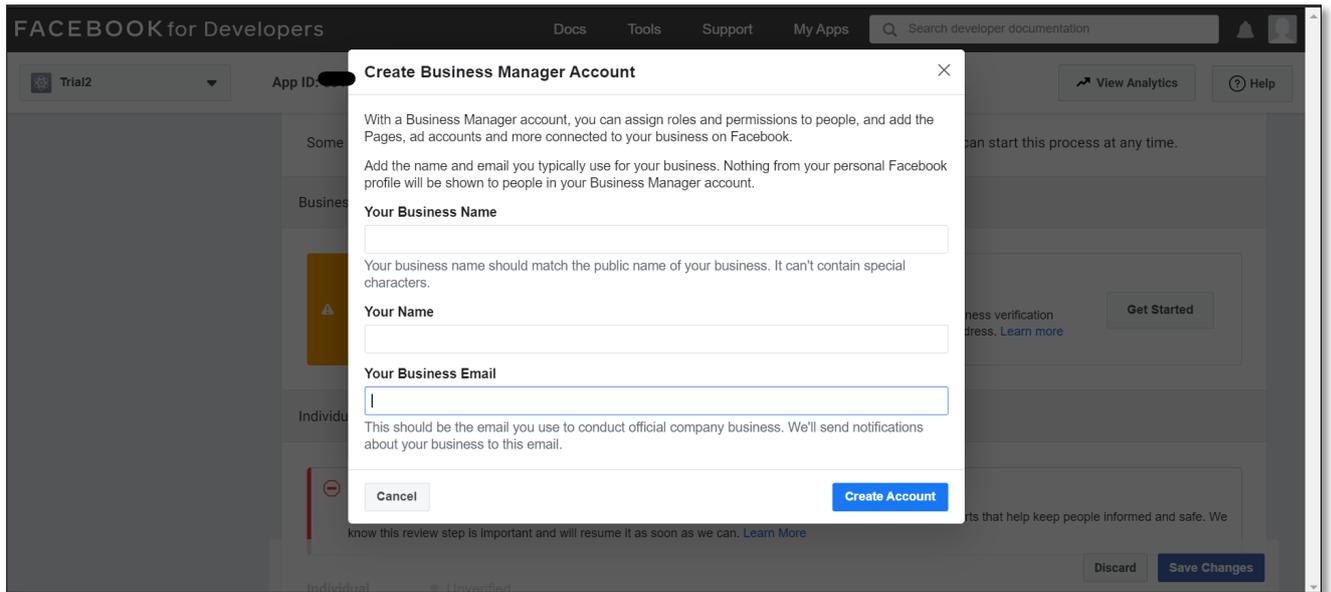
Please note that, as of this writing, regardless of what is written in the Facebook documentation, if you want to have Public Pages API access to real (Facebook calls it “live”) data, you *must* complete Business Verification in addition to the App Review.

Go to the front page of your newly created app (Trial2 in our test case). Do this by clicking on “My Apps” in the top tool bar, and selecting the app from the pop-up menu.

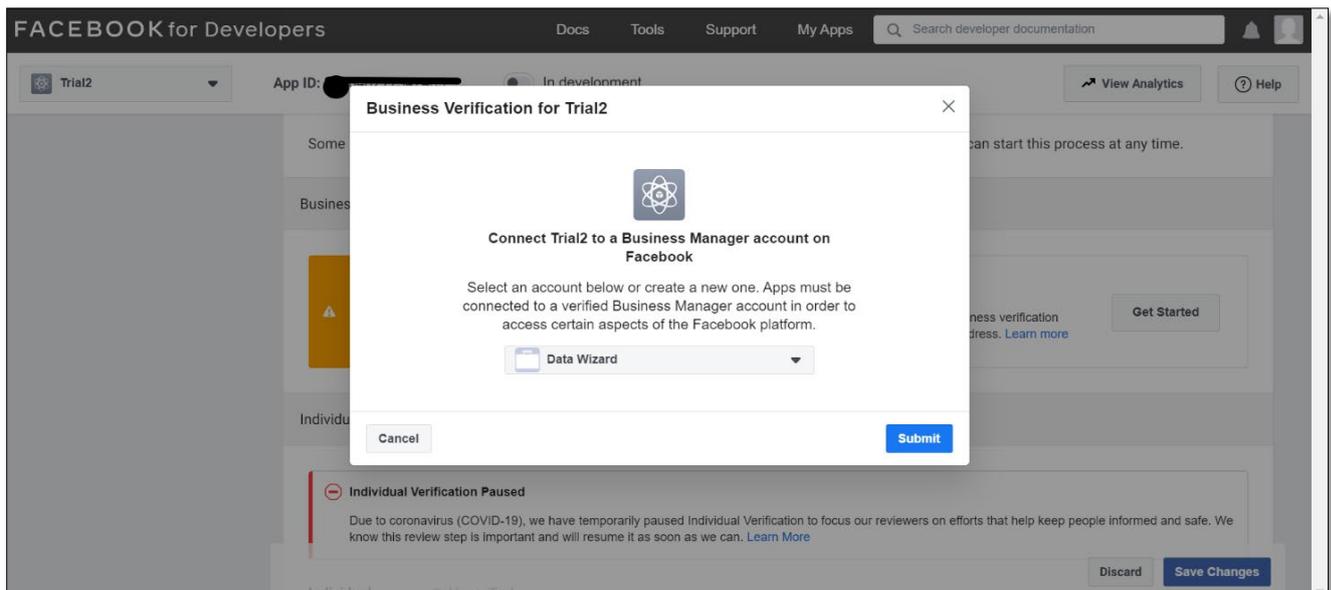


In the menu on the left side of the page - click on “Settings”, and then on “Basic”. On the right side, a page will appear, and one of your options is “Verify Your Business on Facebook”. Click on “Get Started”.

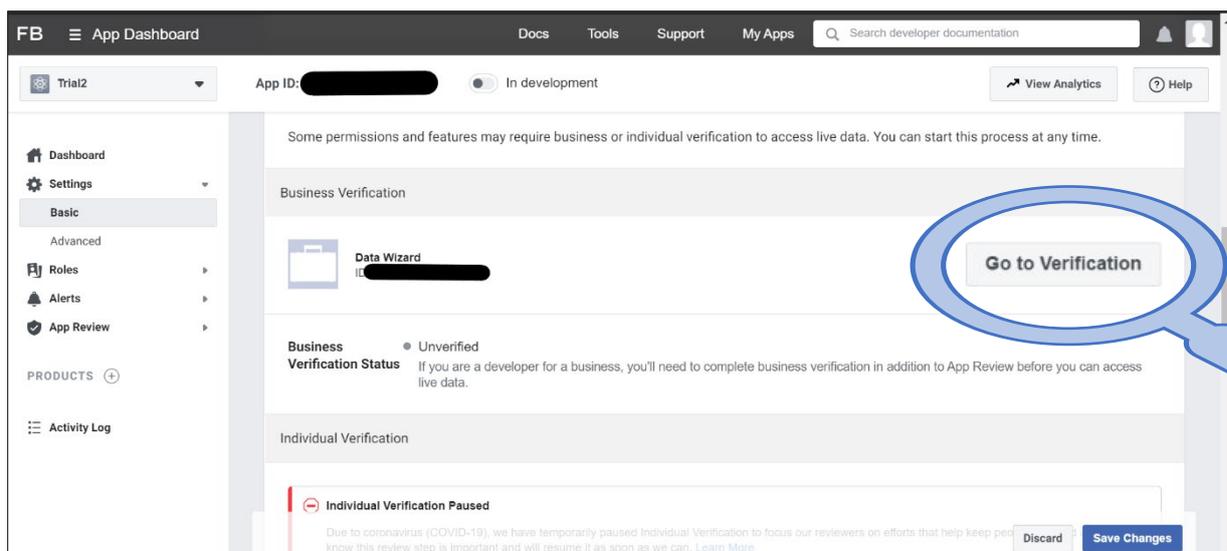




Fill in the information requested. If you are working for (or creating the app for) an NGO, everywhere you see “Business”, think of “NGO”, and fill in the information accordingly. For the test, a business account called “Data Wizard” was created.



Once you are done, you can connect your app to the newly created Business Account (“Data Wizard” in this case). If you return to the Basic Settings page of your app, you will now see that under the heading “Business Verification”, you have a Business Account connected (see screenshot below). Whenever you are ready to verify your business/organization, you can click on the “Go to Verification” button.



The business verification procedure involves an exchange of documents with Facebook, until they are convinced that your business truly exists, and will comply with Facebook rules. Depending on the status and size of your NGO, this can be a trying procedure, as the Business Verification was created with businesses, and not NGOs, in mind. That said, our experience has been that the “Page Public Content Access” app review was the most difficult of the procedures involved in getting full Public Pages API access - so if you’ve made it this far, take courage and persist!

3.5. After Public Pages Access has been granted:

Once public pages access has been granted, it is possible to start downloading data from public pages on Facebook. A few hurdles still await. Some of the major ones the authors have faced:

- Lack of open source libraries (in the programming languages R and Python) for working with Facebook, since most of the pre-existing libraries are no longer maintained after losing their public pages access in 2018¹⁵.
- Frequent changes in the Facebook API, requiring constant adaptation of code
- The volume of data that can be downloaded is limited by the number of users the app has. This holds true, even if the app is a server-to-server app, and not intended to have users. Furthermore, the documentation is not clear on exactly how the volume of data is being computed. If the app exceeds its permitted allotment of data, it is blocked for one hour. If this happens too often (again, the documentation does not clearly state how often is too often), the app could be permanently blocked.
- If the Page Public access is not used for a longer period of time, it can be revoked. It is not clear how long a period of inactivity has to be in order to cause the loss of access - in our experience, this has ranged from 3 months to over one year of inactivity.

3.6. Obtaining access to the Facebook Ad Library API:

In this section we describe how to get access to the so-called Ad Library, which is an initiative by Facebook to increase transparency on ads about social issues, elections and politics. Since May 2018, specific and mostly aggregated data about advertisers on Facebook (and recently by other Facebook products like Instagram, too) can be accessed by literally everyone.

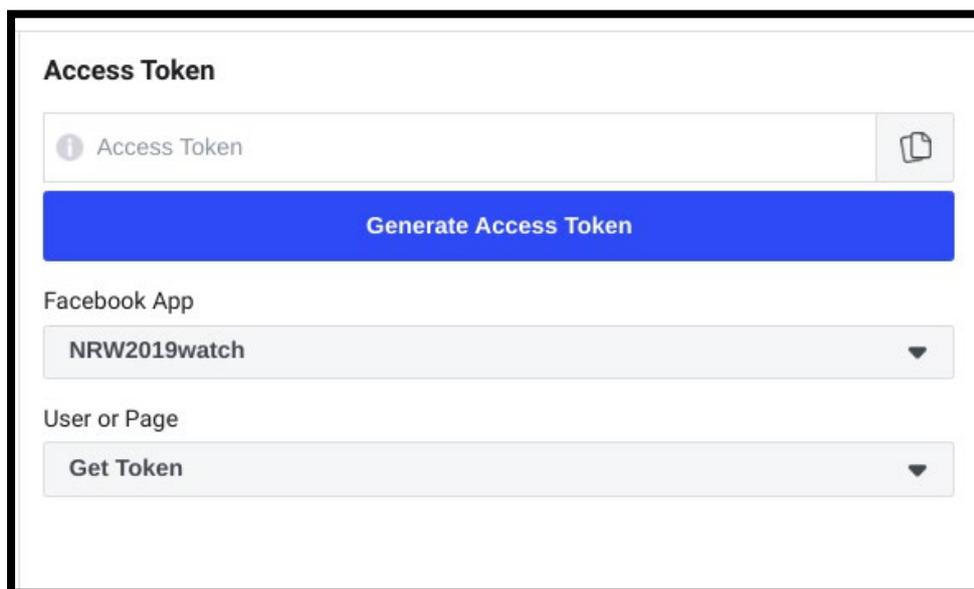
¹⁵ <https://theconversation.com/facebook-data-lockdown-is-a-disaster-for-academic-researchers-94533>

There are two distinct approaches for accessing data from Ad Library. Either by using Facebooks Graph API, which is the more complicated, code-centric, but more powerful way. Or you can view and download aggregated statistics about ads on social issues, elections and politics as reports. That way makes accessing data on such ads easily accessible to everyone (it is even possible without having a Facebook account), but comes with the drawback of being less powerful to access specific data of interest.

3.7. Access Ad Library using Graph API:

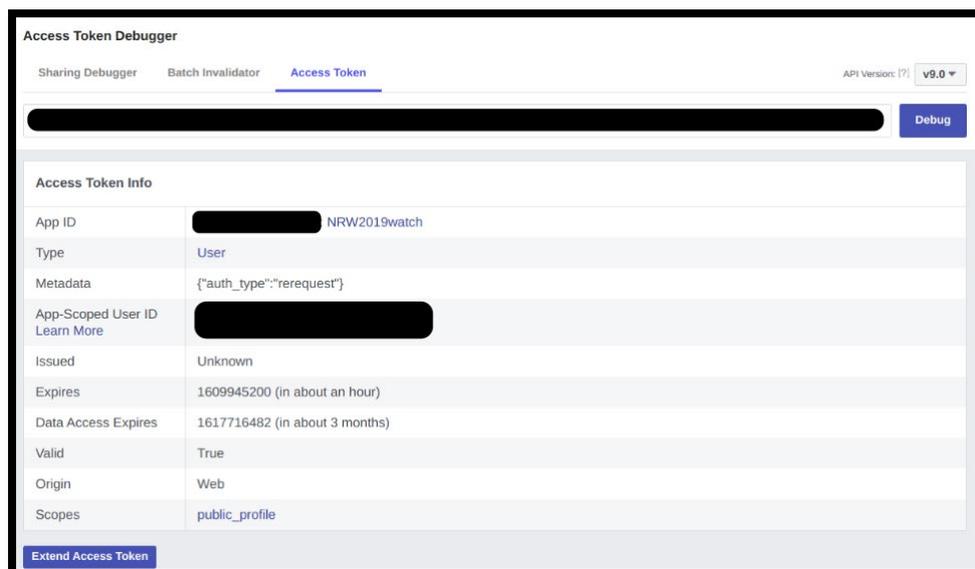
For accessing data from Ad Library programmatically, you need to have an access token. That token is mandatory for querying Facebooks so-called Graph API, which can be accessed by taking advantage of one of the many libraries in various programming languages¹⁶ or by making a plain HTTP request in a browser or on command line by yourself. As prerequisites to obtain an access token you need a Developer account (see “Creating a Facebook Developer Account” above) and an active app (see “Creating an App” above). As providing those data to the public is an initiative of transparency by Facebook, every registered Facebook Developer gets access to them, which means a Business Verification is not necessary.

In the header of your app dashboard there is a “Tools” menu. Head over it and click on the link “Graph API Explorer”. In the Graph API Explorer, you can generate an access token. On the right side of the window (image below) make sure, that the correct Facebook App is selected in the dropdown menu and then click on the “Generate Access Token”-button. Leave the preset for “User Token” as it is and there is no need to add extra permissions.



To use this token, just copy it to clipboard and paste it into a HTTP request, respectively put it into the configuration file of a programming library which you want to use to access Graph API. But make sure to keep it secret. Note, this access token expires in an hour. As it would be quite tedious to go through this described procedure repetitively, there is an option to extend the expiration date of the token. To accomplish this, copy the token to clipboard and head over to “Access Token Debugger”, whose link can be found again in the header of Facebook for Developers site under “Tools”. Paste the token into the search field and click on the “Debug”-button. You will get metadata on the token in an overview table (see image below). Then you will want to click the “Extend Access Token”-button at the bottom of the overview. After that the token can be used for the upcoming two months. To verify that extension of expiration date was successful, you can again debug the token and check if the “Expires” field is set to a date which is in about two months.

¹⁶ Which usually are wrappers for HTTP requests to the API.



By taking advantage of the obtained access to Graph API, you now have access programmatically to data from Ad Library. There are various ways to filter down the results of ads by specific parameters (for a comprehensive list of parameters, have a look at [Ad Library API documentation page](#)). In the current Version of the API (which is v9.0), the most valuable filter parameters in the context of social media monitoring during an election phase are:

- `ad_reached_country` - to specify in which country the ad had been delivered.
- `ad_delivery_day_min` & `ad_delivery_day_max` - start and stop date of ad delivery. In former versions of the API these parameters were not available. That means, in some edge cases, it was not possible to safely determine, if an ad were delivered within the date range of monitoring phase or not.
- `Publisher_platforms` - on which Facebook product the ad had been delivered.

Important to notice is, that the parameter “`ad_type`” does not need to be specified explicitly as “`POLITICAL_AND_ISSUE_ADS`” to filter for ads about social issues, elections and politics only, because this is the default value of that parameter when querying Facebooks ads archive.

Results of the query against the API should result in a response, which contains various information about the ads (for a comprehensive list of result fields, have again a look at [Ad Library API documentation page](#)). The most valuable fields highly depend of course on the specific question you want to answer with research on social media monitoring during an election phase. Some of them are:

- `ad_creative_body` - text body of the ad.
- `currency` - currency in which the ad was paid. Which is especially useful to determine if foreign actors interfered in election campaign.
- `funding_entity` - to determine which actors participated in election campaign.
- `impressions` - how often the ad was delivered.
- `publisher_platforms` - on which Facebook product the ad had been delivered.
- `spend` - amount of money spent running the ad.

If demographics of ad consumers and where they are living are of an important research interest, these two result fields are of use:

- `demographic_distribution` - demographic distribution of age and gender the ad reached.
- `region_distribution` - geographic distribution the ad reached.

Special attention is needed when dealing with the response you get back from Ad Library API. One field, namely “`ad_snapshot_url`”, provides a quick link to a view of the resulting ad in a browser. But to accomplish that special view, your Graph API access token is attached in this URL (Uniform Resource Locator). As this token should be kept secret and shared with nobody, if you persist the result as it is you generate a high risk security issue which also violates Facebooks regulations of handling that access token. Therefore, what you definitely want to do is cutting the access token out of the URL before persisting that snapshot URL of an ad to disk.

From an analytical perspective, one important issue to highlight is that these data on reach or cost of an ad are in theory valuable information to analyze social media campaigns during election phase. In practice it turns out that potential useful fields like impressions or amount spent are not returned as numbers, but in ranges¹⁷. This practice undermines transparency significantly. Because aggregating data of such wide ranges leads consequently to an inaccurate final metric (due to its high variance) on which competitors (e.g., political parties in an election campaign) should be compared against. A more fine-grained regime of discretization would be needed for getting more transparency.

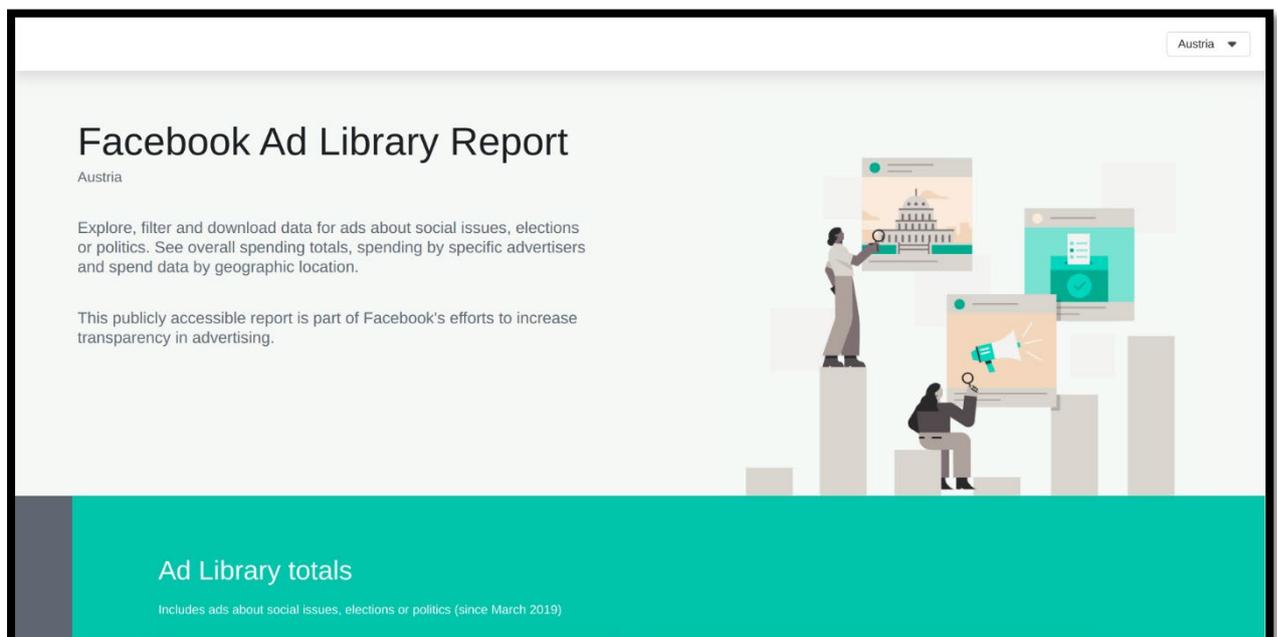
3.8. Reports:

Ad Library Reports make aggregated statistics on ads about social issues, elections and politics easily accessible for researchers, analysts or activists who might not have the skills to query an API. All that needs to be done is accessing a webpage and set specific filters for the scope of the desired reports. Thereafter you can see the aggregated statistics within interactive tables on the webpage or download them as a report. We will walk you through this stages in the following.

One important thing to notice is that the main drawback of using Ad Library Reports is that it is not possible to set a custom date range for the report. There are only 5 options for date ranges (described below) to choose from, which might collide with your chosen monitoring period within an election phase, but usually not. That makes the whole feature pretty useless in most of the cases or at least tedious to work with, when date ranges of monitoring period and Ad Library Reports overlap. Because the schedule on which time these statistics are aggregated is only known to Facebook itself - usually they are generated and published at the next day of the last date in the date range, but this rule of thumb is not reliable and days for aggregation can be intransparently skipped during weekends or public holidays.

To access Ad Library Reports, first go to its webpage, which is at the time of writing (05.01.2021): <https://www.facebook.com/ads/library/report>

If you want to access the aggregated statistics within the webpage interactively, make sure to disable any Ad-blockers in your browser. When you just want to download the reports, this step is not necessary.

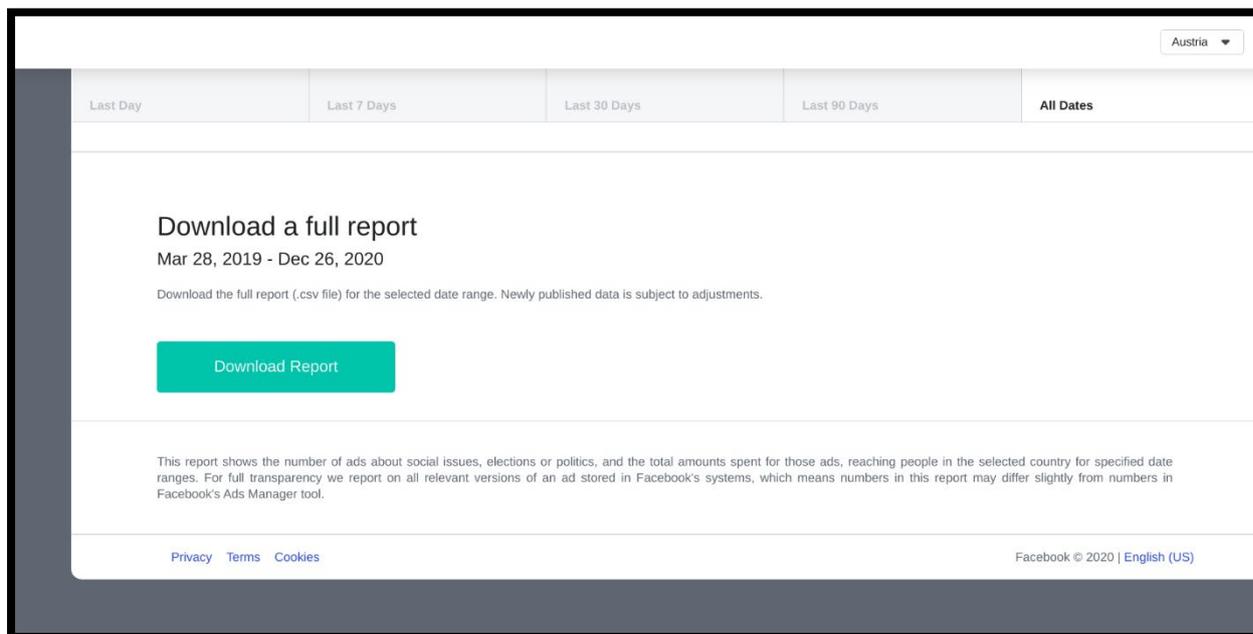


¹⁷ E.g., the amount spent ranges (in specified currency) are: <100, 100-499, 500-999, 1K-5K, 5K-10K, 10K- 50K, 50K-100K, 100K-200K, 200K-500K, >1M. Where usually most of the ads belong to one of the two lowest ranges, if the currency is EUR or USD. Therefore, this method of binning makes it difficult to effectively monitor campaigns in an age of micro-targeting.

In the upper right corner, there is a Dropdown-Button to filter the reports you want to access by country.

Download Reports:

If you just want to download aggregated statistics about ads on social issues, elections and politics in the specified country as reports, scroll down to the bottom of the webpage and a filter by date. There are only 5 options for date ranges - reports on ads from the last day¹⁸, last 7 days, last 30 days, last 90 days and all date. These fixed filters for date ranges and a lack of a custom filter makes working with the Ad Library Reports laborious in the context of monitoring an election phase. Because there is no possibility to filter the reports based on the exact monitoring period for which you want to have information about the political ads.



After having specified one of the options for the date range, you can download the reports by pressing the 'Download Reports'-Button at the bottom of the page. This invokes the download of a .zip compressed archive, which needs to be extracted on your local device to access its content. It contains the following:

- 'README.txt' file - which holds information about the content of the downloaded archive and on Ad Library Report in general.
- '*.csv' file - which is the actual report as a spreadsheet for the specified country.
- 'regions' directory - which is a folder containing individual reports¹⁹ of regions within the specified country.

Each report, either for the whole country or an individual region, contains aggregated statistics about social issues, elections and politics on *all* Facebook products. Therefore, there is no way in distinguish between individual products, like Instagram or Facebook itself. Reports contain the following fields:

- Page ID - ID of advertisers Fanpage.
- Page Name - Name of advertisers Fanpage.
- Disclaimer - Disclaimer field of Fanpage, which is mandatory to run ads on social issues, elections and politics. It should include information on the entity who paid for the ads (see [policy on disclaimers by Facebook](#)).
- Amount Spent (EUR) - total amount of money (in EUR) spent by advertisers within specified country (resp. region) and date range²⁰.

¹⁸ Attention, this is not necessary the previous calendar day, but the last day when the reports were aggregated. The exact date from which the report is generated is shown below in the section which contains the download link.

¹⁹ Again as '*.csv' files.

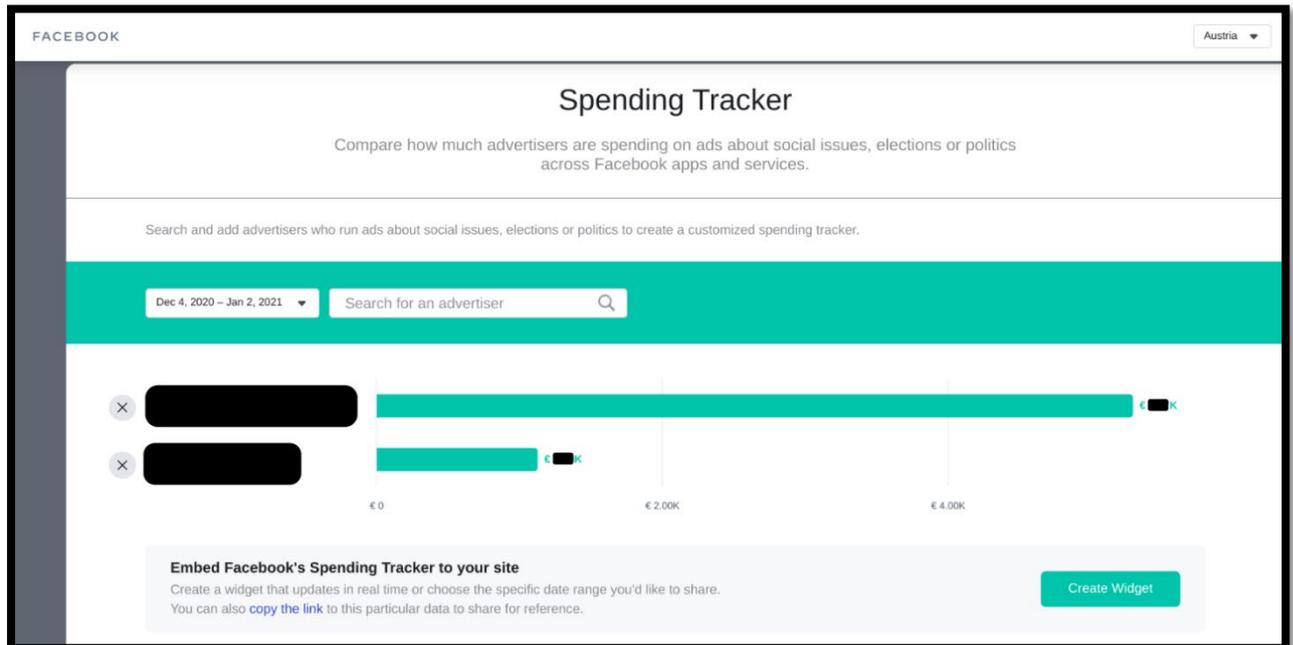
²⁰ Attention: This metric includes estimates on spending for actually running ads, therefore might be slightly inaccurate compared to the true total amount spent by advertisers within the data range.

- Number of Ads in Library - total number of ads delivered on Facebook products within specified country (resp. region) and date range.²¹

View Reports:

If you want to interactively explore data about ads on social issues, elections and politics within the webpage, the workflow is a bit enhanced. At the day of writing (05.01.2021) there are three ways to interact on the data about ads.

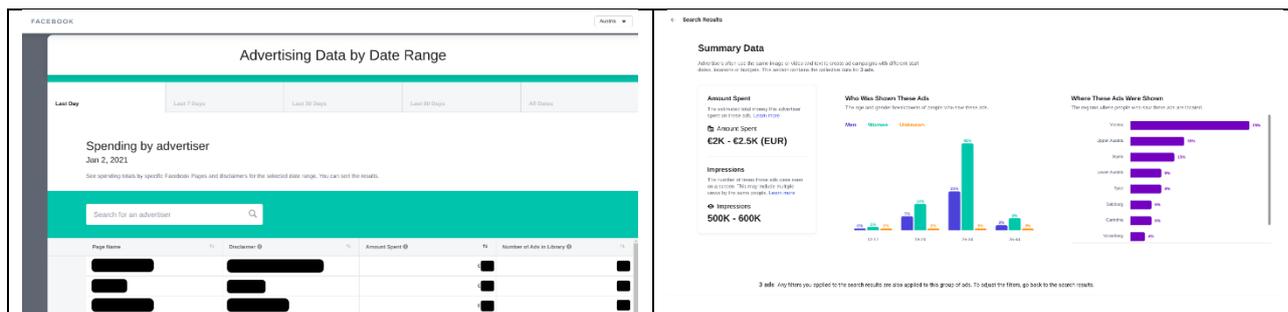
First, there is a so-called “Spending Tracker” with which you can compare total amount of spending only between different advertisers. Scroll down to the “Spending Tracker”-section (see image below), select a date range using the dropdown button and then search for names of advertisers using the text search field. You can select multiple Fanpages to compare against their total amount of spending within the date range.



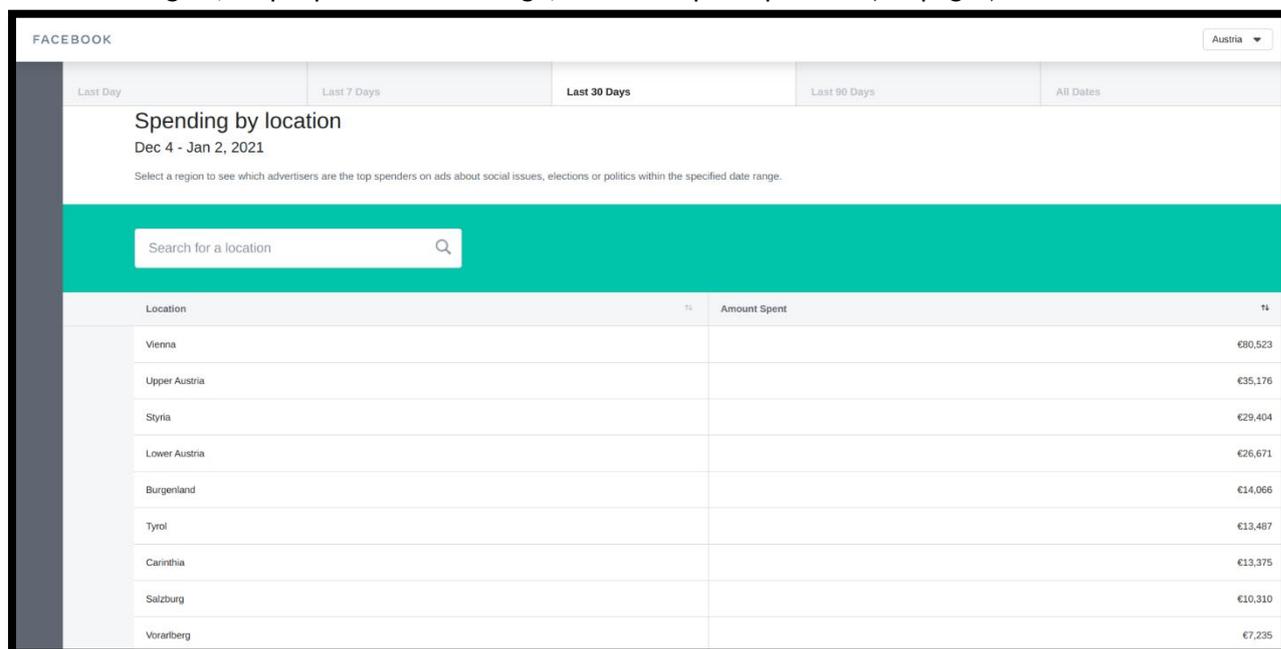
Second, you can explore Ad Library statistics by advertisers. By taking this route it is possible to get the most fine-grained statistics without using the Graph API. First scroll down to “Spending by advertiser”-section. Choose one of the 5 options for date ranges (described above). That generates a table of all advertisers, who ran ads about social issues, elections and politics sorted by the total amount spent within the timeframe (left image below). To get more details on each advertiser, click on its row and you will be redirected to the Ad Library overview page of that Fanpage. On that page you can access basic information, like follower count or total amount spent on ads. More importantly you can access all the ads who were deployed by that Fanpage. By clicking on the “See Ad Details”-button at the bottom of the ads, you can access statistics on the ad itself and to which audience it was shown (right image below)²².

²¹ Attention: This metric includes all versions of an ad and not the count of different ads.

²² Apart from potential reach, actual impressions and amount spent, it is even possible to see in which region the consumer of the ad was located and to which age, respectively gender category he/she/it belongs to.



The last way of interacting with aggregated statistics using the web interface of Ad Library is by the location of spendings. Scroll down to the “Spending by location”-section, choose a data range as described before and select a region of interest. After that, a sidebar pops up on the right side of the screen, which shows the total amount of spendings about ads on social issues, elections and politics within that region, resp. specified date range, and the Top-10 spenders (Fanpages).



3.9. CrowdTangle

CrowdTangle is an off-the-shelf tool that provides access to Facebook and other Social Media content. When it was purchased by Facebook in 2016, Facebook started restricting access to this popular research tool, but as of July 2020, it has started expanding access again to at least some academic researchers²³. As we do not have first-hand experience in using this tool²⁴, we cannot give any guidance on getting access or its usage, but wanted to mention it for the sake of completeness.

Some differences between Public Page Access and CrowdTangle that we are aware of are:

- CrowdTangle also includes insights on Instagram
- CrowdTangle includes Posts, but not Comments or Replies to Comments

More information is available on the [CrowdTangle help page](#).

23 <https://research.fb.com/blog/2020/07/crowdtangle-opens-public-application-for-academics/>
 24 As non-University based researchers, we are currently not permitted access.