

Product Document

Team Number: 2

Team Members:

Mark Daly
Evan Dunbar
Conor McGavin
Conor Heeney
~~Joel McKenna~~

Website URL: <http://tweetguard.me>

Disclaimer 1: While interacting with this application, you will view offensive and inappropriate tweets. These words and statements do not reflect the views of the developers. The very nature of the app is to aid users in minimising exposure to these tweets. We have, however, redacted offensive tweets from this document.

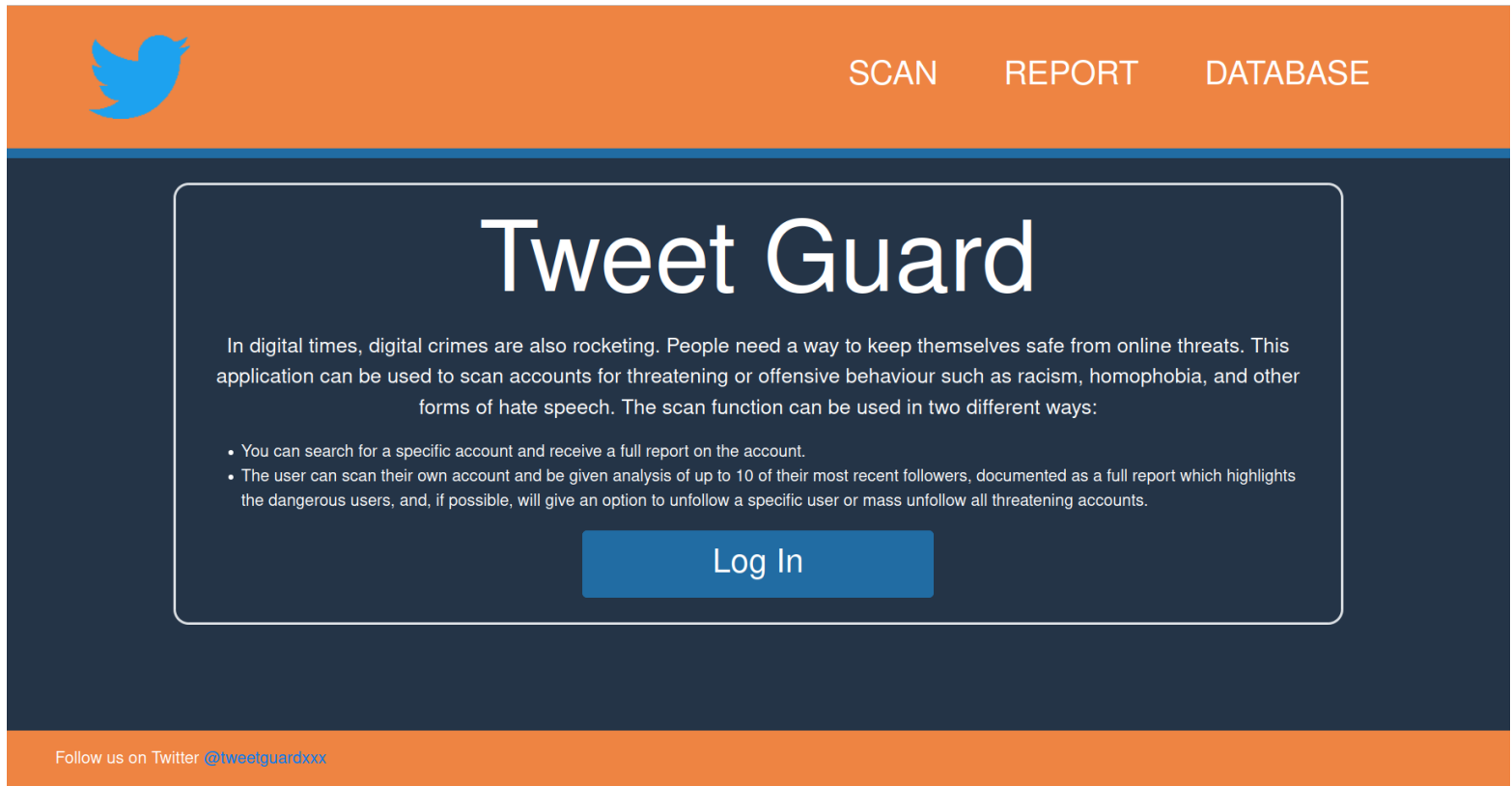
Disclaimer 2: To avail of the full functionality of the application, a Twitter account is needed. If you do not have a Twitter account, and would prefer not to make one, you can use the credentials for our admin account during log in:

Username: TweetGuardxxx

Password: DigitalCrimes5

Disclaimer 3: There are a number of input fields throughout the app, where you can enter in a user's Twitter handle. **You should not include the "@" symbol at the beginning of the handle.** For example, a user's Twitter handle is @firstname.secondname99. To act upon this account, simply type "firstname.secondname99".



Home Screen



This is the entry point to the application and the first page you will see. To avail of the full functionality of the application you need to log into Twitter. On the top-right are the different sections of the website. This is the main navigation bar of the site. You can click these links to go to each section, but you will need to log in to actually use the functionality. If you are not logged in, and you attempt to scan or report, you will be prompted to log in. The Twitter bird on the top-left can be clicked on any page to be redirected to the homepage. On the footer is a link to our admin Twitter account's page. To get started, press the "Log In" button. This will take you to the Twitter API's OAuth screen to log into Twitter.

Note: If you login with our admin account, there is a small possibility that the "Log In" button will not change text to be "Log Out". If this happens, please append "/logout" to the base url. Then when you log back in you will have access. This error can only occur on our admin account as it was originally entered in the database.

OAuth Log In

 tweetguardxxx

Authorize TweetGuardxxx to access your account?


Authorize app

Cancel

This application will be able to:

- See Tweets from your timeline (including protected Tweets) as well as your Lists and collections.
- See your Twitter profile information and account settings.
- See accounts you follow, mute, and block.
- Follow and unfollow accounts for you.
- Update your profile and account settings.
- Post and delete Tweets for you, and engage with Tweets posted by others (Like, un-Like, or reply to a Tweet, Retweet, etc.) for you.
- Create, manage, and delete Lists and collections for you.
- Mute, block, and report accounts for you.

Learn more about third-party app permissions in the [Help Center](#).



TweetGuardxxx

tweetguard.me

Defender of Twitter Users

We recommend reviewing the app's terms and privacy policy to understand how it will use data from your Twitter account. You can revoke access to any app at any time from the [Apps and sessions](#) section of your Twitter account settings.

By authorizing an app you continue to operate under Twitter's [Terms of Service](#). In particular, some usage information will be shared back with Twitter. For more, see our [Privacy Policy](#).

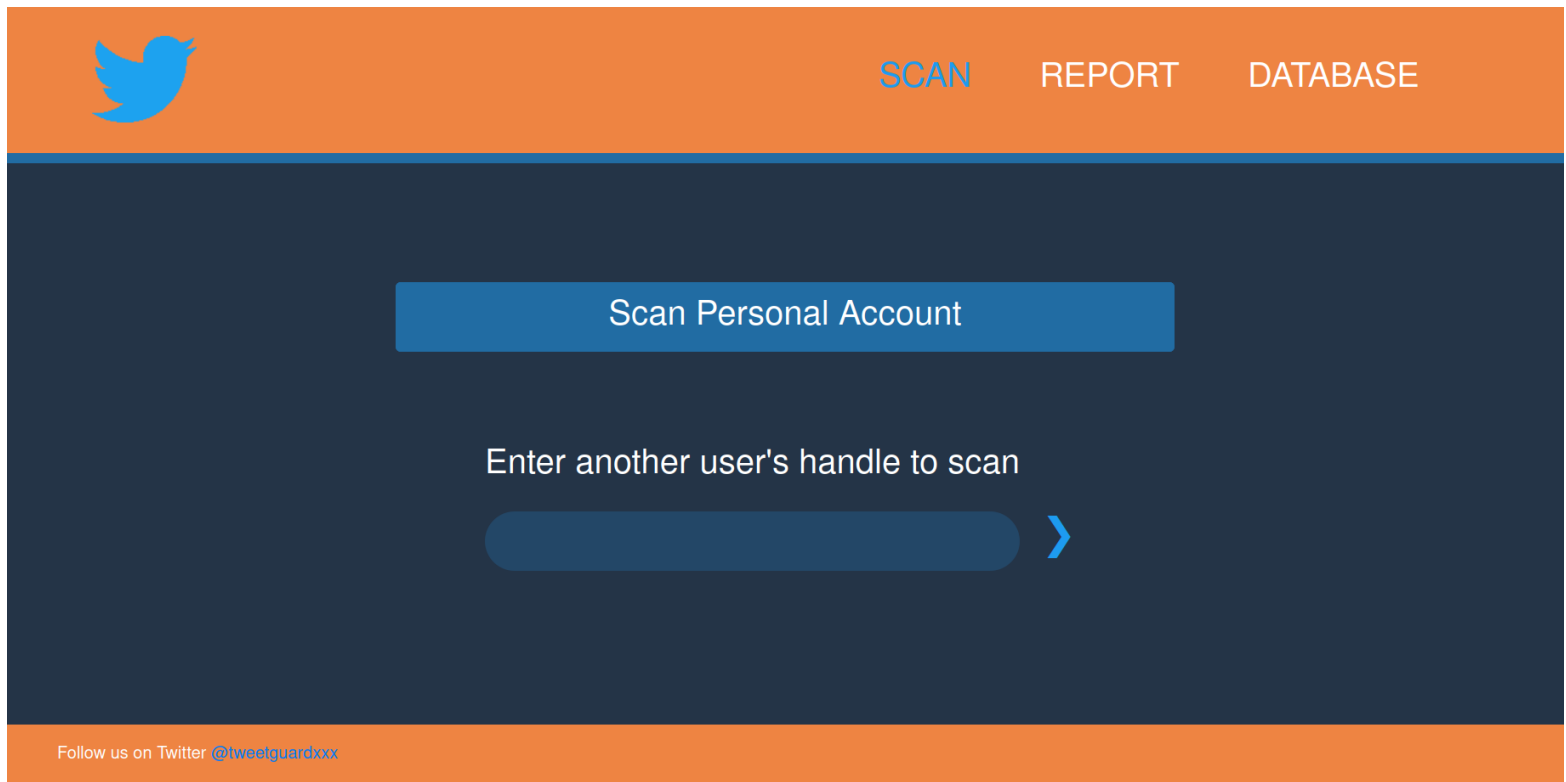
After clicking “Log In” on the homepage, you’ll be redirected to Twitter’s OAuth page. Here you can log in with the Twitter account you want to use. To log in, enter your username/email in the first input, and your password in the second input. You may check “Remember me” before you log in so that anytime you log in again in the future, you won’t need to enter in your username and password. If you forgot your password, click the “Forgot password?” link to start the password reset process on the Twitter website. There is also the “Sign Up for Twitter” link in the top-right. If you don’t have a Twitter account, you can create one at this link. A reminder, however, that you can use our admin credentials for the purposes of the beta. Credentials are at the top of the document. Once you have entered your username and password, click “Authorize app” to be redirected to the site.

Back to Home Screen



Once you log in, you will be redirected back to the home screen. As you can see, the “Log In” button text has now changed to “Log Out” you can now begin navigating the website. If you look at the top navigation bar, you will see the “Scan” link is highlighted blue. This is the hover effect on the links. You can click a blue link to be redirected to that section. We will begin with the scan section. Press the scan link in the nav bar.

Scan Home




This is the homepage of the scan section. This section is to scan for racist tweets, as determined by our algorithm that uses logistic regression.

There are two parts to the scan functionality. You can either tell the application to get a list of your followers. You will then be able to select which users to receive a scan report for, or you can enter a specific user to scan. This page is where you choose one of these options. To view a list of your followers, click the “Scan Personal Account” button. This will take you to a new page where you can choose who to scan.

Alternatively, you can enter in a specific user handle into the text input box. Click the right arrow to submit the user’s handle and receive a report. There may be three outcomes from the submission of the input box. The first, and most desirable is being redirected to the scan report page. The second is the case where the account is protected. This happens if they have protected Twitter privacy settings. You will receive a red flash message that tells you that the account can’t be scanned. The third situation is the case where the user handle doesn’t match any account. You will receive a red flash message to tell you the account can’t be scanned. First, we will look at the option to view a list of your followers. Press the “Scan Personal Account” button to begin. We will come back to this page to demonstrate the second input later.

NOTE: Depending on the amount of followers you have, there may be a delay in displaying them. This is due to limitations with the Twitter API. While the application is collecting your followers, the Twitter bird will be replaced by a loading icon. This is how you know the application is working on it. You will also see this icon on when the application is scanning users, as this may have a delay as well.


Scan Choose




SCANREPORTDATABASE

Follower list for @tweetguardxxx

Select from your followers using the list below


@thisisatest419


@markdalyy

Scan Selected

Or choose up to 10 of your most recent followers to scan!

Follower Count:

0

Go

Follow us on Twitter @tweetguardxxx

Once the application has collected your followers, you will see them listed here. Again, you have two options on this page. You have the option to select certain followers to scan from the list, or you can use the slider to scan up to ten of your most recent followers. Both of these options will redirect you to the same page, with the results depending on what you select. To select specific followers, click on their profile pictures. When a user is selected, their picture will fade to indicate they've been selected. You can unselect a selected user by clicking their profile picture again. Once you've selected the users you want to scan, press the "Scan Selected" button. The Twitter bird will be replaced by the loading icon and scanning will begin.

To select a number of your most recent followers, simply slide the slider until the number underneath indicates the amount of followers you want. Once you have selected the desired amount of followers, press "Go!" beside the slider and scanning will begin.

Note: Our algorithm can take a while to scan users depending on how many tweets they have and how many followers were selected. Selecting a small amount of followers to scan will save you waiting too long for results.

Selected Followers



Tweet Guard
@tweetguardxxx

Click on a threatening user's handle to **unfollow**, **mute**, **block**, or **view user-made reports** on them!

Threatening Follower 1



please work
@thisisatest419

Total Reports on Account: 0

Total Tweets Scanned: 8

Total Tweets Flagged: 2

Danger Level: 2

Nigga nigga nigga Nigga Nigga nigga nigga Nigga nigga nigga Nigga Nigga nigga nigga Nigga nigga nigga
Nigga nigga


[View Tweet](#)

This is the result page for the scan choose page. When you select specific followers, or you select a certain number of followers on the slider, the results of the scan are shown here. On the top, you will see your Twitter profile picture, name and handle. Then you will see a list of users whose scan returned racist tweets. For each user, there will be a bordered box with a summary of their Twitter info, and underneath that, will be a list of their racist tweets. You can press the “View Tweet” button underneath each tweet which will send you to the Tweet displayed on the Twitter website. If no threatening tweets were found for any of the selected users, the page will tell you no threatening tweets were found.

You also have the ability to unfollow, block and mute flagged users by clicking on their user handle.

Now is the time to head back to the homepage of the scan section to touch on the second option, where you can enter in a specific user's Twitter handle. Click the "Scan" link in the navigation bar to return to this page.

Returning to Scan Home



SCAN REPORT DATABASE

Scan Personal Account

Enter another user's handle to scan

>

Follow us on Twitter @tweetguardxxx

Recall that you can enter in a specific user handle into the text input field to scan their account. This works for users who are not followers of yours, however, as previously stated, if the account is protected, you won't be able to scan them. Once you enter a user handle, press the right arrow and the scan will begin.

Scan User



please work
@thisisatest419

Total Reports on Account: 0

Total Tweets Scanned: 8

Total Tweets Flagged: 2

Danger Level: 2

Threatening Tweets Found:


Nigga nigga nigga Nigga nigga nigga nigga nigga
nigga nigga nigga nigga nigga nigga nigga nigga
nigga nigga nigga

[View Tweet](#)

The resultant page is very similar in design and content as the selected followers page. You will see the scanned user's twitter information and a summary of their account. Then there is a list of that user's threatening tweets with an option to view each one on Twitter. These tweets will be paginated to make navigation easier, with a maximum of 10 tweets per page.

Now it's time to move on to the report section. Click the "Report" link in the navigation bar to view this page.

Report



SCAN REPORT DATABASE

Account Name

Report Summary
(at least 10 characters)

Threat Type

Racism

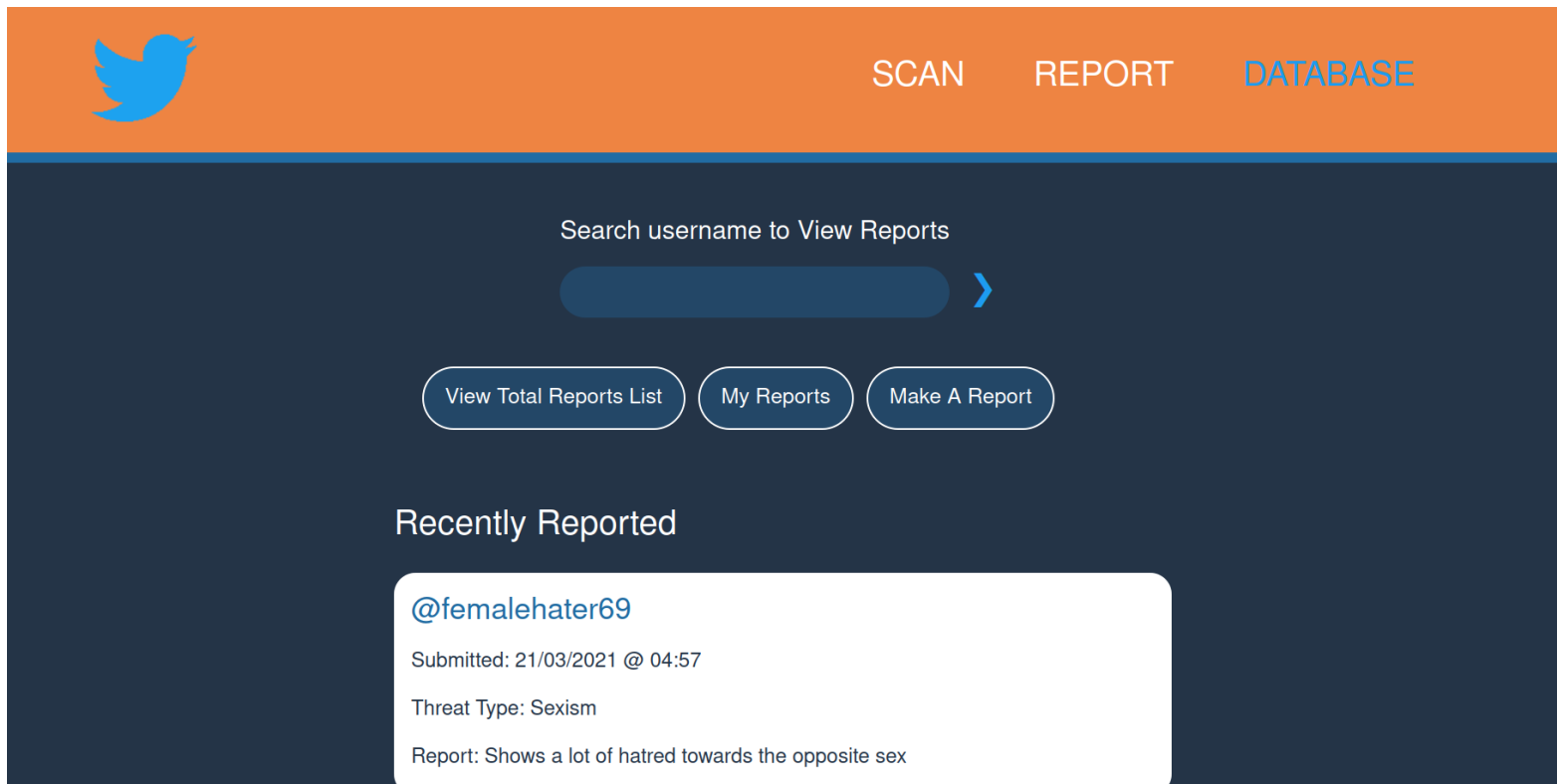
Report Account

Follow us on Twitter @tweetguardxxx

On this page, you can report a specific user based on a number of categories in the Threat Type dropdown menu. This adds an extra layer to the investigation of the account by adding a human element to reports. To make a report, enter a user's Twitter handle, a summary of the inappropriate behaviour and select the category it falls under. Once you are satisfied with the report, press the "Report Account" button. This report will be entered into our database and you can view it, alongside other user reports in the Database section. As with the scan section, if you report an account that doesn't exist, a red flash message will appear, and the report won't be submitted. When a report is successfully submitted, a green flash message will appear telling you the report was successful.

Make sure to enter a valid account username, and to put at least 10 characters of a summary in of what happened.

Database Home



The database home page hosts two main features, and also has access to other sub-pages.

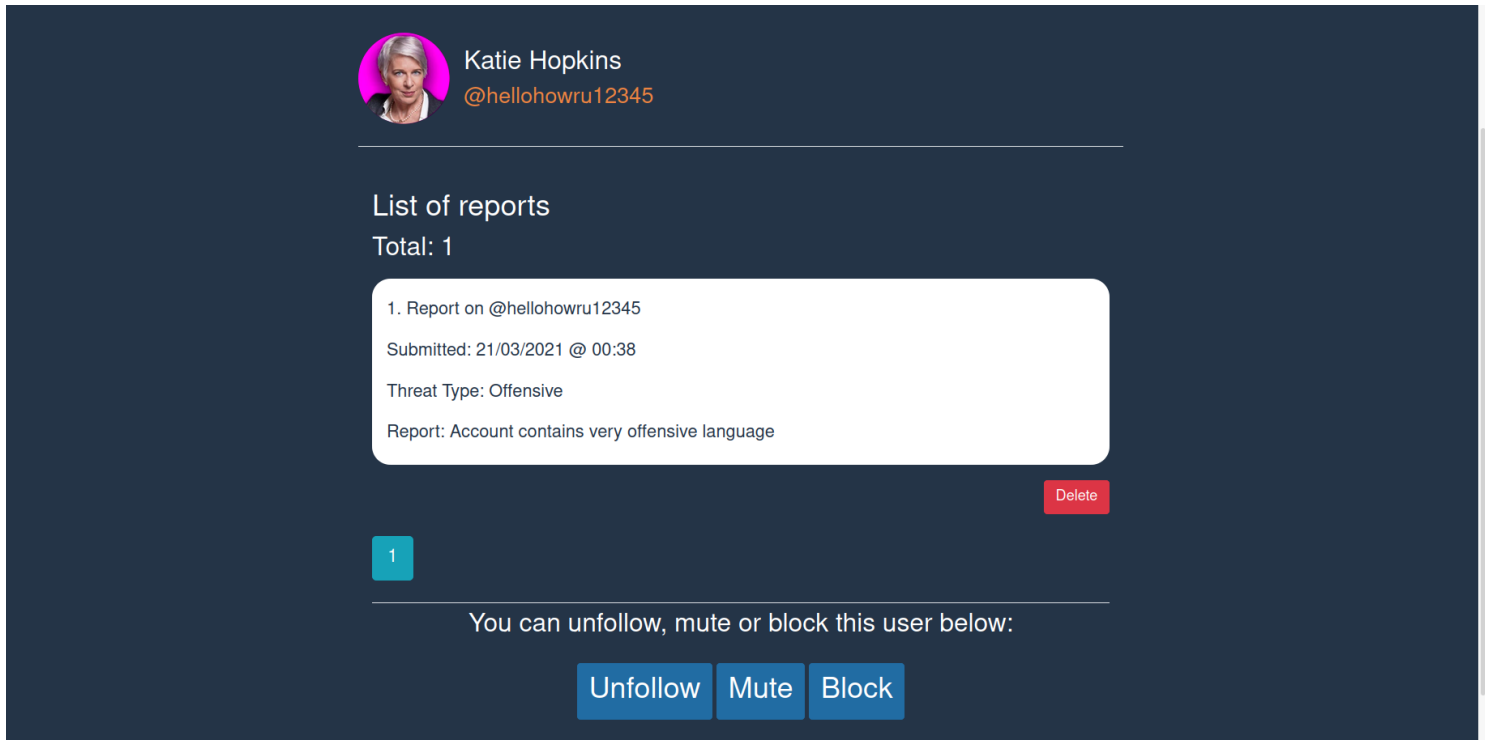
Firstly of the two features is a database search, where the user can enter an account name to search up all the associated reports with it. The account name can only be submitted and searched if it is a valid Twitter account name. Twitter users with protected accounts are still searchable in this database, and can be reported in general.

Secondly on this page, we see a recent reports section in which the most recent user-made reports are shown. This section aims to keep our customers informed on the most dangerous recently active accounts on Twitter. The usernames are clickable, returning the database search page of the account with that username. Note also that these reports are paginated, so click the numbers underneath to flick through the most recent reports.

Between the recent reports section and the search section there are three links: one to the “total reports list” page, one to the “make a report” page, and one to the “my reports” page. The contents and functionality of these pages are detailed in this document.

This page is available without a login.

Database Search



The screenshot shows a user profile for Katie Hopkins (@hellohowru12345). Below the profile information, there is a section titled "List of reports" with a total of 1 report. The report details are as follows:

Report ID	Report Text	Submitted At	Threat Type
1	Account contains very offensive language	21/03/2021 @ 00:38	Offensive

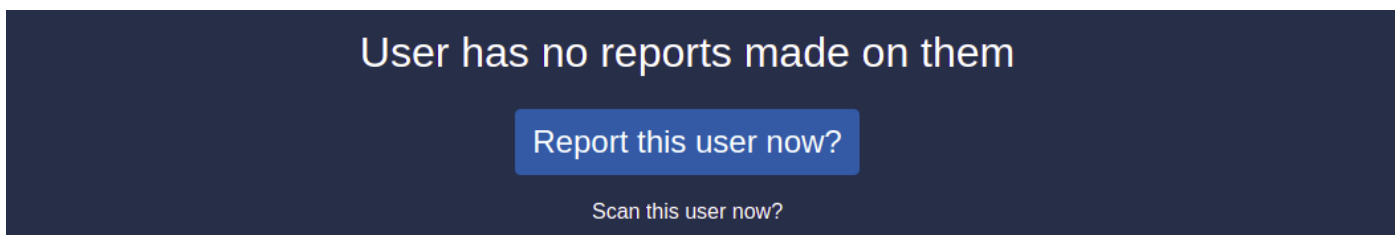
Below the report list, there is a pagination button labeled "1". At the bottom of the profile section, there are three buttons: "Unfollow", "Mute", and "Block".

The database search page can be found by either clicking a username that is a link, or by using the search feature on the database home page. This feature is available without a login.

This page content is simple and only contains some small info on the account, and then the list of user reports that were made on the account. These are also paginated so to look through them use the numbered buttons underneath. You also have the option to unfollow, mute or block the account by clicking on the buttons underneath.


If you have made the report on the user yourself, you are given the option to delete that report if you feel you have made an error or no longer wish to account for it on our website. Simply click the red delete button, and then choose confirm on the box that comes up to delete a report.

If the user has no reports made on them, you will see a section like this under their username, which gives options to make a report or to scan that user.



The screenshot shows a user profile with the text "User has no reports made on them". Below this text, there are two buttons: "Report this user now?" and "Scan this user now?".




Total Reports



SCANREPORTDATABASE

Total Reports On Each Account


Sorted by number of reports


Account	Reports
<div><div>ant @jewkillerant</div></div>	1
<div><div>Jake Paul @jakepaul</div></div>	1
<div><div>Katie Hopkins @hellohowru12345</div></div>	1

This is a simple page that ranks the accounts that have been most reported in the past. They are ranked in order of the total number of reports made. It aims to provide customers with a list of the most dangerous accounts.

It is accessible with or without login.

My Reports

LOG INREPORTDATABASE



Tweet Guard

@tweetguardxxx

List of reports

Total: 8

1. Report on @femalehater69

Submitted: 21/03/2021 @ 04:57

Threat Type: Sexism

Report: Shows a lot of hatred towards the opposite sex

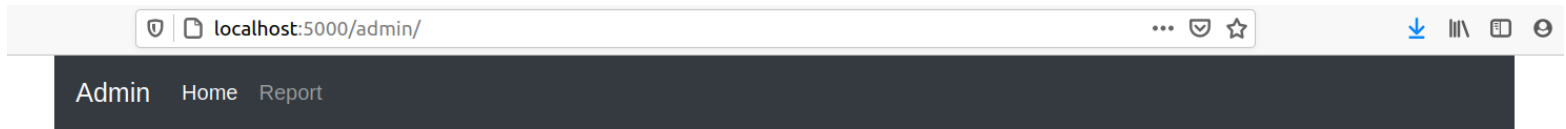
Delete

2. Report on @jewkillerant

This page contains all the reports that the logged in user has made. Its main aim is to allow users to manage their made reports, and gives the main functionality of deleting any reports that they are not happy with, or simply just don't want on the site anymore.

This page is naturally only available if you are logged in.

















Admin System



 View Reports →

 Exit →

To access the admin system, you must be logged in with our tweetguard account (credentials on page 1). You can then add “/admin” to the base URL and will be directed to the page above. The View Reports arrow will bring you to view all of the reports in the database. The Exit arrow will return you to the web application home page. The admin system allows us to delete reports from users as shown below.

Admin Home Report						
List (8)	Create	With selected ▾				
<input type="checkbox"/>		Threat Type	Summary	Date Submitted	Account Id	Author
<input type="checkbox"/>	 	Offensive	Tends to use the "B" word quite a lot	2021-03-21 00:25:56.344162	dojacat	<User 1>
<input type="checkbox"/>	 	Offensive	Generally a delinquent	2021-03-21 00:34:45.020095	jakepaul	<User 1>
<input type="checkbox"/>	 	Offensive	Account contains very offensive language	2021-03-21 00:38:37.764060	hellohowru12345	<User 1>
<input type="checkbox"/>	 	Violence	Contains some violent sexual comments	2021-03-21 04:38:51.321561	fuckthispunter	<User 1>
<input type="checkbox"/>	 	Racism	Completely and utterly racist account	2021-03-21 04:46:33.798343	fuckblackpeopl8	<User 1>
<input type="checkbox"/>	 	Racism	Great example of a pure racist	2021-03-21 04:50:14.120853	fuckfaggots22	<User 1>
<input type="checkbox"/>	 	Homophobia	Continuously uses the word "faggot" and is racist towards jewish people	2021-03-21 04:54:26.963248	jewkillerant	<User 1>
<input type="checkbox"/>	 	Sexism	Shows a lot of hatred towards the opposite sex	2021-03-21 04:57:41.496947	femalehater69	<User 1>