



USER MANUAL

JOKITOT

Abstract

A complete guide to installing and using tool 1 and tool 2

Team 21

Contents

Overview	2
Installation	2
1. Get Repository	2
2. Open a terminal in the directory where the code is stored	2
3. Create a virtual environment	2
4. Open the virtual environment	2
Using Windows PowerShell	3
Using Windows Command Prompt	3
Using Linux/macOS	3
5. Install the Required libraries	3
6. Run the tool	3
Creating AWS LightSail Service	3
Requirements	3
1. Delete the virtual environment	3
2. Navigate to the folder containing the code in your terminal	3
3. Build docker container	4
(Optional) Test the docker build	4
4. Create container service	5
5. Push image to LightSail	5
6. Change deployment version	6
7. Deploy the image	6
8. Check status	6
9. Open website	7
(Optional) Delete container	8
Updating AWS LightSail	8
1. Download the new code from the repository	8
2. Follow the steps from create AWS LightSail service	9
Tool 1	9
File Upload	9
Interaction Menu	9
Multiple Documents Analysis	10
Tool 2	11
At a glance	11
Upload Your CSV File:	11

Overview

Jokitot is a plagiarism detection tool consisting of two tools. Tool 1 is designed to analyse word documents and present its metadata. Tool 2 is designed to detect contract cheating using lms logs to determine the location of users from their IP addresses.

The GitHub repo to the code can be found at https://github.com/lujor20/CITS3200_Project

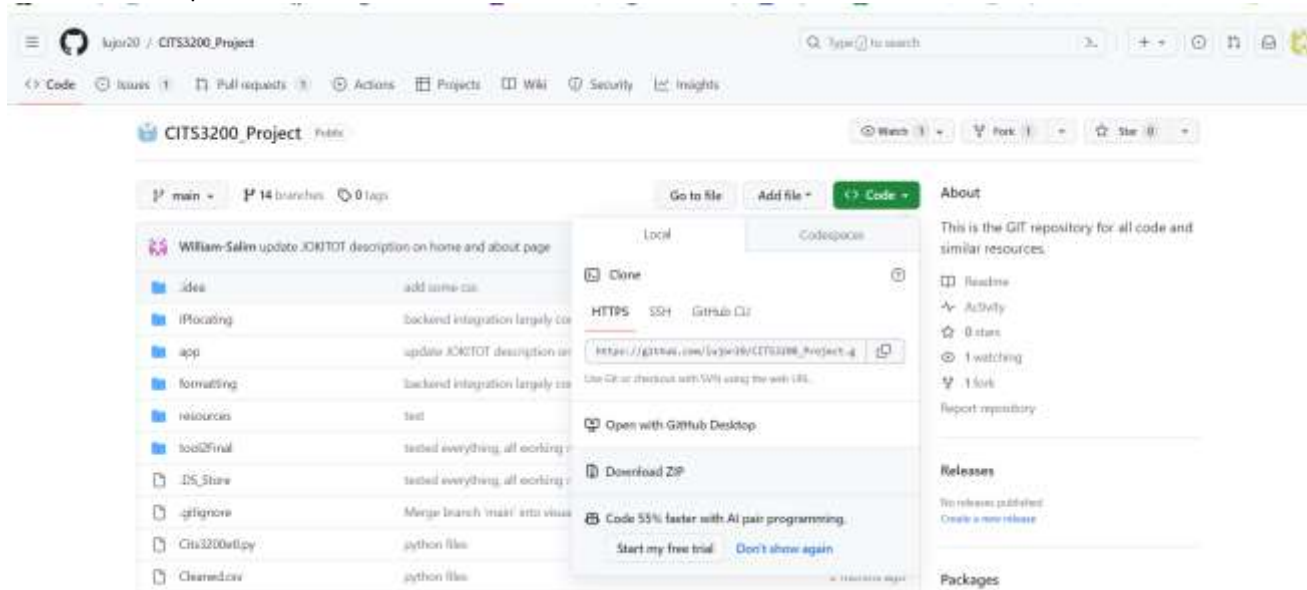
Installation

1. Get Repository

This step is to get the latest code from the repository, which can be done by a command in the terminal:

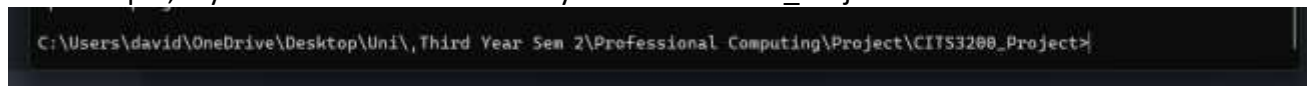
```
git clone https://github.com/lujor20/CITS3200_Project.git
```

Or by downloading the zip file by going to the website https://github.com/lujor20/CITS3200_Project and clicking the download zip button under the button "Code".



2. Open a terminal in the directory where the code is stored

For example, my terminal is in the directory called CITS3200_Project



3. Create a virtual environment

This creates a virtual environment to run the code in.

```
python3 -m venv dev-env
```

4. Open the virtual environment

The next step is to enter the virtual environment. A different version of the same command will need to be run depending on the type of terminal used.

Using Windows PowerShell

```
dev-env\Scripts\Activate.ps1
```

Using Windows Command Prompt

```
dev-env\bin\activate.bat
```

Using Linux/MacOS

```
source dev-env/bin/activate
```

5. Install the Required libraries

This will install all dependencies the code needs to run. It will take approximately 3 to 5 minutes.

```
pip install -r requirements.txt
```

6. Run the tool

```
flask run
```

Creating AWS LightSail Service

Note: This is not required to run the code on a single machine!! This will deploy the website using AWS so the website can be accessed by anyone.

Requirements

For the tool to be deployed using LightSail, the requirements needed are:

- AWS account
 - <https://aws.amazon.com/getting-started/guides/setup-environment/>
 - You will need to use a debit/credit card when creating your account.
- Configured AWS Command Line Interface (CLI)
 - <https://docs.aws.amazon.com/cli/latest/userguide/getting-started-install.html>
- Docker installation
 - <https://docs.docker.com/engine/install/>
- LightSail control (lightsailctl) plugin
 - https://lightsail.aws.amazon.com/ls/docs/en_us/articles/amazon-lightsail-install-software

Note: If installing on a university computer, you will need to contact the University IT department for temporary admin privileges to install the above requirements

Once all the requirements are installed and the code within this repo downloaded, we can start the steps below.

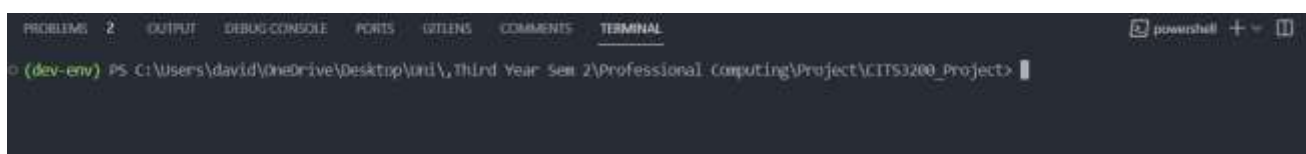
1. Delete the virtual environment

We will no longer need the virtual environment since building the docker container will download all the requirements again.

Delete the folder called dev-env from the folder containing the code. If you don't have this folder, skip this step.

2. Navigate to the folder containing the code in your terminal

I have my code stored in the folder CITS3200_Project



3. Build docker container

Run the command in the terminal

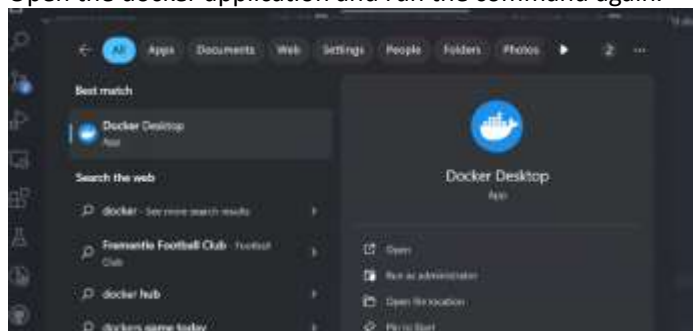
```
docker build -t flask-container .
```

This may take from 15 mins to 30 mins to finish.

If you see this error:

```
PS C:\Users\DavidOneDrive\Desktop\AI\Third Year Sem 2\Professional Computing\Project\CITS3000_Project> docker build -t flask-container .  
error during connect: This error may indicate that the docker daemon is not running.: Post "http://25273f3232/pipfile/docker/engine/v1.24/build/buildings/5d786fca  
chfe9e268d5c/ggoprogram-kcpuried-08cquqao-&kpusetcpu=&kpushares=&kpushares:&ddockerfile=Dockerfiles&label=3700/0memory=&nesssap=&networknode=default  
time-lifetime=00-flask-container-target=multisite-buildversion=1": open //./pipe/docker engine: The system cannot find the file specified.  
PS C:\Users\DavidOneDrive\Desktop\AI\Third Year Sem 2\Professional Computing\Project\CITS3000_Project>
```

Open the docker application and run the command again.



This is what will be displayed if successful.

[illegible]

(Optional) Test the docker build.

To test the build, run the following command.

```
docker run -p 5000:5000 flask-container
```

Click on the <http://127.0.0.1:5000> and the website should work as intended.

```
PS C:\Users\david\OneDrive\Desktop\Uni\third Year Sem 2\Professional Computing\Project\CI55200_Project> docker run -p 5000:5000 flask-container
* Debug mode: off
WARNING: This is a development server. Do not use it in a production deployment. Use a production WSGI server instead.
* Running on all addresses (0.0.0.0)
* Running on http://127.0.0.1:5000
* Running on http://172.17.0.2:5000
Press CTRL+C to quit
```

Close the server by clicking the stop (square) button in the docker application.



4. Create container service

This step will create a container on your LightSail account. The name of the container is team21plagiarismdetection.

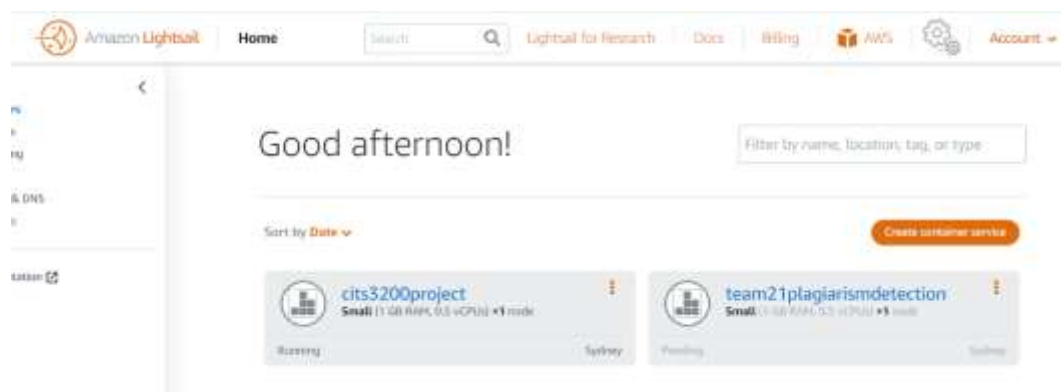
```
aws lightsail create-container-service --service-name  
team21plagiarismdetection --power small --scale 1
```

If it responds with a configuration error, make sure to configure (login) to the AWS CLI first and run the command again. (This is extremely complicated and there are guides that explain it better on the internet). See below websites for more detail.

- <https://docs.aws.amazon.com/cli/latest/userguide/cli-chap-authentication.html>
- <https://docs.aws.amazon.com/cli/latest/userguide/cli-configure-ssso.html>

The container will appear as "team21plagiarismdetection" in the LightSail website <https://lightsail.aws.amazon.com/ls/webapp/home/containers>

The "cits3200project" container was created previously.



5. Push image to LightSail

This will upload the docker image to the AWS container. It may take from 5 mins to 10 mins.

```
aws lightsail push-container-image --service-name team21plagiarismdetection --  
label flask-container --image flask-container
```

If successful, it will output:

Take a note of the output. In this case, it tells me I should refer to this image as **:team21plagiarismdetection.flask-container.7**

```
00060bc1f1e5: Pushed
1e8c11cb2977: Pushed
2f50256a8712: Pushed
a1c2f058ec5f: Pushed
cc2447e1835a: Pushed
Digest: sha256:99ad952de88aaf5ab497d6e7ba90786f45393ef71f4240225e30fe1d1e288cce
Image "flask-container" registered.
Refer to this image as ":team21plagiarismdetection.flask-container.7" in deployments.
PS C:\Users\david\OneDrive\Desktop\Uni\Third Year Sem 2\Professional Computing\Project\CITS3200_Project>
```

6. Change deployment version

Open the file "containers.json" in the folder containing the code (with notepad)

```
You, 3 weeks ago | 1 author (You)
1 {
2   "flask": {
3     "image": ":cits3200project.flask-container.6",
4     "ports": {
5       "5000": "HTTP"
6     }
7   }
8 }
9
```

Edit the line:

"image": ":cits3200project.flask-container.6"

and replace with the deployment version the terminal just outputted from the step above. In my case, this is my new file.

```
You, 1 second ago | 1 author (You)
1 {
2   "flask": {
3     "image": ":team21plagiarismdetection.flask-container.7",
4     "ports": {
5       "5000": "HTTP"
6     }
7   }
8 }
9
```

7. Deploy the image

This will make the website for you. This could take from 5 mins to 10 mins.

```
aws lightsail create-container-service-deployment --service-name
team21plagiarismdetection --containers file://containers.json --public-
endpoint file://public-endpoint.json
```

8. Check status

Open the LightSail website.

<https://lightsail.aws.amazon.com/ls/webapp/home/instances>

Navigate to the containers and click on the container. You will see in the deployments tab (when you scroll to the bottom) that the status is deploying.

Health check path
/
[More details](#)

Deployment versions

All containers deployed to your container service are recorded as versions. You can roll back your current deployment to a previous version.
[Learn more about deployment versions](#)

Date	Version	Status
October 15, 2023 - 2:16 PM	1	Deploying

1 of 1 items

Wait until it is active. It usually takes around 3 mins.

[More details](#)

Deployment versions


All containers deployed to your container service are recorded as versions. You can roll back your current deployment to a previous version.
[Learn more about deployment versions](#)

Date	Version	Status
October 15, 2023 - 2:16 PM	1	Active

1 of 1 items

9. Open website

At the top of the page, there is a hyperlink labelled "public domain:".



team21plagiarismdetection

Container service
Small (1 GB RAM, 0.5 vCPU) • 1 node
Sydney

Public domain: team21plagiarismdetection.6qn3nfmf5p1bc.ap-southeast-2.cs.amazonlightsail.com
[View this site on my domain with my container service?](#)
Private domain: team21plagiarismdetection.service.local

[Getting started](#) ×

[Deployments](#)

[Capacity](#)

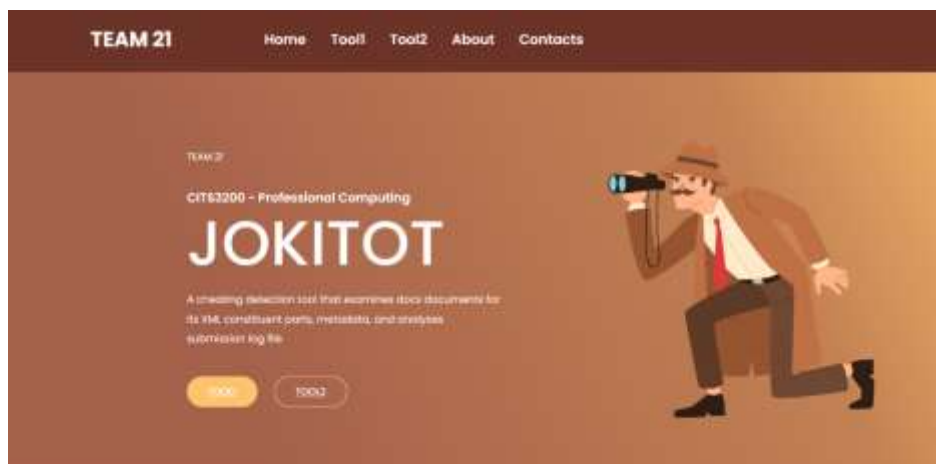
[Images](#)

[Custom domains](#)

[Metrics](#)

[Create new deployment](#)

Clicking on the link will lead you to the website.



IMPORTANT!! By having a container (either running or disabled), you will pay a monthly fee (\$15 aud/month for this one). The only way to stop paying is to delete the container, which is outlined below.

(Optional) Delete container.

To delete the container, go to the container menu on the LightSail website, click the three dots and choose “Delete”.



This is irreversible and can **only** be installed again by following the instructions.

Updating AWS LightSail

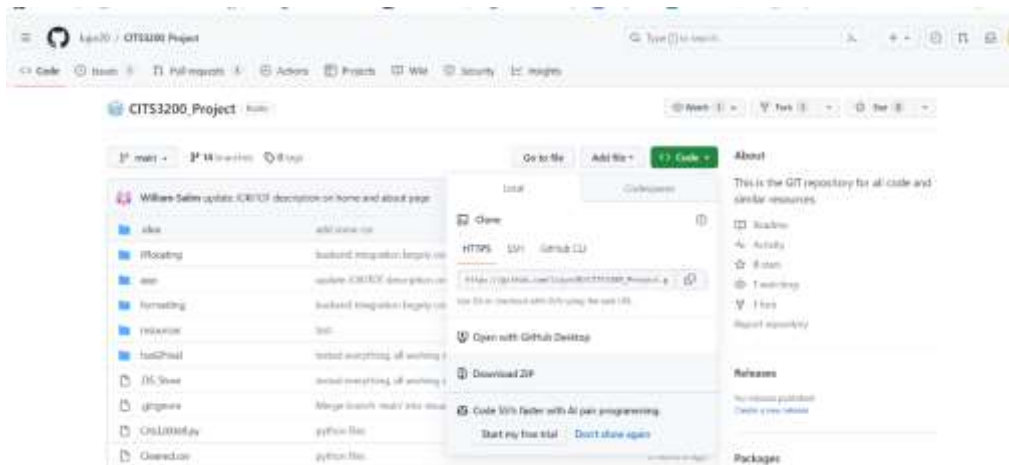
Go through these steps **ONLY** to update LightSail to a newer version of the code.

1. Download the new code from the repository

The new version of the code can be found at:

https://github.com/lujor20/CITS3200_Project/tree/main

Click Code (Green button), then download zip.



Extract the folder from the zip format and store the folder in a permanent location.

2. Follow the steps from create AWS LightSail service

The next steps are the same as the steps to set up the AWS LightSail service EXCEPT for step 4, which can be skipped. To clarify, **follow steps 1, 2, 3, 5, 6, 7, 8 and 9.**

Tool 1

File Upload

- Tool 1 allows docx document to be uploaded by clicking on the 'Browse..' button or by drag and drop the file to the specified area.
- 'Submit File' button will start the analysis of the document uploaded by the user

Interaction Menu

- User has the ability to click on a specific part of the text and change the text's corresponding colour on the Interaction Menu by clicking 'Change Colour' button
- 'Hide this' button allows user to unhighlight a specific RSID value from the result
- 'hide all' and 'unhide all' buttons allow users to remove and reapply colour coding to the output
- 'toggle collapse' button is a functionality that toggles between a formatted view and non-formatted view. A non-formatted view will remove any spaces between paragraphs, while a formatted view maintains the structure of the document.
- User also has the options to increase, decrease, and reset font size of the output.
- 'Download PDF' button allows user to download the result as a PDF for further investigation.



The 'Foundations of Group Behaviour and Work Team' lecture was delivered by Dr. Hassan Farag. This lecture highlighted the differences between a group and a team and how they function in the workplace. Understanding their differences has helped many businesses and organizations to improve efficiency and satisfaction in the workplace. I think that this lecture has taught me an invaluable lesson in both social and psychological elements that come into play in an organization.

Groups are defined as two or more individuals, interacting and interdependent, who come together to achieve particular objectives. They are formed both formally (e.g., flight crew) and informally (e.g., social hangouts). While it is natural for people to form groups, the stages of group development are not linear and in fact may regress to previous stages depending on the situation that arises. Dr. Farag explained that there are five stages of group development, which are Forming, Storming, Norming, Performing, and Adjourning. Based on my experience, I relate to this point when working in my past time job. Although the members in the group have been fully functional (Performing stage), we regressed to the 'Norming' stage due to an internal role change. This lecture made me aware of the reasons behind this and better understand the intricacies of group properties such as members' roles, norms, and status that come into effect.

While larger groups are likely to develop better ideas and solve problems faster compared to smaller groups, they are more susceptible to make such as mistakes and groupthink. Groupthink is defined as the perceived obstacles that split a group into two or more subgroups based on individual differences such as gender, race, and religious groupthink. On the other hand, the phenomenon in which the norm for consensus overrides the realistic appraisal of alternative courses of action. These risks have been proven to be detrimental to group functioning and performance. I believe that this is a crucial point to understand especially if we're to take the opportunity to be in a managerial role. To limit these risks, I will start the group size to less than 10, seek input from all members, appoint a 'devil's advocate', and discuss alternative solutions.

Dr. Farag further explained that teams have something more than groups have, a 'collective memory'. Members in a team have complementary skills, committed to a common purpose, set of performance goals, and common associations for which they

Tool 1 page – shows an example of a docx document being analysed

Multiple Documents Analysis

- Tool 1 also offers the ability to upload multiple files at once
- Statistics such as 'Characters per unique RSID' and 'Characters per run' are provided

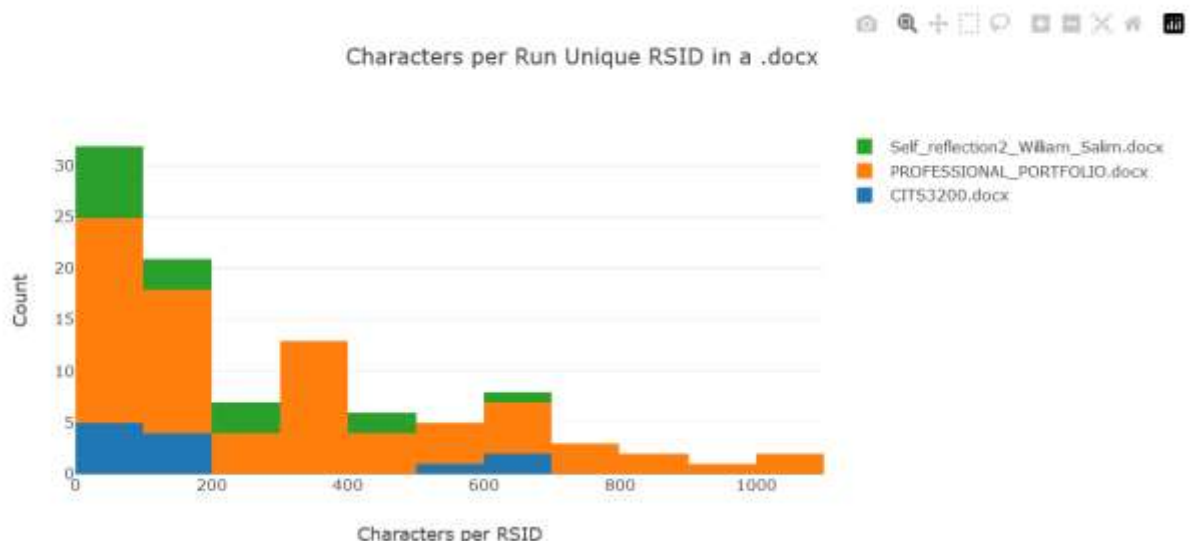
Upload New File

Tip : You can select multiple files by holding ctrl and clicking the files in the folder. Alternatively you can hold shift and click two files to select those files, and all those files inbetween.

Multiplefile No files selected.

Graph of count of char per unique RSID in a given document

Typically values above the range of 300 - 400 chars for a given .docx file warrants further investigation



Tool 2

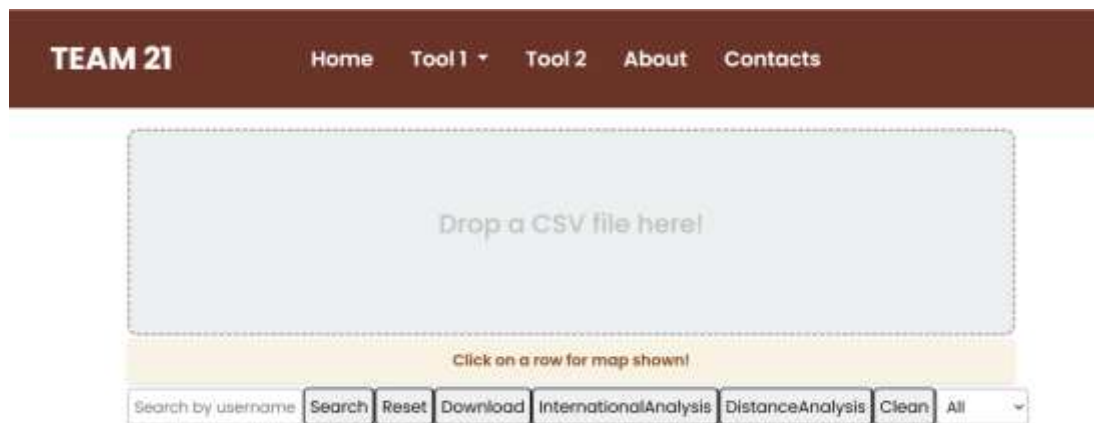
At a glance

Upload Your CSV File:

- Drag and drop your CSV file into the designated box.
- Rows will be displayed with relevant information under the box.

Choose Your Analysis:

- Click on either "International Analysis" or "Distance Analysis," or download the current data.



Preparation

- Tool 2 relies on specific keywords to proceed with its analysis, ensure the csv file has:
 - Headers for each column
 - Two headers "Last Edited by: IP Address" and "Last Edited by: Username" respectively.
- Refer to 4. Errors in your csv file doesn't satisfy the above

```
i,ip = index,row['Last Edited by: IP Address']  
i,id = index,row['Last Edited by: Username']
```

Figure 1- initialClean.py code snippet

Initial analysis

- To begin, Drag-drop the csv file into input box.
- The initial analysis filters out unnecessary data and geolocates each IP address.
- The geolocation process duration depends on the number of rows in the CSV.
- After the initial analysis, you'll see a table with country, city, longitude, latitude, and a flag indicating international/domestic.
- You can download the current data or click on a row to visualize the IP location on a map.

In depth analysis

- Choose between "International Analysis" or "Distance Analysis."
- International Analysis identifies higher-risk international countries for academic integrity breaches.
- Distance Analysis finds rows with the same user ID but different IPs, suggesting tests taken at different locations.
- These in-depth analyses are quick to perform.

Search

- Entering the id in the search box and clicking search will filter out other users.
- Click "Reset" to redisplay all ids.

Clean

- Click the "Clean" button before dropping another csv file.

Errors

- If your csv file lacks the 2 headers, or headers in general, you can replicate this by adding “,” for each non important column and “Last Edited by: IP Address” or “Last Edited by: Username” for corresponding columns. Note: when inputting, please leave out the “”
- Due to budget restrictions, tool 2 utilises a free geolocating database which limits lookups. Too many lookups within given time frame leads to it temporarily blocking its services. Generally, 1000 lookups within 1 day is safe. If your csv file has more than such rows, this can be solved by dividing csv file into ≤ 1000 and ≥ 1000 and running respective csv files on different days.
- If interface is stuck on loading, please don't worry, the geolocating process often takes around 40m for 1000 rows.