



Ollscoil Chathair  
Bhaile Átha Cliath  
Dublin City University

# eMot

## User Manual



Figure 1: eMot logo

Michael Savage - 17313526

---

Gerard Slowey - 17349433

Supervisor: Jennifer Foster

Date Completed: 01/05/2021

# Table Of Contents



<b>Table Of Contents</b>	<b>2</b>
<b>Abstract</b>	<b>3</b>
<b>Glossary</b>	<b>3</b>
Terms	3
Colours	3
<b>Overview</b>	<b>4</b>
<b>Installation</b>	<b>5</b>
Python	5
Docker	5
eMot Application	6
Install Script	6
Start Application	7
Using Docker Desktop	7
<b>User Guide</b>	<b>8</b>
Choosing Parameters	8
Browser	8
Filter	8
Starting New Instances	9
About eMot	9
Changing Preferences	10
HTML Tags	10
URL Addresses	10
Pressing the Go! Button	11
Reading the Metrics	11
Split Chart	12
Line Chart	12
Pie Chart	13
Bar Chart	13
Viewing The Wordclouds	13
Most Positive and Negative Sites	14
Most Positive and Negative Sentences	14
<b>Screenshots Of eMot</b>	<b>15</b>
Home Page	15
Processing Page	15
Preferences Page	16
About Page	16
Metrics Page 1	17
Metrics Page 2	17
Metrics Page 3	18

# Abstract

eMot is a customisable client based emotion classifier. The system uses data extraction, natural language processing, emotion classification, and machine learning to judge the sentiment of textual material being read online by you the user. The application has a simple user-friendly interface and provides graphical representations of the textual material.

## Glossary

### Terms

GUI	Graphical User Interfaces are systems of interactive visual components for computer software.
PYQT	A Python binding of the Qt C++ cross-platform framework. It is a GUI module.
Docker	A tool designed to make it easier to create, deploy, and run applications by using containers. Containers allow a developer to package up an application with all of the parts it needs, such as libraries and other dependencies, and ship it all out as one package.
VENV	VirtualENV is a tool for creating isolated Python environments containing their own copy of python, pip, and their own place to keep libraries installed from PyPI. It's designed to allow you to work on multiple projects with different dependencies at the same time on the same machine.
Headless Browser	A web browser without a GUI. eMot uses one to render websites in the background.
CMD	<p>The Command Prompt (terminal) is an interface in which you can type and execute text based commands. You can open it as follows:</p> <p>Windows - press the  <b>Win</b> key on your keyboard and type cmd.</p> <p>Mac - press  and the space bar simultaneously and type terminal.</p> <p>Linux - press <b>Ctrl+Alt+T</b> and type terminal.</p>

### Colours

<i>Filenames</i>	A script file in the eMot repository.
Commands	Text that can be executed in the CMD.
Buttons	Interactable buttons in the PYQT GUI.

# Overview

The system aims to provide users with insight into the material they have viewed online. This is achieved by collecting the users browsing history links based on their browser choice and date filter choice. The browsers supported by eMot include Chrome, Edge, Firefox, Safari, Brave, and Opera. Each URL of the browsing history is then rendered in a lightweight browser and the textual material is scraped (extracted) from the web page. Your browsing history is only collected temporarily and is never transmitted outside the machine that eMot runs on.

Each sentence in the textual material is processed and classified into an emotion such as anger, fear, joy, surprise, happiness, and sadness. The user will then be shown statistics on their browsing history in the form of pie charts, bar charts, and line charts.

The user is presented with a simple GUI that operates linearly towards the goal of showcasing these statistics. By choosing the initial two parameters (browser and time period) on the start screen, the user will only have to click next to proceed to the latter screens to view statistics. They will see eMot processing some information when it prints output to the screen.

---

**Note:** Due to the nature of the technologies used, the installation of this project may be difficult. Docker is a quite large application to install and if your PC is old you might not have the required PC specs to install it.

---

# Installation

For this project, the user will need to have Python 3.9.4 installed, Docker Desktop installed and running, and the source code downloaded and unzipped.

## Python

---

The application is written entirely in Python 3.9.4 so it should operate cross-platform on Windows, Mac, and Linux machines.

You will be required to install Python3 from the following link:

<https://www.python.org/downloads/release/python-394/>



Figure 2: python logo

```
C:\Users\micha>python --version
Python 3.9.4
```

Figure 3: python version in cmd

You can confirm that you have the right version of Python installed by opening your CMD and entering `python --version`.

---

## Docker

---

eMot scrapes web pages by using a lightweight javascript rendering service called Splash. This is a lightweight headless browser that runs in a Docker container in the background away from the user.

You will be required to install Docker Desktop by selecting the link associated with your PC's operating system:

[Windows - Docker Desktop](#)

[MacOS - Docker Desktop](#)

[Linux - Docker Desktop](#)

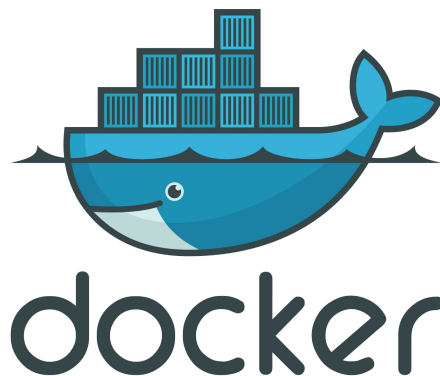


Figure 4: docker logo

```
C:\Users\micha>docker --version
Docker version 20.10.5, build 55c4c88
```

Figure 5: docker version in cmd

You can confirm that you have Docker installed by opening your CMD and entering `docker --version`.

**Note: The program will not work if Docker Desktop is not running.**

---

# eMot Application

---

After successfully installing Python3 and Docker Desktop, the eMot application can be downloaded from the following Google Drive Link by right-clicking on the emot-1.0 folder:

<https://drive.google.com/drive/folders/1NgfzMaWEXUnBIPPr17biZ5my6j-JiYig?usp=sharing>

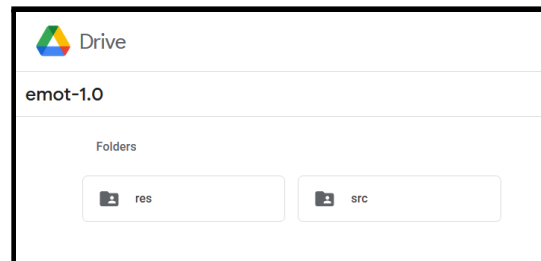


Figure 6: eMot project files

After extracting the zip files to your PC, the eMot application needs to install some dependencies. You can find the appropriate scripts for your operating system in the 'res' folder of eMot.

---

## Install Script

You can find the install scripts by double-clicking on the 'res' folder in eMot.

If you are using a Windows PC, then you must click on the **Install.bat** file located in the **Windows** folder.

If you are using a Linux PC, then you must click on the **Install.sh** file in the **Linux** folder.

If you are using a Mac PC, then you can click on the **Install.sh** file in the **Mac** folder.

By double clicking the install file in the root folder, Python will create a VENV and install all the required components into the VENV.

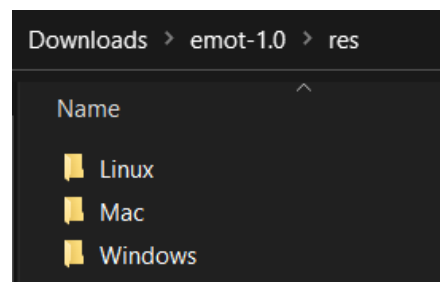


Figure 7: the contents of the res folder

---

**Note: If you prefer to use the command line, you can open the install file in a text editor and extract the commands that you need or use the following commands:**

```
cd src
python -m venv env
env\Scripts\activate
python -m pip install --upgrade pip
python -m pip install -r requirements.txt
docker pull scrapinghub/splash
```

## Start Application

---

You can find the start scripts by double-clicking the res folder in eMot.

If you are using a Windows PC, then you must click on the **Start.bat** file in the **Windows** folder.

If you are using a Linux PC, then you must click on the **Start.sh** file in the **Linux** folder.

If you are using a Mac PC, then you can click on the **Start.sh** file in the **Mac** folder.

The start file will start the main emot application window in the VENV and run the splash docker image in the background.

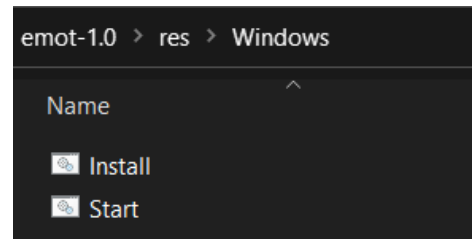


Figure 8: contents of the Windows folder

---

**Note:** If you prefer to use the command line, you can open the start file in a text editor and extract the commands that you need or use the following commands:

```
cd src
call env\Scripts\activate.bat
docker run -d --name splash -i -p 8050:8050 --rm scrapinghub/splash
python emotQT.py
```

## Using Docker Desktop

---

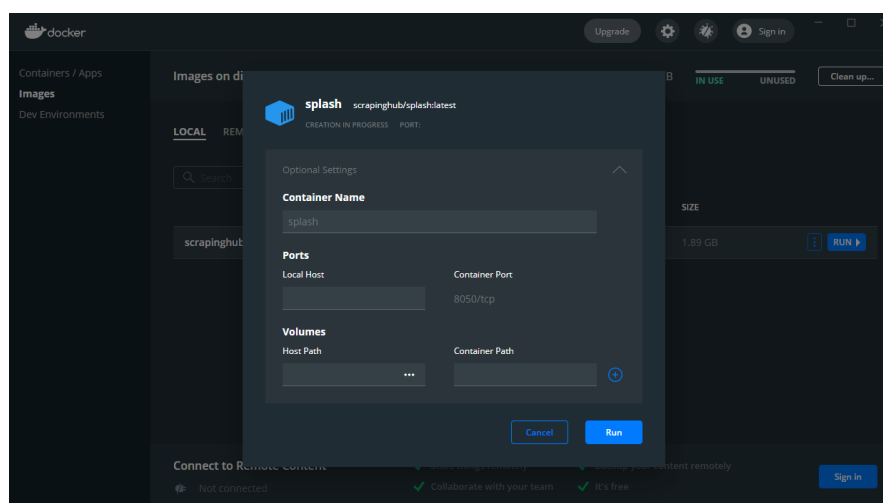


Figure 9: Docker Desktop showing scrapinghub/splash image.

---

If the docker instance does not start, you can easily use Docker Desktop to start the container. You can run the scrapinghub/splash container in the Docker GUI by going to the Images tab. Enter *splash* as the container name and click Run.

---

# User Guide

## Choosing Parameters

---

There are a large variety of browsers available to use and many people don't rely on a single browser. eMot gives the option of selecting from six of the top browsers when you land at the home page.

The options include [Chrome](#), [Edge](#), [Firefox](#), [Safari](#), [Brave](#), and [Opera](#). A browser must be selected to be able to click the [Go!](#) button.

By default, the date filter is set to [All](#). This can be changed to [Hour](#), [Day](#), [Week](#), [Month](#), and [Year](#).



Figure 10: Chrome browser



Figure 11: Safari browser



Figure 12: Edge browser



Figure 13: Brave browser



Figure 14: Firefox browser



Figure 15: Opera browser

---

## Browser

---

The browser selection must be a choice other than the default [Select Browser](#) when you click the [Go!](#) Button. Otherwise a dialog window will appear telling you that you must choose a browser from the dropdown menu.

If you choose a browser that is not installed on your PC, you will have to start the application again. This is simply done by clicking the [start again?](#) button that has replaced the [Go!](#) Button.

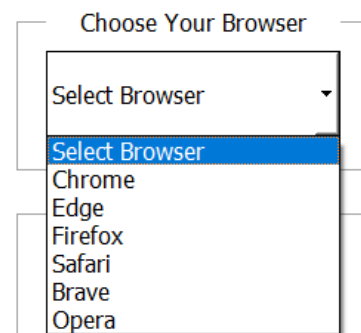


Figure 16: browser dropdown menu

---

## Filter

---

The date filter is set to [All](#) history by default. This is convenient for those who want to check every URL in their history. The more history that a user has, the longer the system will take going through each URL.

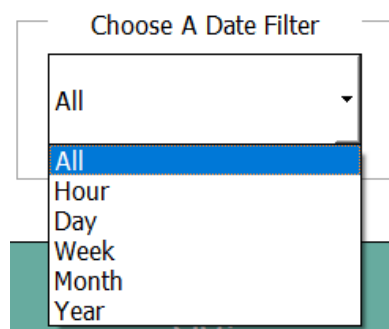


Figure 17: date filter dropdown menu

---



## Starting New Instances

---

At any point during the application, a user can start a new instance by clicking the file menu item and clicking [new](#). This will close the current window and restart the application.

In some cases, a new instance will need to be started when eMot has nothing to scrape. The message *Nothing to Scrape!* will appear in the print box in three situations.

1. The chosen browser is not installed.
2. The chosen browser is not compatible with the PC.
3. The parameters returned no history.

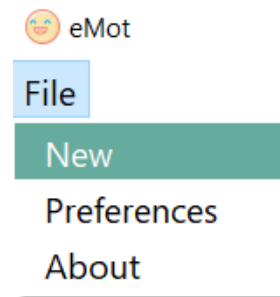


Figure 18: new instance location

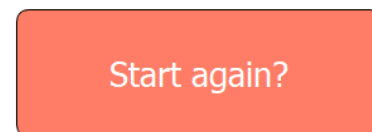


Figure 19: start again button

---

## About eMot

---

eMot has an about page that briefly explains what the application is about. It also links to this user manual if you want to find out some information.

To view the about page, click on the file menu item on the top left of the window and click on [About](#).

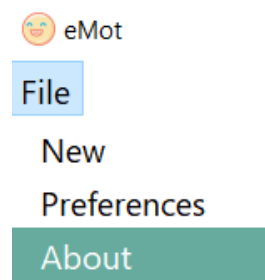


Figure 20: about eMot location

---

## Changing Preferences

---

eMot comes with some predefined preferences that can be edited. The preferences include HTML tags and URL web addresses that are blacklisted from the scraper function.

To add or remove preferences, click on the file menu item on the top left of the window and click on [preferences](#).

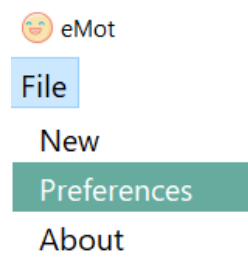


Figure 21: preferences location

## HTML Tags

---

HTML tags make up the skeleton of web pages. Excluding a tag will make the scraper ignore any textual material inside that tag.

For example, some predefined tags include *title*, *style*, and *meta*.

You can exclude tags by entering the name of the tag into the HTML blacklist prompt and clicking [add](#).

Similarly, re-including tags previously added can be deleted by filling in the text prompt and clicking [delete](#).

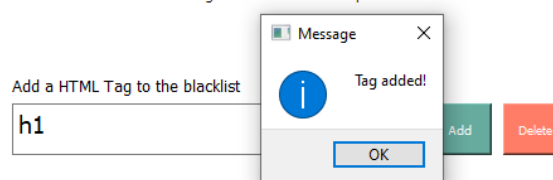


Figure 22: adding the h1 tag

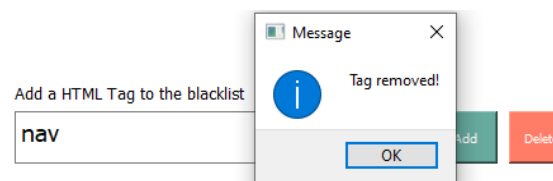


Figure 23: deleting the nav tag

## URL Addresses

---

URL addresses can be excluded to reduce the irrelevance of scraped data and prevent scraping of URLs that you don't want.

For example, some predefined URLs in the blacklist include *google.ie*, *mail.google.com*, and *gitlab.computing.dcu.ie*.

You can add more URLs to the blacklist by filling in the text prompt and clicking the [add](#) button. Leaving out *https* and *www* before the domain will not impact the blacklist.

You can delete URLs from the blacklist by filling in the text prompt and clicking the [delete](#) button.

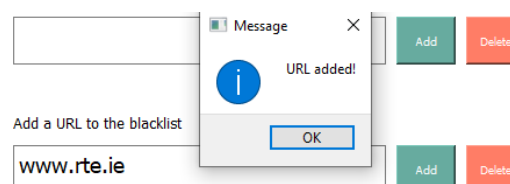


Figure 24: adding RTE to url blacklist



Figure 25: deleting reddit from url blacklist

## Pressing the Go! Button

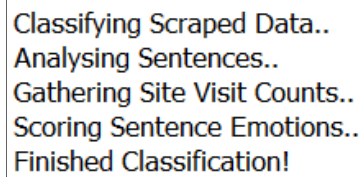
---

Choosing two successful parameters will allow eMot to start the processing stage. During this stage, text will be printed to the screen as the application progresses along. The button underneath the scrolling text will display *Scraping...* in orange until every URL has been scraped. It will then display *Classifying...*.

The screen will print information about the following:

- History being retrieved
- Sites being scraped
- Data being classified
- Sentences being scored

When eMot has finished classifying, the button will turn green and display *Show Results!*. You can click it to see some interesting graphical representations of the data.

A rectangular box with a black border containing five lines of text: "Classifying Scraped Data..", "Analysing Sentences..", "Gathering Site Visit Counts..", "Scoring Sentence Emotions..", and "Finished Classification!".

```
Classifying Scraped Data..  
Analysing Sentences..  
Gathering Site Visit Counts..  
Scoring Sentence Emotions..  
Finished Classification!
```

Figure 26: example text output

A green rectangular button with rounded corners and a thin black border, containing the text "Show Results!" in white.

Show Results!

Figure 27: show results button

---

## Reading the Metrics

When *Show Results!* is clicked, a new window will appear with three pages that can be toggled by using the *Next Page* and *Previous Page* buttons.

The first page shows a split chart of emotions, a line chart of emotion confidence levels, and a collection of positive sentence examples.

The second page shows a colourful pie chart of emotions, bar chart of visited sites, and a collection of negative sentence examples.

The final page visualizes the sentences into two generated wordclouds and links to the most positive and negative websites found. You will also be able to see your initial browser choice and filter choice on this page.

A teal rectangular button with rounded corners and a thin black border, containing the text "Next Page" in white.

Next Page

Figure 28: metrics next page button

A salmon-colored rectangular button with rounded corners and a thin black border, containing the text "Previous Page" in white.

Previous Page

Figure 29: metrics previous page button

---

## Split Chart

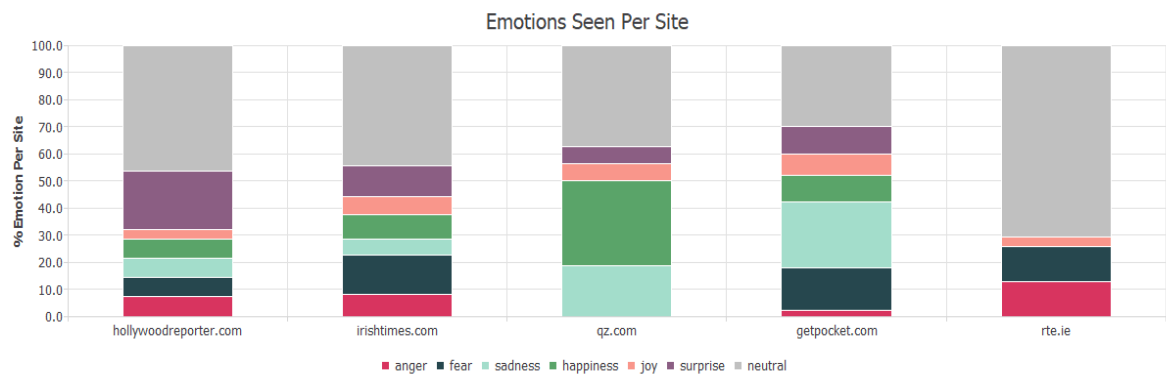


Figure 30: split bar chart of emotions per site

The split chart shows the emotions shown per site. A percentage of each of the six emotions is conveyed in a bar with the colours being represented by the six emotions.

In figure 30, we can see that the middle split line bar, qz.com, indicates that articles showed a lot of happiness. We can also see that the articles read on rte.ie exhibited angry emotions.

## Line Chart

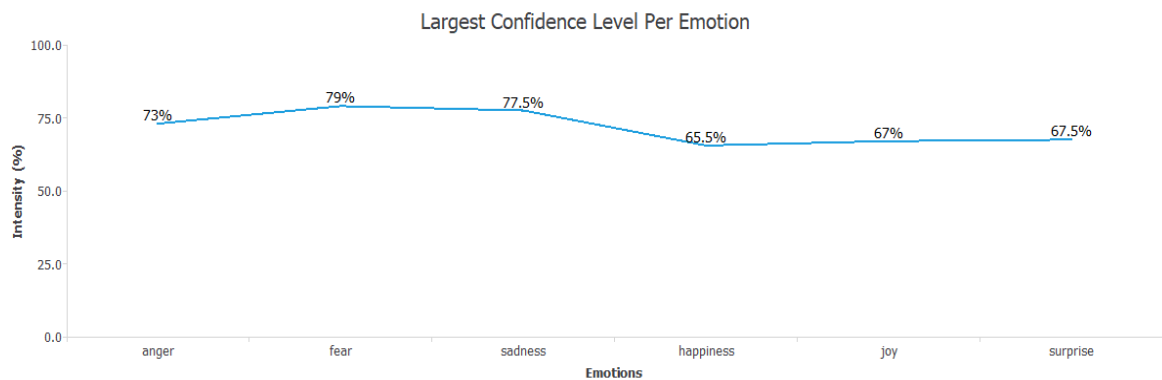


Figure 31: line chart of emotion intensities

The line chart represents the strongest intensity of each emotion. This shows how confident eMot was for choosing an emotion.

In this particular instance, we can understand that eMot was 73% sure that a sentence was an angry sentence. Similarly, eMot was 67% certain that a sentence showed joy.



## Most Positive and Negative Sites

---

Underneath the wordclouds, there are two click here buttons. You can click on these to bring you to your default browser.

The buttons mask the most positive website and most negative websites you have viewed.

Most Negative Site

[Click here](#)

Most Positive Site

[Click here](#)

Figure 35: click here buttons

---

## Most Positive and Negative Sentences

---

The metrics pages show examples of positive and negative sentences in a scrolling window.

eMot extracts ten sentences with various emotions so one screen could show two positive sentences, and the other screen could show eight negative sentences. There will not always equal emotions.

The emotion will be visible in uppercase with the corresponding sentence underneath it. In figure 36, we can observe that a joyous sentence was “smiles and nerves as lockdown lifts in northern ireland.”

### Most Positive Sentence Examples

JOY

smiles and nerves as lockdown lifts in northern ireland.

HAPPINESS

european commission president ursula von der leyen pictured last week has said that fully vaccinated american tourists will be able to visit the european union over the coming months without restrictions photograph john thys getty.

Figure 36: examples of positive sentences

### Most Negative Sentence Examples

FEAR

people arriving into the country from the us who are not fully vaccinated are required to undergo mandatory hotel quarantine because the us is one of 71 countries specified as designated states where there are higher rates of covid-19 infection or variants of concern in circulation.

Figure 37: example of a sentence with fear

---

# Screenshots Of eMot

## Home Page

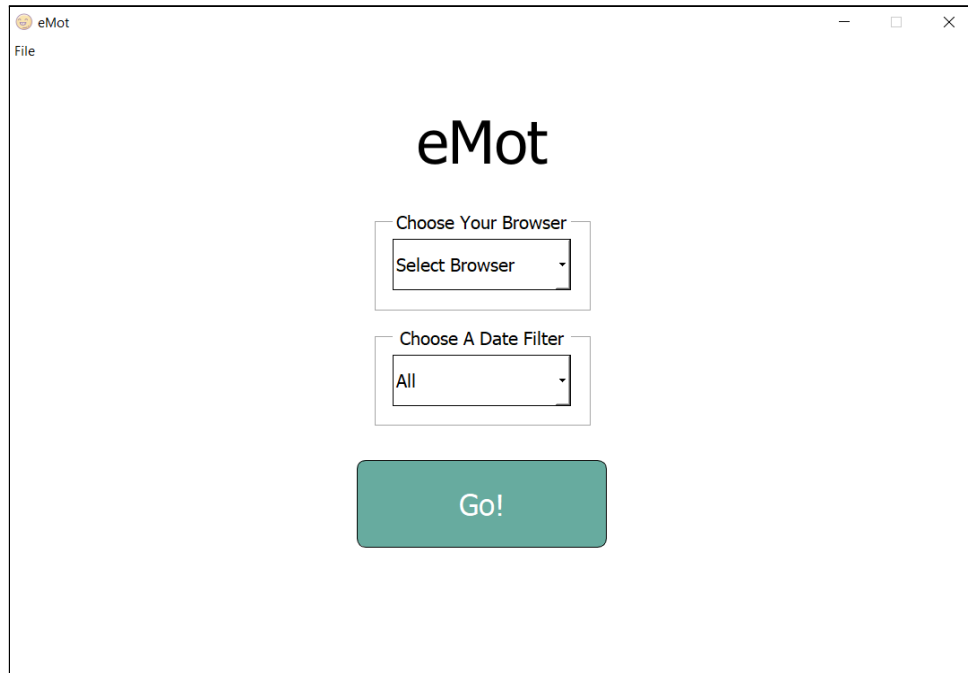


Figure 38: home page with default selections

## Processing Page

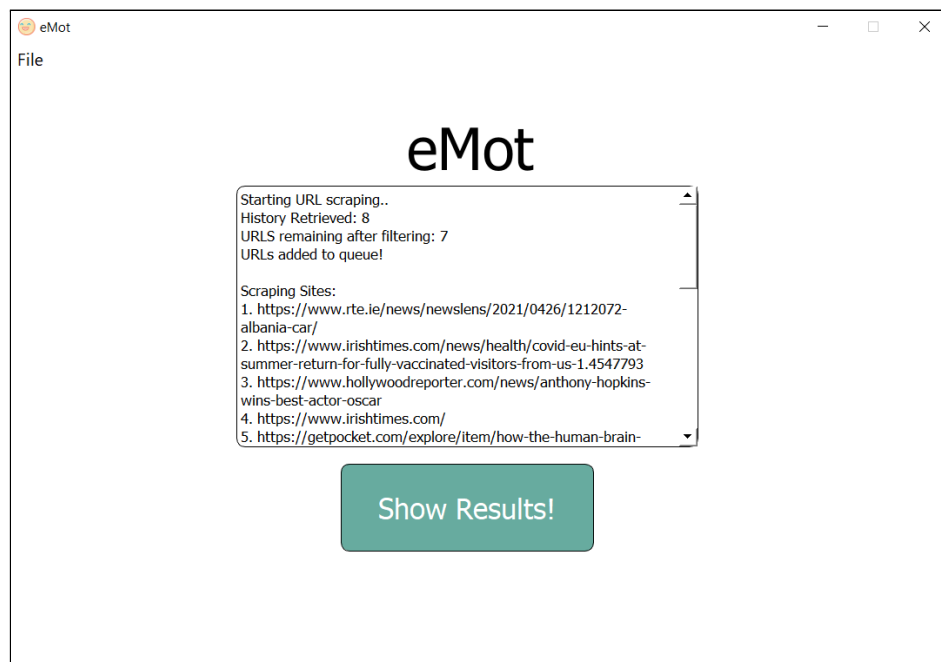
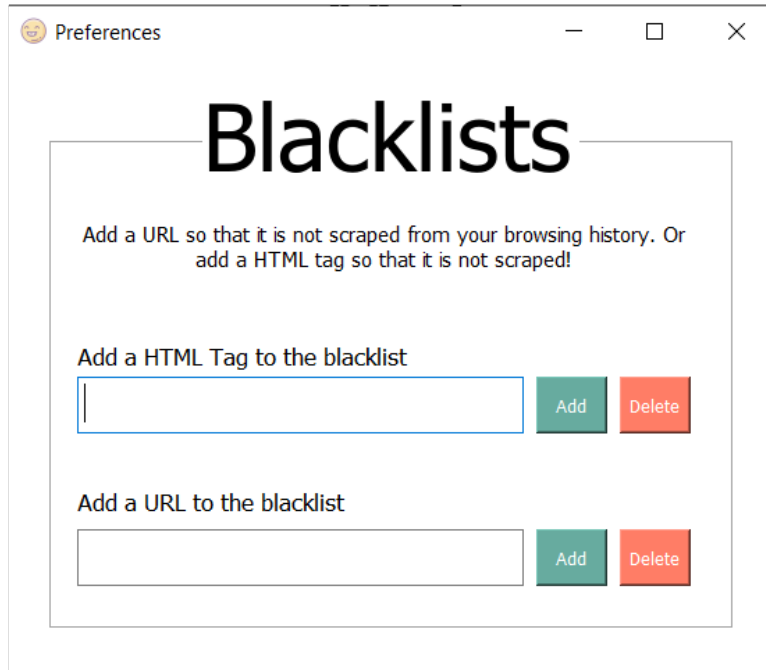


Figure 39: eMot printing progressing information

## Preferences Page



The screenshot shows a window titled "Preferences" with a close button. Inside, the "Blacklists" section has a heading and a description: "Add a URL so that it is not scraped from your browsing history. Or add a HTML tag so that it is not scraped!". Below this, there are two input sections. The first is "Add a HTML Tag to the blacklist" with a text input field, an "Add" button, and a "Delete" button. The second is "Add a URL to the blacklist" with a text input field, an "Add" button, and a "Delete" button.

Figure 40: default blacklist preference page

## About Page



The screenshot shows a window titled "About" with a close button. Inside, the "About" section has a heading and two paragraphs of text. The first paragraph explains that eMot aims to provide users with insight into the material they have viewed online by collecting their browsing history based on browser choice and date filter choice. The second paragraph explains that the text from these websites is processed and classified to give an indication of what sort of material you are reading, and that your browsing history is never stored at any point. Below the text, there is a link to the "eMot User Manual".

Figure 41: about page briefly explaining eMot



## Metrics Page 1

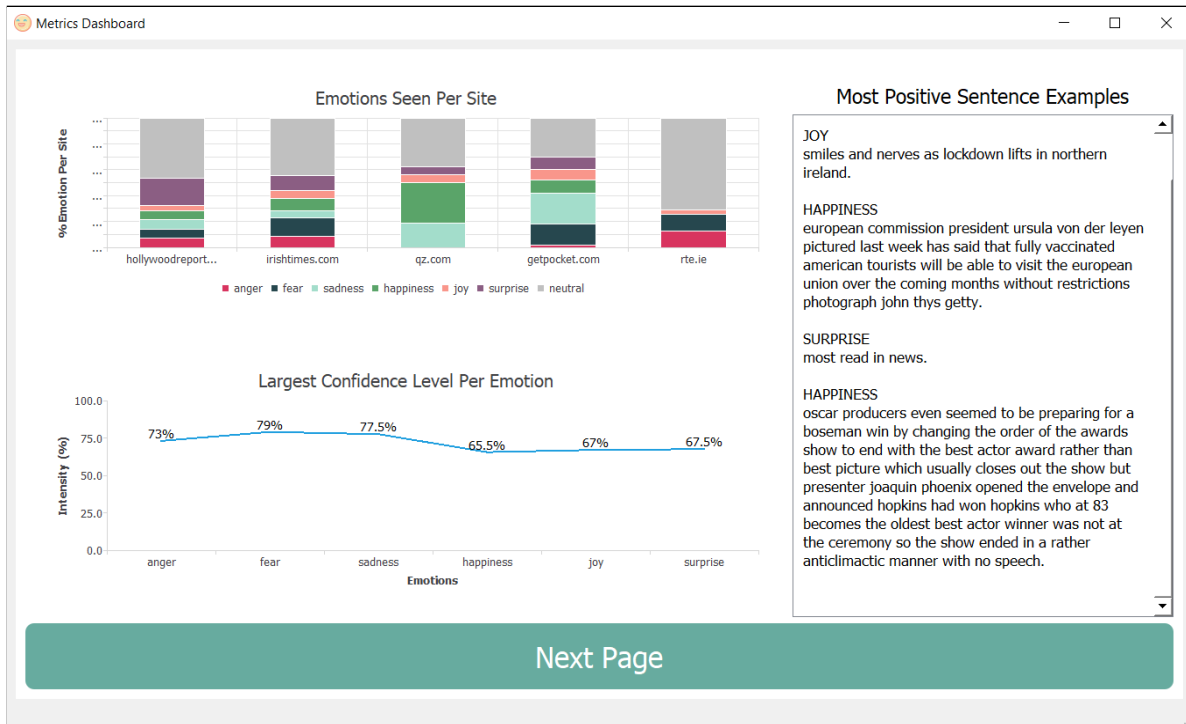


Figure 42: first metrics page showing split line chart, line chart, and positive sentences

## Metrics Page 2

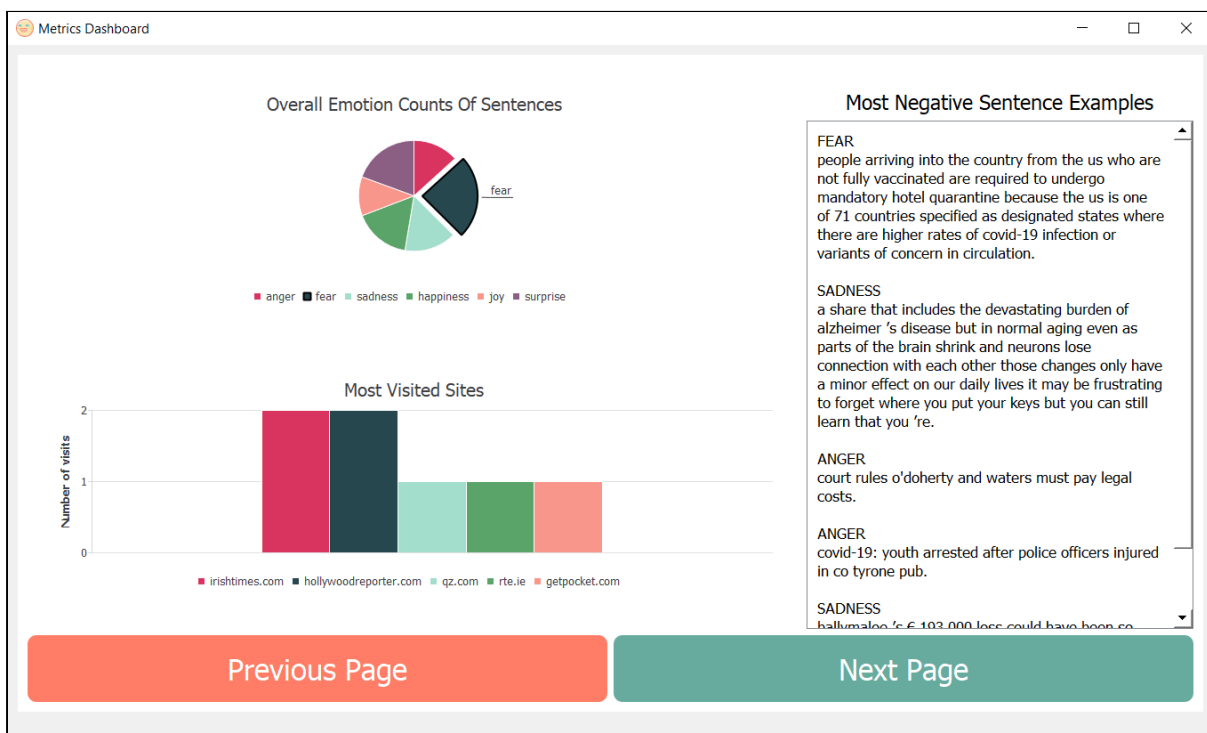
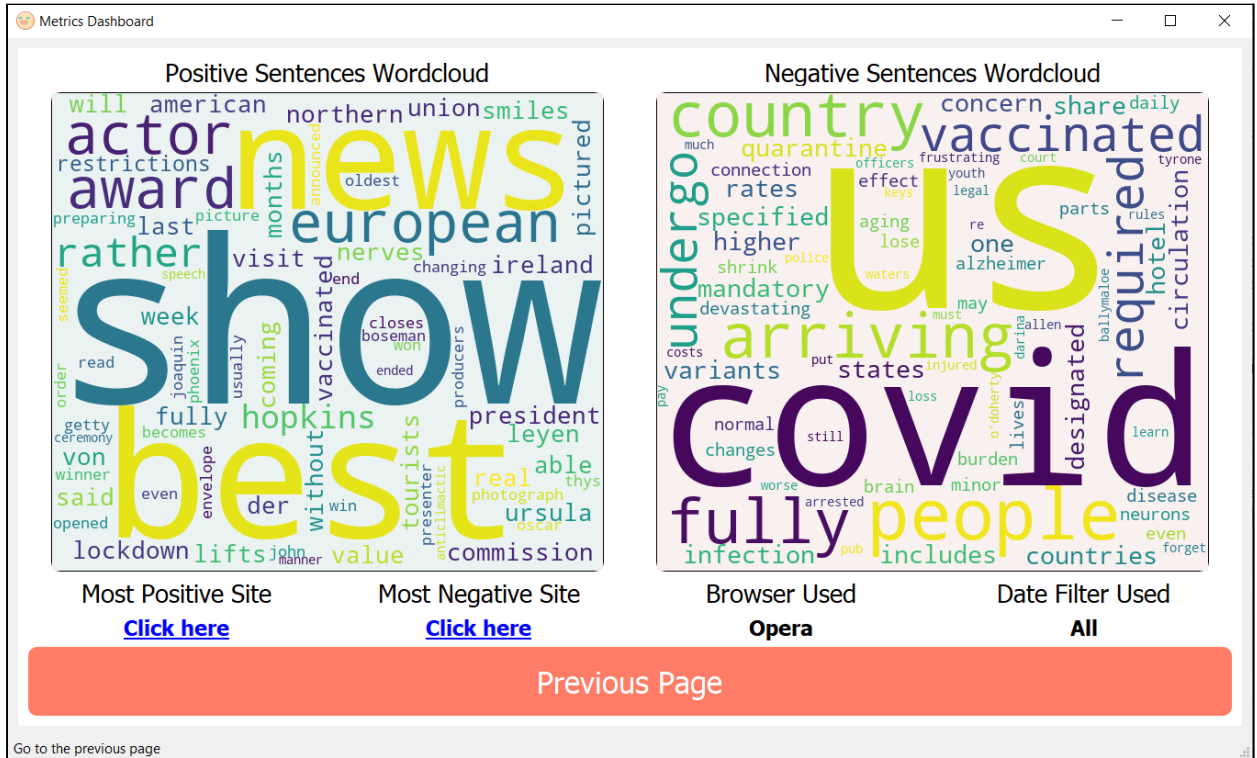


Figure 43: second metrics page showing pie chart, bar chart, and negative sentences

## Metrics Page 3



**Figure 44: third metrics page showing positive and negative wordclouds and links to websites.**