University of Venice Ca’ Foscari

SOCIAL NETWORK ANALYSIS

*Sentiment analysis of Twitter users*

MARCH 2023

Authors:

Jinpeng Zhang 886854

Samuel Takyi Oben 881431

Index

[1. Introduction 3](#_Toc129419279)

[1.1 Context 4](#_Toc129419280)

[2. Queries 5](#_Toc129419281)

[3. Data Analysis 5](#_Toc129419282)

[3.1 Data collection 5](#_Toc129419283)

[2.2 Analyzed events 6](#_Toc129419284)

[3.3 Data elaboration 6](#_Toc129419285)

[3.4 Sentiment Analysis 6](#_Toc129419286)

[3.5 Data visualization 6](#_Toc129419287)

[4. Conclusions 6](#_Toc129419288)

# 1. Introduction

This report will detail the behaviour of users on a social media platform after witnessing actions committed by members of a certain social movement regarding the climate crisis, and then observing opinions regarding it. The movement in question is fighting to disseminate a warning message to as many people as possible who are willing to listen. Recently, activity has increased due to a single event that led to the spread of similar behaviour by other members of the same movement, resulting in an increase in online searches and discussions about it.

Especially in recent times in Italy, we have heard a lot about Just Stop Oil, a movement that has been around for some time in Great Britain, where climate activists vandalize works of art in museums to attract the attention of as many people as possible. The goal is to ensure that the government commits to stopping all authorizations for the development, export, and production of fossil fuels and climate change.

Many famous painters around the world have been targeted by young activists from environmental groups such as Just Stop Oil, Extinction Rebellion, and Ultima Generazione (an Italian group). These groups of young people stage demonstrative acts in defense of the environment and our planet. They demand that science and the words of climate scientists be heard.

The first work of art targeted with a cake was Leonardo da Vinci's Mona Lisa, housed at the Louvre Museum in Paris. At the National Gallery in the British capital, activists glued themselves to John Constable's painting The Hay Wain, causing minor damage to the artwork in this case. They then targeted the Botticelli Room at the Uffizi Gallery in Florence, where they hung a banner reading "Ultima generazione, No Gas No Carbon," before gluing themselves to Botticelli's painting "La Primavera."

In most cases, these are not demonstrative acts aimed at damaging works of art, but rather to destabilize, create discomfort, and provoke reactions to talk about such a delicate issue as climate change.

Activists have realized that art, for us, represents a weak point to provoke our reaction, as they are objects of inestimable value, created by artists who lived centuries ago.

## 1.1 Context

After observing these events, caused by representatives belonging to various movements that fight for climate change and above all to spread a warning message to the whole planet, we have noticed a strong growth in research regarding these events. We will analyze how these events have affected users of a social platform through the analysis of Twitter using the API (Application Programming Interfaces, that is the programming interfaces of the applications) provided by the company to determine if they have had a positive or negative impact on the general public.

The events considered for the analysis are crimes committed in the months of October/November/December 2022 and the countries considered for the analysis are: Italy, United Kingdom, Germany, France, Spain, and the United States.

The environmental protest actually started on May 29, 2022, at the Louvre in Paris, where the Mona Lisa by Leonardo da Vinci was targeted and hit by a man dressed as an old lady with a cake. He was subsequently detained in a psychiatric hospital.

In July, a series of English museums were targeted by Just Stop Oil, including the Courtauld Gallery in London, the Kelvingrove Art Gallery in Glasgow, the Manchester Art Gallery, and the National Gallery in the capital, where young activists glued themselves to a John Constable painting, causing minor damage.

In Italy, in July, a protest took place in the Botticelli room of the Uffizi Gallery in Florence, with activists gluing themselves to the glass protecting the painting of La Primavera. At the end of the month, Ultima Generazione activists attacked the Museo del '900 in Milan and glued themselves to the structure supporting Umberto Boccioni's sculpture Forme uniche della continuità nello spazio.

In October, they targeted Vincent Van Gogh's Sunflowers, kept in London, with tomato soup. Monet was also targeted by Ultima Generazione activists at the Barberini Museum in Germany.

On October 27, Vermeer's painting Girl with a Pearl Earring exhibited in the Mauritshuis Museum was hit. Two members of Futuro vegetal glued themselves to the frames of Francisco Goya's paintings at the Madrid museum.

Fortunately, these works were protected by glass, so all the paintings affected showed minor or no damage.

# 2. Motivation of our research

Sentiment analysis regarding individuals who damage works of art is an important research for several reasons. Firstly, works of art are a cultural and historical heritage that represents a significant part of our identity and history. Damaging them is a crime that goes against our cultural and historical heritage. Secondly, sentiment analysis can provide useful information to prevent these acts of vandalism and protect works of art for future generations. For example, sentiment analysis can identify the feelings and opinions of individuals about the protection of works of art, and this information can be used to formulate more effective policies and improve public education. Additionally, sentiment analysis can help to understand the motivations behind vandalistic behavior and find solutions to prevent it. In summary, sentiment analysis regarding individuals who damage works of art is an important research to protect and preserve our cultural and historical heritage.

The motivation behind acts of vandalism that damage such important works must be equally important to them; in fact, the reason behind their actions is the spread of policies against environmental pollution and the prevention of such a crisis.

The climate crisis is a very current topic that will have an important impact on future generations but is not always at the center of discussions on online platforms, despite everyone knowing how critical the current situation is.

For psychological reasons, the problem of global warming is not often discussed even though we are aware of it.

According to psychologist Espen Stoknes, there are five reasons why the fight against climate change involves little people:

The temporal and spatial distance of the negative consequences of climate change;

Destiny, we perceive it as an inevitable thing;

Dissonance between what we do every day and what we know we should do;

Denial of the problem, which comes naturally to us when we do not want to feel responsible;

Cultural identity and political values that in people with conservative orientation lead to rejection.

# 3. Queries

After carefully analyzing the situation, we have posed some questions:

* Were these actions really necessary to attract the attention of the general public?
* Does committing crimes really have a positive impact on the masses?
* How can we help our planet for ourselves and future generations?
* How was this message perceived by users on the Twitter social platform?
* What are the feelings and opinions of individuals regarding the protection of artworks?
* What are the motivations behind vandalistic behavior?
* What policies can be implemented to protect artworks?

To answer these questions, a detailed analysis of the behavior of a large group of individuals within a platform is necessary, that is, carrying out sentimental analysis and deriving results from it. The analysis of the platform can offer important information and answers to certain questions, such as those we have previously posed. The collection and analysis of data on a platform, such as Twitter, can provide a vast amount of information on user behavior, opinions, and preferences. This information can be used to answer many questions. The analysis of the platform can provide answers to these questions and many others using techniques such as text mining, sentimental analysis, and network analysis. These techniques can be used to analyze the data collected from the platform, such as tweets, and extract relevant information. Despite these limitations, the analysis of the platform can still provide valuable information and answers to certain questions. By using rigorous and appropriate methods, researchers can obtain reliable and meaningful results that can be used to understand the attitudes of users belonging to a particular platform.

# 4. Data Analysis

Analysis is a process that involves examining and interpreting data or information to draw conclusions. The analysis process begins with collecting and organizing data, which can be in any format, in our case sentences or texts. Next, these data are processed using mathematical, statistical, or machine learning methods to identify patterns and trends. Finally, the obtained information is interpreted by creating graphical models that represent the data and relationships between them to formulate conclusions later.

The information we needed and will work with was collected from Twitter through the API provided by the company; the data will then be stored and saved locally within data frames and later converted to excel and csv files.

The collection was carried out periodically to avoid the loss of possible important information, and if redundant information is presented, it will be eliminated to avoid potential errors in the analysis process.

Since Twitter's API is very limited, collecting all necessary information was not possible, so the platform analysis will not be complete. Nevertheless, we have decided to work and process the information available to us.

## 4.1 Data collection

The data has been collected through the SNScrape library suggested by our professor; since the API given to us is not able to search for tweets older than 7 days we had to scrape old tweets.

The data we collected was not random but we followed a criteria:

1. First we needed to decide which events to analyse
2. After finding the events we had to determine the day they took place
3. Collecting tweets using specific keywords and hashtags 15 days prior to the event and 15 days after

With this kind of data collection we can take a better look at how the users were talking about the interested topic before and after the event.

## 4.2 Analyzed events

The events we decided to analyze are the acts of vandalism that we discussed in the introduction of the report, which occurred during the year 2022.

The events considered for the social media analysis include:

* Trends: Keyword and phrase trends are one of the most important factors considered during social media analysis. These trends can provide information about users' opinions and preferences regarding certain topics.
* Sentiment: Sentiment is another important factor considered during social media analysis. This factor seeks to identify whether a tweet or post contains positive, negative, or neutral emotions, helping to understand users' viewpoints on a particular topic.
* Advertising campaigns: Social media advertising campaigns can have a significant impact on users' opinions. Social media analysis can help evaluate the effectiveness of these campaigns and understand whether they are achieving their goals.
* User interactions: Social media analysis can also consider user interactions, such as likes, comments, and shares. These interactions can provide information about users' level of engagement regarding a particular topic.

The 4 main events that we’ve analysed each took place in a different country based on their official language:

1. The incident occurred at Louvre in Paris: around the month of may in 2022 a climate activist disguised as an old lady on a wheelchair managed to snuck inside the museum with a cake and threw it on the Mona Lisa.
2. The incident occurred at the national gallery of art in London: two activists of the movement “Just Stop Oil” during the month of october 2022 entered the national gallery of art in london and threw tomato soup on Van Gogh’s Sunflowers and proceeded to glue themselves to the wall.
3. The incident occurred at the Uffizi gallery in Florence: three activists of the italian movement “Ultima Generazione” (Last Generation) entered the Uffizi’s gallery during july 2022 and glue themselves on Botticelli’s Primavera.
4. The incident occurred at the Prado’s museum in Madrid: at the beginning of the month in november the environmental activist group Futuro Vegetal glue themselves to Francisco Goya’s paintings, “The Clothed Maja” and “The Naked Maja” whilst writing “+ 1.5°C” on the wall.

## 4.3 Data elaboration

For each analysed event we’ve used a set of keywords and hashtags to reduce the unnecessary data we’ve collected; avoiding data noise its quite impossible to accomplish so even after using specific keywords we had to further filter the dataframe.

The total amount of data collected is around 5.7GB which is a great amount that will need further filtering in order to avoid irrelevant data.

After creating different dataframes using keywords and hashtags we merged all the .csv files into a single one, for each event there will be a dataframe containing all the data 15 day before the event and another one containing all the data after the event.

From the file we defined different functions in python in order to elaborate the sentiment of the tweets by analysing their text with a python library called Vader that uses an algorithm to estimate the emotion of each user based on the raw content of the tweet; we decided to use Vader over TextBlob because after some research we found out that in terms of accuracy Vader is slightly more reliable than TextBlob.

**But how can we understand the emotion of a user from different tweets?**

There were 2 ways to analyse if a users changed their opinion after the event:

* Analyse only two tweets per users, which are the ones closer to the event
* Gather all the tweets for each user and determine the most relevant emotion (MRE)

We decided to rather use the second option than the first one because it’s not really reliable to base the sentiment analysis on a single tweet before and after the event.

After creating the necessary dataframes the only thing left to do is compare the same users from before and after and observe if there was a change in their behaviour on the platform and their opinion.

## 4.4 Sentiment Analysis

As mentioned in the previous paragraph the collected data has gone through a sentiment analysis.

What is a sentiment analysis?

It’s an approach to natural language processing (NLP) that identifies the emotional tone behind a body of text through machine learning algorithms, it is a very powerful tool that can help us understand and analyse the expression of the users in a community on a social media platform and eventually find patterns and trends in their behaviour.

Through the results we managed to acquire from this process we can estimate the percentage of each emotion per event and observe how the behaviour and opinion of user have changed based on the events.

What needs to be taken into account is the possibility of the sarcasm and irony from some users that the algorithm might have confused for real opinions.

## 4.5 Data visualization

//snippet

From the dataframe we created the networks, the nodes as username, edges as mentions and we then set the attributes of the nodes from the .csv file and assigned the emotions

The emotion that is most relevant is the neutral one, probably because the algorithm used (Vader and TextBlob) assign the emotion to neutral when they aren’t able to evaluate the sentiment

# 5. Conclusions