

Emoji Visualization in Twitter Feeds

FOURMOND Jérôme
LANGLOIS Aurélien

Université Claude Bernard Lyon 1

ABSTRACT

In this paper we describe a method for visualizing frequency of emoji usage in Twitter feeds in order to show the impact and the emotion of an event. Previously work, like Emoji Tracker [1] produced by Matthew Rothenberg, had realized similar work by tracking in real time all emoji on Twitter.

Keywords: visualization, data, emoji, twitter, Javascript, d3, Node JS

1 INTRODUCTION

Nowadays, social networks are a way to express freely. emoji allowed people to share a feeling, an emotion, or a wish with readers through smileys or ideograms. Emoji are ideograms and smileys used in e-mail, Web pages and mostly on social networks. Thousands of emoji are registered, from the simple smiley, to the national flag in the Unicode Organization list. [2] In 2010, the first major update to the Unicode emoji characters has been made, by including more than 1000 emoji: diversified emotion smiley (wink, sweat, open face...), animals, people and objects.

The next update to the Unicode emoji characters is scheduled to June 2017 with currently 56 new emoji.

Most computing company have developed their own fonts to display emoji, thus making harder to have a unique visualization.






Chart	Apple	Google	Twitter	Windo ws
				

Table 1. Figure 1 Same emoji from different company

Our method allows to observe the emoji frequency during a scheduled time. In our case, we aim the impact and the emotion of the French release of the movie “Rogue One: A Star Wars Story”, the 14th December 2016. In order to achieve that 3 visualizations have been realized, and to keep it simple we used only one language: Javascript, and its library d3.js [3] for manipulating documents based data and create visualization.

“Rogue One: A Star Wars Story” is an American epic space opera movie directed by Gareth Edwards. It is the first from the Star Wars standalone movie. It has been chosen because we thought it would have an important impact, as all the previous movies, on the social media, and hopefully on Twitter [4].

2 FETCHING AND FILTERING TWEETS

There no visualization without data. Tweets are our data. The tweets fetch is the first step to our project.

We used several Node JS scripts to accomplish different fetching and filtering tasks.

2.1 The Fetch

The Twitter API is used with the help of the Javascript library realized by Craig Cook, TwitterJSClient [5]. It is used as a node module. Only the search method is employed but these method is limited: for each call, 100 tweets maximum in JSON format are fetched according to keywords. It hasn't got a way to fetch a real time stream. To approach such a comportment, we call these methods each minute during approximately two days.

From the 13th December 2016 evening, to the 15th morning we fetched more than 190 000 tweets.

2.2 The Filter

Tweets fetched this way, are too bountiful. Some were retweets other were duplications. So it is needed to be filtered in order to optimize the computation and displaying time.

Thus, we first remove all the retweets and the duplications. We went from more than 190 000 tweets to 83 587 tweets, but the number is still too high, and not necessarily useful. Therefore, we kept only the tweets which contain one or more emoji.

The final count is 7 603 tweets.

3 VISUALIZATION: EMOJI CLOUD AND BAR CHART

Tweets alone are not useful for our visualization. We needed emoji data in order to cross the value. We retrieved the full emoji data of the Unicode Organization list and converted those on a csv file format.

As a first approach we used Tableau Public [6] in order to show the number of addition to the Unicode emoji character by years. [Figure 2].

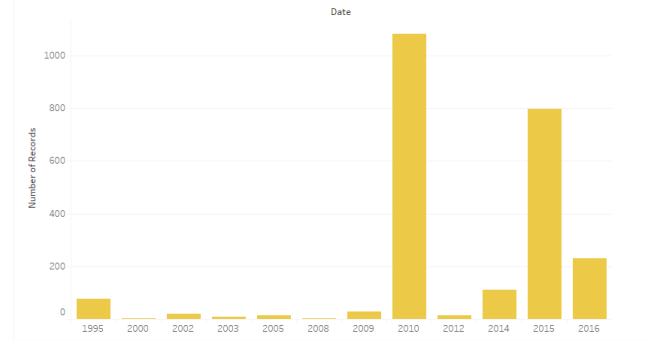


Table 2. Figure 2 Apparition of Emoji by year

On the previous figure, it is possible to identify the different major update made to the Unicode emoji character list that we previously detailed in the introduction.

Each of the visualization starts by reading tweets JSON file and emoji CSV file. Then we count the occurrences of each emoji in the 7 603 tweets, which is injected into the emoji JSON structure in memory.

3.1 Bar Chart [Figure 3]

The first realized visualization is a simple bar chart with interaction. It shows emoji occurrences in the 7 603 tweets fetched, but the visualization has been limited to emoji with more than 100 occurrences, in order to only keep the major emoji. Animations have been implemented: when the pointer goes over, rectangles which represent the number of occurrences of emoji are highlighted and emoji police and occurrences police number get bigger.

Sur 7603 Tweets uniques : 11215 emojis

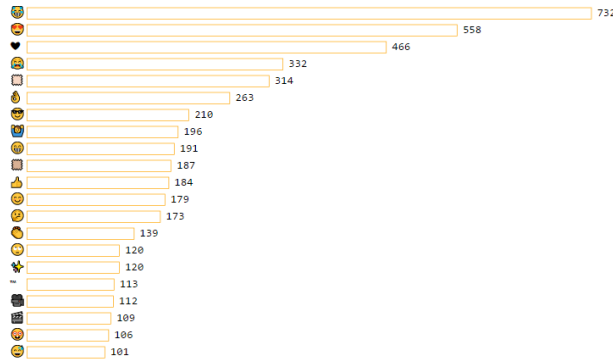


Table 3. Figure 3 Bar Chart

3.2 Bar Chart according to time [Figure 4] [7]

A second bar chart have been realized, these one according to time, in order to visualize the emoji publication from the premiere of the movie "Rogue One: A Star Wars Story" to the second day of exploitation.

Le 14 Décembre 2016 de 12h à 13h

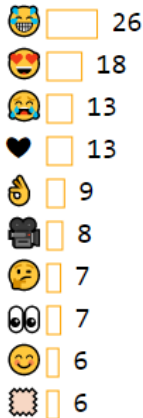


Table 4. Figure 4 Bar Chart with slider

3.3 Emoji Cloud [Figure 5] [8]

The second realized visualization is a tag cloud, renamed here as an emoji cloud. It shows all the emoji used in the 7 603 tweets, their size is proportional to their occurrences.

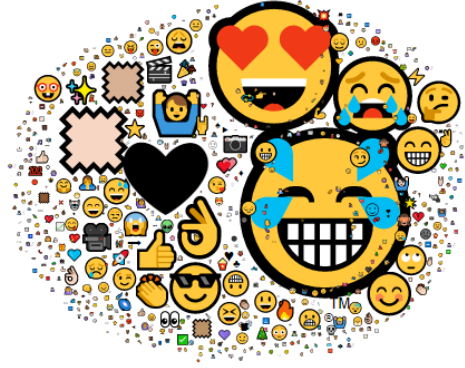


Table 5. Figure 5 Emoji Cloud

3.4 Emoji Cloud and Bar Chart [Figure 6]

Our final visualization is the coupling between the previous bar chart and emoji cloud.

When the pointer passes over an emoji from the emoji cloud, rectangles from the bar chart are highlighted and texts get bigger.

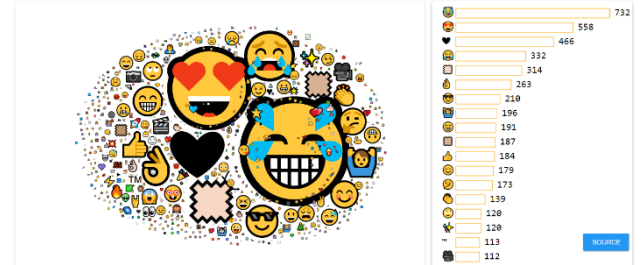


Table 6. Figure 6 Final Visualization

DISCUSSION

The visualizations may be somehow degraded due to different images from different vendors, version and source information.

The visualizations are not responsive.

On each visualization, skin tone emoji are present. While they are represented as a standalone emoji, it may be a wrong interpretation. Skin tone are mostly used in pair with other emoji as hand, or people to indicate a coloured hand or people.

More visualizations could have been useful, like, for example, a correlation matrix between the emojis, or the sentiment analysed from a tweet [9] and the emoji used in it.

Fetching in real time tweets would have been a better and simpler opportunity. This way we could have avoid to duplication in our first raw data.

4 CONCLUSION

Our method and visualization permit to see that the major reaction to the French release of the movie "Rogue One: A Star Wars Story" are "face with tears of joy", "smiling face with heart eyes", "red heart" and "loudly crying face". The first three could be analysed as positive emotion, but the last one is a little more ambiguous. It may be disappointment linked to the movie, but several other emotions.

Emoji transmit with efficiency some emotion, but some may be ambiguous without the text associate with it.

REFERENCES

- [1] <http://emojitracker.com/>, Matthew Rothenberg

- [2] <http://unicode.org/emoji/charts/full-emoji-list.html>, Unicode Organization emoji list
- [3] <https://d3js.org/>, D3 js library
- [4] <https://twitter.com/>, social network Twitter
- [5] <https://github.com/BoyCook/TwitterJSClient>, Craig Cook GitHub project: TwitterJSClient
- [6] <https://public.tableau.com/views/Emojis/Feuille1>, Jérôme FOURMOND Emoji Visualization from Tableau Public
- [7] https://jfourmond.github.io/Emoji_Viz/bar_view_time.html, Bar Chart with slider visualization
- [8] https://jfourmond.github.io/Emoji_Viz/cloud_view.html, Emoji Cloud visualization
- [9] <https://github.com/ayushoriginal/Sentiment-Analysis-Twitter>, library for sentiment analysis