Revista Mediterránea de Comunicación
Mediterranean Journal of Communication ISSN 1989-872X

Karines RODRÍGUEZ-DÍAZ
Universidad de Oriente. Cuba. karines@uo.edu.cu

Dra. Yamile HABER-GUERRA
Universidad de Oriente. Cuba. yhaber@enet.cu

Sentiment analysis on Twitter applied to Donald Trump's #impeachment

Análisis de sentimientos en Twitter aplicado al #impeachment de Donald Trump

Dates | Received: 02/03/2020 - Reviewed: 08/06/2020 - In press: 24/06/2020 - Published: 01/07/2020

Abstract
This article presents conclusions around the sentiments expressed on Twitter concerning Donald Trump's impeachment. A mixed approach, of a descriptive type, enabled us to obtain results on emotional aspects associated with users' discourse from the United States in the public sphere. The perspective employed was a transdisciplinary strategy to understand the relevant feelings of tweets in the context of their production. In this case study, data-mining methodologies and their typologies were applied: text, graph and multimedia, as well as an analysis of multimodal discourse for the extraction and evaluation of affective values concerning the trial. We took into consideration all supports, codes and practices of discursive production in microblogging: written and hypertext texts, social links and reactions, multimedia content and users' identities in a systemic relation of meanings. The publications of virtual communities in social networks about political processes are important to study in electoral situations and media events with a high degree of polarization.

Resumen
El presente artículo expone conclusiones sobre los sentimientos emitidos en Twitter con respecto al impeachment de Donald Trump. Asumimos un enfoque mixto, de tipo descriptivo, que permitió la obtención de resultados sobre aspectos emotivos asociados al discurso de usuarios de Estados Unidos en la esfera pública. La perspectiva investigativa utilizada reside en una estrategia transdisciplinar para entender los sentimientos relevantes de los tweets en el contexto de su producción. En el caso de estudio se aplican metodologías de minería de datos y sus tipologías: texto, grafo y multimedia, así como el análisis del discurso multimodal para la extracción y evaluación de valores afectivos respecto al juicio. Proponemos tener en cuenta todos los soportes, códigos y prácticas de producción discursiva en el microblogging: los textos escritos e hipertextos, enlaces sociales, reacciones, contenidos multimediales y las identidades de los usuarios en una relación sistémica de significados. Las publicaciones de las comunidades virtuales en las redes sociales sobre procesos políticos son de vital estudio en situaciones electorales y acontecimientos mediáticos de alto grado de polarización.

Keywords
Esfera pública; impeachment; comunidades virtuales; análisis de los sentimientos; análisis del discurso multimodal; minería de datos

Palabras clave
Public sphere; impeachment; virtual communities; sentiment analysis; multimodal discourse analysis; data mining

Rodríguez-Díaz, K. & Haber-Guerra, Y. (2020). Sentiment analysis on Twitter applied to Donald Trump's #impeachment. *Revista Mediterránea de Comunicación/Mediterranean Journal of Communication*, *11*(2), 199-213. https://www.doi.org/10.14198/MEDCOM2020.11.2.23

1. Introduction

Analyzing the opinions of the public sphere, and even published opinions, is a complex process that inevitably structured as a partial, but useful, view in communicative contexts. The Networks as a technological advances, do not replace the public square, they magnify the perception of democracy. However, at least, they provide a way to citizen participation of users concerning public issues. Media audiences generate significance and emotions, usually about content that has been treated by the media or institutions with greater communicative opulence. The users distribute arguments that make emotional allusion, either directly (adducing the source) or indirectly in their publications, to official news content. Hence, explains the theoretical arguments about gatewatching (Burns, 2011; Burns and Highfield, 2015).

We live in a society of networks and media, where audiences cause and reproduce their creative capacities (Martinez, 2013), in this context it is easier to study their views on event polarization politics since we agree to what constitutes a published opinion, a public discourse. We can find voters on social platforms that are instantly participating, commenting and reticulating with content that other users can react by approving, distributing, denying, and/or refuting in virtual communities [1]. Although according to Enguix (2017), "the great contribution from social networks to digital or analogue media is the undeniable increase in the dissemination of their publications."

Several authors have assessed the potential of Twitter as a space for political communication (Moya and Herrera, 2015; Pérez, Haber, Díaz and Zamora, 2017; Gómez, 2014). According to Orihuela (2011) Twitter and Social networks are giving ordinary people a chance to have a public voice. Engesser and Humpretch (2015) stablish Twitter as a platform where journalists and politicians are especially active. The microblogging service is the 13th most popular service worldwide. Moreover, it has about 340 million active users [2]. It is one of the most used in digital spaces for political and propaganda purposes, as demonstrated by Donald Trump's campaign that provide him a victory in his first period. Likewise, when Jack Dorsey, founder of Twitter, announced the restrictions concerning political advertisements on this social network, he expressed the role this has played in electoral contexts and its influence on virtual groups [3].

Nevertheless, in the analysis of these public voice of this processes, which include the construction of a collective identity based on emotions, cultural processes, identity, symbolism, social pressures and sources to which these categories have been exposed, in most cases, beyond the possible study of factual content. Thus, public opinion is composed of different valid approaches: the interests and perspectives of senders, controllers and receivers (Briones, 2018); the articulation of its discourse is crossed by three dimensions: the cognitive, the social and the systemic (Portillo, 2004). But, in a functional way, it is possible to analyze the state of public opinion in microblogging through auditable objects: the discourse, its characteristics, practices and the social relations are derived from it.

The current scientific production, associated with the extraction of emotions published by users on Twitter, focuses on sentiment analysis as an effective methodology based in text mining on big data [4]. Although in our opinion it should take into account the evaluation of emotional meaning in the multi-support contents and the social networks that symbolically represent affective issues in the speech.

In this aspect, this article sets out the use of a mixed approach based on the use of several types of data mining and multimodal discourse analysis as a methodology. The choice of such methodological framework suggests evaluating, in a generalized way, the emotional content without segmenting it to a specific support and including categories such as social relations, social influence, the links and the analysis of feelings in multimedia files. We present the results of applying these techniques and research methods in a study case on discursive production with the hashtag #impeachment published by users of the United States of America on Twitter.

1.1. Research background

"Sentiment analysis (SA) is one of the main techniques for studying data large-scale texts (big data) used in social science and communication research politics. Its objective is to recognize and evaluate the emotional value behind the texts that are analyzed, through their structure, classifying them into positive, negative or neutral. Nowadays, this methodology is mainly applied in the interpretation of texts that are spread in social media like Twitter" (Arcila, Ortega, Jiménez y Trullenque, 2017: 975) and with a focus associated to attributes. The syntactic lexicon (Vilariño et al, 2015) allows the qualification of the text.

The analysis of feelings is also known as opinion mining. Other authors have called it analysis of subjectivity, the matter has connections with affective computing, the computer recognition and the expression of emotion. As a field of research, it is a part of computational linguistics, an area that studies natural language processing and mining of texts whose object is discourse (Fornari, Abeille, Ferrero, Pérez, and Boglione, 2019).

"Unlike the manual feeling analysis with human coders or the automatic and computer assisted analysis, text mining uses procedures from an automatic supervised machine learning to generate models based on previous data and thus, to be able to predict with a significant degree of reliability the actual feeling of messages: it also allows you to run filters on queries, for examples, of dates, languages, geographic locations or labels included in the text of the message that will analysed" (Arcila et al, 2017).

The lines of research in virtual community spaces, where a large number of open-access data to obtain results on public emotions and perspectives merge, have turned in four essential aspects:

- Theoretical and methodological proposals for the use of software, algorithms [5] and approaches to cybernetic sciences for the detection of patterns focused on the feelings of users (Colle 2002, 2013, 2017; Daniel, 2010; Cohen and Hamilton 2011; Mancera and Pano, 2013a, 2013b; Diakopoulous, 2015; Coddington, 2015; Perez, Haber and Duvergel, 2016; Perez et al, 2017; Hermida and Young, 2017; Vállez and Codina, 2018; Conde, Pullaguari and Prada, 2019).

- Use of text mining (Zappavingia, 2011; Reyes, 2012; Ortega, Fonseca, Gutiérrez and Montoyo, 2013; Vilares, Alonso and Gómez, 2013; Abascal, López, and Zepeda, 2014; Almgrem and Olssom, 2015; Vilariño et al, 2015; Arcila et al, 2017, Figueira, and Guimarães, 2017; Verbeke, Berendt, d'Haenens, and Opgenhaffen, 2017; Hernandez, 2017; Van Hee, 2017; Kwabla, Kwame, and Katsriku, 2017; Martinez, 2017; Reyes, Paniagua and Sánchez, 2017; González, Hurtado and Pla, 2018; García, Henríquez and Herrera, 2019).

- Qualitative analysis for the detection of emotional affective values (Gómez, 2014; Ventura, 2016; Marín and Quintero, 2018).

- Proposal for the use of mixed tools or techniques (quantitative and qualitative) for the detection of feelings and affections (De Uribe, Pascual, and Gascón, 2016; Suau, Percastre, Palá and Pont, 2017; Bavaria, 2017; Rodríguez and Haber, 2017; Aguirre, Hernández, Briceño and Marín, 2018; Vallejo, 2018)

In the theoretical-methodological proposals, it is worthy to highligh the scientific production of Daniel (2010), who exposes methodologies from various sciences including perspectives and paradigms of qualitative, quantitative and mixed approach analysis which implies a transdisciplinary analysis for the study of virtual communities. Furthermore, Colle (2017) exposes the need for algorithms that, although they are not 100% effective, they allow you to detect patterns and qualify the context in large data sets. The use of polarity detection and opinion texts is quite effective, although it is constantly improvement based on its linguistic characteristics and still needs more effective algorithms in the detection of resources such as irony and sarcasm, say authors such as: Reyes et al (2012), Ortega et al. (2013), Hernandez (2017) and Van Hee (2017).

In text mining, symbolic techniques are used that make use of lexical resources to examine the polarity. Kwabla et al (2017) explains that through corpus linguistics and text-matching or clustering-based mining [6]. The task of identifying the feeling in a written text is a complex task even for a human being. For this reason, the automated sentiment analysis requires continuous deliberate development and refinement which has been approached from two perspectives: semantic approaches and computerized learning techniques.

However, in text mining, the filtering of feelings replies to a specific support on the written text, its characteristics and qualities, but, without taking into consideration all the symbolic constructions of the analyzed sample. The text remains the important part of the communication that we cannot forget, it is the extraction and qualification, which gives as way to specialized software that is capable of measuring polarity [7], links and reactions in large data sets. Nevertheless, given the properties of the user's posts on social media, we cannot eliminate in the context focused on the detection of affection categories such as social relations and multimedia content in expression of emotions.

The tweet is a 280 characters unit of meaning that cannot be separated from its retweets, likes, videos, images or links that are part of the content. Likewise, it should be taken into account the number of replies as reaction patterns, although these can be studied separated from the main text (or main tweet) as they come from another user. The reactions (likes and retweets) and replies within the tweet, although measurable attributes through the text filter, are the direct result of the articulation of social relations. According to Cansino et al (2016: 16): "(...) the tweet is persuasion. A tweet that nobody reads is a tweet that never existed. To transcend, a tweet has to seduce, persuade, convince, motivate... Only then, it will stand out from millions of tweets that are born and die every second (...)"

Although most of the preceding research is oriented towards text mining, there are few exceptions that have paid attention to social relations and the multimedia content. Authors such as Welbers and Opgenhaffen (2019), Wu, Hofman, Mason and Watts (2011), and Gruzd, Wellman, and Takhteyev (2011) noticed a particular importance concerning the social links in conversation in virtual communities as a determining factors in the production of sense and emotional speeches. Social links are measurable through graph mining, technique that provides tools for mapping data structures and finding new connections (relations) between objects or nodes (Chakrabarti, Papadimitriou, Modha, and Faloutsos, 2004).

Regarding the analysis of multimedia content, Conde et al (2019) makes a quantitative comparison of the production on Twitter, which mentions the multimedia of political messages published by presidential candidates in Spain, Ecuador and Colombia; paying attention to the messages in different supports in the context of microblogging: notwithstanding, the authors do not go into detail about the emotional content of the sample. In addition, Vásquez (2019) talks about a multimodal analysis in the semiotic relationship between the text of the tweet and the added image. We subscribe that this relationship is a unit of meaning.

What we denominate as emotions, feelings and opinions from the point of view of interpretation cannot be only determined through the analysis of written texts in text mining because it assumes a segmented form of the feelings expressed in the public sphere. The project has to identify and take into consideration the graph mining, multimedia supports, the identity and social influence factors; in order to determine their context. Multimedia mining is useful for categorization based on form such as: file type, duration, and size. However, it is less effective in content analysis, so it is permissible the application of qualitative methodologies of manual type to determine the effects used in supports of audio, video and interactive content.

Studying not only the written text, but including multimodal values is important since "it is not in the data structured where the largest volume of information in the world is found, but in multimedia such as images, text, audio and video" (Oviedo and Velez, 2016:127).

Image mining commonly performs a characterization by means of color histograms, texture and shape characteristics that can be extracted. Audio mining makes an advantage to extract low-level characteristics in the time and frequency domain. Finally, mining of videos is presented as the conjunction of text, image and audio features. It can be identified that multimedia mining presents dimensioning problems, as represent it, which affects the performance of analytical methods based on content. Therefore, a qualitative analysis is relevant for emotional data extraction in images, audio and video. Arnáiz and Filardo (2020) present a multimodal approach in the study of images on Twitter.

What we propose is not an overcoming of the arguments and theses of the authors who were above mentioned. We advocate the inclusion of their presentations in a more comprehensive perspective. Due to its nature, most of the researches set out in this section, and probably because of the objective that the researches were after, have segmented the analysis of feelings and opinions into a specific support. The studies associated to the discourse, which is the productive space where we find the representation of feelings, have to use multiples methodologies for the detection of patterns in their multi-supports samples. They also need complete softwares. These ones facilitate the detection of affective elements in text, graphics and multimedia files. The researcher has to present an active attitude in the review of the filters and the detection of the effectiveness of the algorithms, a human accompaniment.

Similarly, we suggest not to suppress the use of qualitative techniques, such as discourse analysis, due to the fact that they are necessary in the research of audiovisual materials; whose expression and content cannot be effectively categorized yet by data mining software. We are talking about an assimilation of the methodologies presented by the authors, all useful, qualitative and quantitative, but more efficient together in a mixed, yet complex and transdisciplinary approach. Furthermore, we propose a generalized vision that integrates several categories or codes of practice on Twitter feeds: texts (tweets, replies, links, hashtags, user identities available at the bio), graphics (social and label relations), multimedia (content and form in photo formats, audio and video) and text-graphic-multimedia relationships. A symbolic system has to be analyzed from of the parts that make up the whole of their practices of meaning and emotional production.

Therefore, in this report, we present a descriptive research on emotions published by users in virtual communities on Twitter. From a mixed approach using quantitative and qualitative tools to obtain results, we focus on the study of particular cases of #impeachment in microblogging.

1.2. Political context of #impeachment

The development of the impeachment process of Donald J. Trump, who is the president of the United States of America, began on December 18th 2019, when, following testimony hearings, the House of Representatives approved the denunciation that charged the president of obstruction of Congress and abuse of power. The first session was held on Tuesday January 21st and the last one on February 5th 2020, when the Senate acquitted and shut down the process after not using any witnesses, evidence or documentation.

The news had a great impact on the international arena due to proximity of the United States elections in November 2020 and it was among the users most followed events. This was significant during the course. The present research report delved into the emotions from their symbolic representation in the speeches. Taking into consideration the social practices carried out during Donald Trump´s trial on Twitter, we use data mining, its various typologies, and the analysis of the multimodal discourse on the tweets subtracted under the hashtag #impeachment and whose geolocation was located in the United States. The analysed publications were published between 21th and 25th January 2020. This date was selected to assess the contents and the beginning of the impeachment. Moreover, the period was also representative due to the prevalence of original tweets concerning quotations and examples replies and retweets, respectively.

Published tweets that did not come from media institutions were taken into account in order to assess the emotional state of the active public sphere that generated content on Twitter during that event. In research, data mining and multimodal discourse analysis behave as useful tools for an analysis of feelings and opinions in the context of microblogging. Likewise, form describes the significance for the public sphere of #impeachment from the perspective of the virtual community that produced meaning concerning this by following the label. Therefore, in the study, the following research questions are answered:

- What polarity did the involved users assume in the impeachment speech?

- What were the emotions with the greatest number of reactions associated with impeachment?

- What were the points of framing, considerations or states of opinion about the beginning of Donald Trump's impeachment trial?

The above research questions determined the following objectives:

- Compare the polarity of the total number of tweets published in order to measure the state of opinion of users about impeachment.

- Determine the emotions in the speech of the tweets with the highest number of reactions (retweets, likes, replies).

- Identify the frame, considerations or states of opinion given by the users of the tweets at the beginning of Donald Trump's impeachment.

2. Methodology

This article proposes the use of data mining and its various typologies (text mining, graphical mining and multimedia mining) for sentiment analysis in the context of Twitter. It also incorporates the emerging methodology of Multimodal Discourse Analysis as a qualitative tool in a multimedia discursive context. The methodological proposal presented here, structured in a case, assumes a contextualized position in the production of emotional meaning in its various codes in microblogging: multimedia content, user identities, influence of relationships and relationships between users, intertextual relationships, distribution relations, label relations, reactions, among others.

We must see the produced speech as an interface that represents a symbolic content, according to Scolari (2018) in the 9th law of interfaces: "The interface is an ecosystem that includes relations of exclusion, inclusion, convergence, divergence, substitution and extension – reduction…", which in consideration of the authors can be identified in the pattern detection by the algorithms that include categorical elements that can be studied in the above-mentioned codes. Therefore, we promote the inclusion of social relationships and the analysis of multimedia files to detect feelings without leaving any of the above categories out of the question.

In order to obtain results, data mining was applied with the Stela software, developed by the Datys company in Cuba. The extraction of the tweets was carried out from January 21st and up to 7 days later of January 25th, in order to download the reactions that the published tweets had with the hashtag #impeachment during the period as such. For the extraction of the tweets, the interface of Twitter application program (API) (REST and Streaming), a freely available resource. The REST API allows you to download and filter the message history of the last 7 days for free. The polarity analysis with the SSA-UO algorithm of the Stela software has a 0.06 % margin of error in polarity detections which is quite acceptable for the identification of qualities in large data sets, an almost impossible task manually. The Stela software processes the messages through various types of mining: text, structure (links), graphics and multimedia, which allows for widespread classification. It also has filters to detect and qualify the tweets according to characters and specific properties.

According to Baeza (2009) data mining in social networks works usefully in determining perceptions. The concept of data mining groups together computational techniques that allow the discovery of information, especially characteristics that relate in an unexpected way - or hard to discover - the values of multiple variables in a large number of records. The methods of data mining reveal this information and transform it into valuable knowledge both retrospectively (historical) as prospective (projections) or "comprehensive" (understanding what is happening). Hence, it is very important for decision making in companies, organizations and governments. Therefore, the data mining is essentially an exploration and discovery methodology.

Data mining does not eliminate human participation to solve the task completely, but simplifies the work significantly and allows an analyst who is not a statistics professional and programming to manage the process of extracting knowledge from the data. There are several methods and the existing software usually includes a battery of programs that operate from different ways and deliver different types of results, mostly accompanied by visual forms to expose the relationships that have been uncovered. The operations that can be performed are, for example: classification, estimation, prediction, relationship detection, modeling, clustering and deviation detection (Colle, 2017).

Automated algorithms play a large role in some services. Concerning this, the objective of data mining is to take advantage of the hyper abundance of information. That is why, in order to work in this technique, it is very relevant to establish specific search terms. The Stela program of Twitter data mining is particularly useful for its wide range of possibilities for extraction and filtering; although it is less factually functional in describing the content of images, audio, videos and interactive media such as surveys. Therefore, multimodal discourse analysis was required to determine the relation between written text and content in other media.

"Multimodal discourse analysis is an emerging paradigm in the field of discourse that extends the study of language per se to the study of language in combination with other resources" (O'Halloran, 2012: 75). This implies that the analysis of multimodal speech is much more open to the variety of media and the multimedia nature of discourse on digital platforms such as Twitter, which includes several semiotic resources such as image, video and reaction values added to the text in microblogging as the retweet, likes, and replies. It also took into account the fact that users will use arguments to validate the emotional approach to the beginning of the impeachment in the corpus analyzed, even if they did not do so from a classical way of argumentation as Ventura (2016) proves on Twitter.

Multimodal discourse analysis allows us to visualize the meaning of a specific event. The theory of metafunctions models and the meaning potential of semiotic resources in: -ideative meaning (i.e. our conception of the world) involves: the experiential, - representation and the way we portray the obtained experience, - the logical; - the interpersonal: the put into action of social relations and the -textual: organization in the form of texts and coherent units (Iedema, 2003). These metafunctions in microblogging are measurable from the subtracted data contextual.

We will study the textual meaning construction, understanding as texts the tweets that were emitted. The fact that they not only have a personal emotional dimension, but are the result of the sources to which users and lobbying groups resulting from social relations have been exposed. In fact, according to Foucault (1973), speeches institute, order and organize the interpretation of society, social practices, social actors and relations between them, by building of versions (symbolic representations), containing values, opinions, etc. In this sense we discover orders, rules and regularities in speech, and we recognize the frames adopted by the public sphere.

Consequently, in the context of this research we based ourselves on a global analysis and then on a fine analysis guide (Wodak, 2003) on the virtual community, understood as the total number of users who actively posted content or reacted to posted content by giving retweet, like or responding with the tag #impeachment in America.

The global analysis guide was applied to the total corpus of 13,661 tweets to determine the forms of most used speeches in the context of microblogging: tweets, retweets, replies and mentions, the comparative measurement of the polarities to which they responded, references to external content to through hypertext, the interrelations between the main hashtag and other textual categories repeated in discourse, as well as the use of resources such as video and image for the production of meaning. On the other hand, the fine analysis was applied to a sample of 100 original tweets considered the most social influence[10]. The tweets were particularly taken into account in the fine analysis because determined an increase in the reactions of exposed users or emitters and a synchronization in the emotional aspect. In order to do so, the following analysis guide was applied:

1. Polarity that was assumed in most tweets.

2. Comparison between the number of likes and the number of retweets.

3. Frames regarding impeachment that were evaluated by the most influential users social (or with a greater number of social relations in the context).

4. Use of images or videos. Assessment of the meaning of visual discourse with respect to written text in the tweet.

In the point number 2 of this guide we differentiate the reactions between the number of likes and retweets given from the point of view of meaning likes to give involves an emotional or primary acceptance. However, the retweet works more directly - as a double acceptance or distribution - because it implies a greater interrelationship, quoting a message: it makes it more strength and support within the discourse of the public sphere.

3. Analysis of the results

As part of the overall analysis, the speech forms used in the sample corpus were determined (as illustrated in Table 1). The Stela data mining tool, used to determine this, provides these results with the content analysis filter, which shows a summary of the total. In the case of study, it is representative of the use of original tweets concerning the types of retweets that contained the lowest percentage of issued. The case behaves in an anomalous way, since in other investigations the conversation is given using mostly the retweet as such, as a conclusion; which denotes a level of participation and publication of relevant content by users.

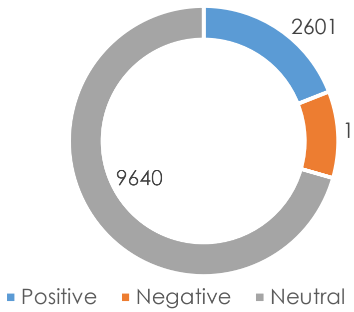
Table 1: Discourse ways used in the sample of tweets corpus

Discourse ways	Published quantity	% about total
Tweets	4485	33 %
Replies	4267	31 %
Mentions	3292	24 %
Retweets	1617	12 %

Source: Self-made based on results obtained through the Stela software.

However, the polarities of the various forms of speech behaved mostly neutral followed by positive statements with the lowest number of negative tweets. (as shown in Figure 1).

Figure 1: Comparative measure of total tweets polarity



Source: Self-made based on results obtained through the Stela software.

References to external content via hypertext are mostly links to publications of media outlets that argue the point of view that was expressed in the tweet. As stated in the Table no. 2 the hypertexts distributed mostly lead to journalistic works of CNN, Rolling Stones, Fox News or Twitter events, although other users also distributed content from these media in an indirect way without sharing the link, for instance, the case of the video where President Donald Trump says he had all the material (referring to the evidence) while "they" did not (see https://bit.ly/3ceLjUd).

Table 2: Hipertexts (urls) most shared on tweets

Hipertexts (urls) shared on the tweet	**Number of tweets that contain the hipertext (url)**
https://fxn.ws/369XIUX	23
https://twitter.com/i/events/1220068532370599936	21
https://www.cnn.com/2020/01/22/politics/trump-lawyers-impeachment-false-claims-scif/index.html	18
https://www.cnn.com/2020/01/23/politics/donald-trump-impeachment-trial-witness/index.html	18
https://twitter.com/i/events/1219696006310199296	16
https://www.rollingstone.com/politics/politics-news/trump-impeachment-evidence-we-have-all-the-material-they-dont-941140/	15
https://www.oann.com/sen-cruz-impeachment-trial-is-excuse-for-democrats-to-continue-partisan-attacks/	15
https://twitter.com/repvaldemings/status/1220017702011535364	13
https://www.foxnews.com/opinion/trumps-senate-impeachment-trial-judge-andrew-napolitano	13

Source: Self-made based on results obtained through the Stela software.

The interrelationships between the main hashtag (#impeachment) and other textual categories that are repeated in speech include the use of elements that were identified through matching filters of texts in the interrelationship with other labels and words such as trial, senate, president, and democratic. The conversation was also repeatedly addressed to the figure of Schiff (Democratic Member - of the U.S. House of Representatives and appointed Chief Prosecutor in the impeachment trial), Pelosi (Speaker of the United States House of Representatives. A member of the Democratic Party and considered one of the main faces in the signing and Trump's political trial), Alan Dershowitz and Ken Starr, both Donald Trump´s lawyers during the impeachment. In addition, President Donald Trump's user was directly mentioned (@realdonaldtrump) about 929 times during the 4 days taken as reference for this study.

The use of resources such as video and image was identified in 42% of the tweets. In most of cases, it completes the meaning of the text that was written in the 280 characters of the tweet or is announced by the text. Something similar happened with the following one with 18 684 retweets (translated into Spanish by the authors):

@tommalone1961: New video about the impeachment The Senate exists for moments like this #HoldtheLine

The tweet (above) presents and completes its meaning with the video it publishes where it expresses the importance of the Senate in cases of political confrontation such as the impeachment trial. In a way The Senate's preponderant role in decision making during the year is expressed in 2:17 min. the trial, here is how @tommalone1961 distributes the content published by @NRSC (user of The Senate Majority), and gives it a double acceptance since it subscribes through the retweet of a verified account of the conservative majority in the US Senate.

The fine analysis applied to 100 most influential tweets, and responding to the analysis guide set out in the methodology:

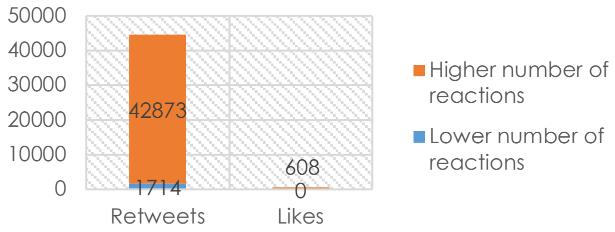
1) most of these replies to neutral polarity, only in some cases such as tweet by @tommychong, verified account (comedian, actor, director and activist in the USA), shows a negative attitude, although the content of the tweet is not one of the most influential, with only 81 retweets and 608 likes. It exposes:

@toomychong: Trump will be eliminated either by impeachment, senate trial or popular vote. No is mentally fit to be President.

He received about 20 replies, most of which refuted the argument, either against it or that in the context of the election won by Donald Trump where the popular vote did not count.

The following graph shows the reactions of the public sphere reflected in retweets and likes. We do not include the replies because the number of them in the most influential tweets was not representative. 2)In this case we conclude that the consensus in the 100 tweets evaluated in the analysis fine in the public sphere of the virtual community that followed the hashtag impeachment was mainly via retweet, that is to say via direct quotation (as we can compare in the Figure 2).

Figure 2. Higher and lower number of reactions per tweet



Source: Self-made based on results obtained through the Stela software.

3) Generally speaking, the tweets with the most considerable social influence were published by community users. The following frames were assumed by the virtual community in order to become judgments on impeachment:

They argue that the impeachment of Donald Trump will have no achievement or end. They framed their arguments and used phrases like the ones below:

- "Miserable failure," a Democratic attempt that will lead nowhere.

- No approval of witnesses or evidence.

- Republican majority in the Senate

- Impeachment is a waste of money and time for the American people

- The Democrats have no case, listening to them is a matter of grace.

- The impeachment is about how Trump will win the election in 2020

- In the trial, the whole day is spent without end, some retweets shared the fact that the senator Jim Rish was the first to fall asleep during the impeachment.

- There are conversations concerning a Democratic impotence using words against Trump unnecessarily. Schiff's impotence is assumed during the exercise.

Other widely shared perceptions of the issue evaluated Donald Trump's attitude as authoritarian and secure given the president's alleged calm and the words distributed by the media where it says that "we honestly have all the material. They don't have the material", in that sense it is worthy to highlight the influence of the tweet from @RepValDemings (Member of the House of Representatives of the America. Representative in the Tenth Congressional District of Florida) which expresses the way in which the American president covered up witnesses and documents for the American people and reinforces it with the use of video.

Other elements, although to a lesser degree of influence valued as repetition patterns by text match and retweet were:

- Donald Trump's impeachment shows divided politicians and a representative contradiction between the two parties. Relevant in the speech in such a case is the tweet with 2066 retweets it expresses: @BrandonStraka: 91-year-old Democrat: "Get rid of this impeachment charge! As Democrat I'm sick of this! #Walkaway. The tweet contains the video (retweet from @Aikens_Josh) in which a Democratic 91-year-old-politician reinforces the idea about his weariness of the process and the division between parties.

- There were discussions about CNN's argument that 51% of Americans approved to condemn Trump. Several users shared, about 398 retweets the following argument and carried out a survey on Twitter making the publication more participatory. This survey resulted in a 98% approval rating for the absolute option to Trump; although it was obtained from of 1096 votes. The tweet text, which calls for action, mentioned: CNN says 51% of voters want to convict Trump in the impeachment. I want to cast my vote because they never asked my opinion. Let's see how people on Twitter think, we're going to vote...

- One of the tweets with 2275 retweets, is an open support to the president, it is shown using the Keep America great again slogan for Trump's 2020 election campaign when he publishes: "The American people don't want impeachment, they want Donald Trump to keep America great again."

- Furthermore, as a parody of the impeachment, tweets were published saying that the event of impeachment was to get pens that Pelosi placed for those present there. It is important to notice that the tweets published for this purpose also contained the hashtag #meme.

3) Concerning the use of images and videos there is a tendency to retweet since official publications such as that of the media and other media institutions. The use of image and video media was participatory in the group from the beginning of a laughable, burlesque point of view.

4. Conclusions

When it comes to the analysis of feelings, it is relevant to consider the meaning, the significance, the state of opinion in audiences and codes in discourse, the results obtained on the basis of symbolic productions are specific to the context and time of their production. This report presents general conclusions associated with the treatment of impeachment by the public sphere of Twitter in America. In the case study most of the content that was provided through publication of original tweets, although in a second place the conversation took place (replies and mentions respectively), and in a smaller percentage distribution (retweets).

The polarity and the emotions that are mostly used in the discourse generated about impeachment are neutral. The users of the virtual community were dedicated to distributing the news from the communication and government institutions, especially indirectly, concerning the state of the matter. However, the most represented affective values, denote a doubt in the effectiveness of the trial for releasing the president of his charge and power. This question is the reflect of the frames of a failed democrat attempt, the majority of shared content samples about the absence of witnesses, documents and factual evidence in the process and the Republican majority in the Senate.

Influential users were also highlighted over the total, such as @RepValDemings, @Aikens_Josh and @NRSC who from different political positions assume the insecurity in the condemnation of Donald Trump in the political judgment in the circumstances, an argument that was most widely disseminated through retweets by members of the virtual community. In opinion mining we see a relationship of the discourses with the proximity of the U.S. elections in 2020, in addition to the allusion to the political division between parties. A generation of visual content including memes to parody the situation and the use of multimedia files was included in 42% of total tweets mainly to make sense or to prove the assertions in the written tweet. Study users' emotions about the #impeachment implies a partial vision, as it does not constitute the whole of the public forum, but if a sample active that generates content available for be audit.

5. Acknowledgement

Translator: Marianne Cándida Jústiz Tamayo.

6. Bibliographic references

[1] Abascal, R.; López, E. and Zepeda, S. (2014). Detección de patrones y grupos de sentimientos a partir del análisis de tuits políticos. *Pistas Educativas*, *35*(108), 1672-1691. https://bit.ly/2VWUAtN

[2] Aguirre, J.L.; Hernández, C.C; Briceño and Marín, A.M. (2018). Emociones e interacción ciudadana: análisis de la comunicación emocional de los presidenciables en Twitter y Facebook. En B. Gómez y J.L. López (Eds.), *Agenda sociodigital de la campaña presidencial de 2018* (pp. 108-130). Torreón, Coahuila: Universidad Autónoma de Coahuila. https://bit.ly/2KUaleO

[3] Almgren, S.M. and Olsson, T. (2015). Let's get them involved... to some extent: Analyzing online news participation. *Social media + society*, *1*(2). http://doi.org/f3m9j4

[4] Arcila, C.; Ortega, F.; Jiménez, J. and Trullenque, S. (2017). Análisis supervisado de sentimientos políticos en español: clasificación en tiempo real de tweets basada en aprendizaje automático. *El profesional de la información*, *26*(5), 973-982. https://bit.ly/2AR9N7H

[5] Arnáiz, S.M. and Filardo, L. (2020). Llamamientos feministas en Twitter: ideología, identidad colectiva y reenmarcado de símbolos en la huelga del 8M y la manifestación contra la sentencia de ´La manada´. *Dígitos*, *1*(6), 55-78. http://doi.org/dx72

[6] Baeza, R. (2009). Tendencias en minería de datos en la web. *El profesional de la información*, *18*(1), 5-10. http://doi.org/fksfc2

[7] Barthes, R. (1971). *Elementos de semiología.* Madrid: Comunicación.

[8] Baviera, T. (2017). Técnicas para el análisis de sentimiento en Twitter: aprendizaje automático Supervisado y SentiStrenght. *Dígitos*, 1(3), 33-50. https://bit.ly/3dTVqOK

[9] BBC News Mundo (31/10/2019). *Por qué Twitter prohibió toda la publicidad política en su plataforma*. https://bbc.in/2werCeC

[10] Briones, G.E. (2018). Nociones de libertad de expresión en disputa: la opinión pública publicada en la prensa ecuatoriana. *Question*, *1*(59). http://doi.org/dx73

[11] Burns, A. (2011). Gatekeeping, gatewatching. Real time feed-back: new challenges for journalism. *Braziliam Journalism Research*, *7*(11), 117-136. http://doi.org/gdxvd9

[12] Burns, A. & Highfield, T. (2015). From News Blogs to News on Twitter: Gatewatching and Collaborative News Curation. In S. Coleman & D. Freelon (Eds.), *Handbook of Digital Politics.* Edward Elgar Publishing.http://doi.org/f9sb9z

[13] Cansino C.; Calles, J.; Echeverría, M.; Rosanvallon, P.; Pasquino, G.; Morris, I.; Sánchez,J.; Aguilar,E.; Caldevilla, D.; Arditi, B.; Calvillo, J.; Ahuactzin, C. E.; Reguillo, R.; Robles, J.; Torres, L.C.; Meyer, J.A. and Islas, O. (2016). *Del Homo Videns al Homo Twitter: democracia y redes sociales.* México: Benemérita Universidad Autónoma de Puebla.

[14] Chakrabarti D.; Papadimitriou S.; Modha,D. and Faloutsos, C. (2004). Fully automatic cross-associations. *ACM SIGKDD*, 79-88. https://bit.ly/2WjowiJ

[15] Coddington, M. (2015). Clarifying journalism´s quantitative turn: A typology for evaluating data journalism, computational journalism, and computer-assisted reporting. *Digital journalism*, *3*(3), 331-348. http://doi.org/gc5t4j

[16] Cohen, S. and Hamilton, J.T. (2011). Computational journalism. *Comunications of the ACM*, *54*(10), 66-71. http://doi.org/cn4t4z

[17] Colle, R. (2002). *Explotar la información noticiosa - "Data Mining" en el Periodismo*. España: Universidad Complutense de Madrid.

[18] Colle, R. (2013). Prensa y "Big Data": El desafío de la acumulación y análisis de datos. *Revista Mediterránea de Comunicación, 4*(1). http://doi.org/dx76

[19] Colle, R. (2017). *Algoritmos, grandes datos e inteligencia en la red. Una visión Crítica*. Colección Mundo Digital de Revista Mediterránea de Comunicación. Universidad de Alicante. http://doi.org/dx77

[20] Conde, M.A.; Pullaguari, K. P. and Prada, O.A. (2019). Comparativa de la competencia mediática en comunicación política vía Twitter durante las campañas electorales de candidatos presidenciales en España, Ecuador y Colombia. *Contratexto*, (32), 41-70. http://doi.org/dx78

[21] Daniel, B. K. (2010) *Handbook of Research on Methods and Techniques for studying virtual communities: Paradigms and Phenomena*. IGI Global. http://doi.org/cnsjmh

[22] De Uribe, C.; Pascual, J.M.; and Gascón, J.F.F (2016). El liderazgo político en la era digital. Un análisis de uso de Twitter por el papa Francisco (@Pontifex). *Obra digital: revista de comunicación*, (11), 85-99. http://doi.org/dx8v

[23] Diakopoulos, N. (2015). Algorithmic accountability: Journalistic investigation of computational power structures. *Digital Journalism*, *3*(3), 398-415. http://doi.org/gc5t4g

[24] Eco, U. (2000). *Tratado de semiótica general,* Barcelona: Lumen.

[25] Engesser, S. and Humprecht, E. (2015). Frecuency or skillfulnes. *Journalism studies*, *16*(4), 513-529. http://doi.org/ggb592

[26] Enguix, S. (2017). Impacto político e informativo de las redes sociales: esferas de actuación y comparación con los medios. *Anàlisi: quaderns de comunicació i cultura*, (56), 71-85. http://doi.org/dgm8

[27] Figueira, Á. and Guimarães, N. (2017). Detecting Journalistic Relevance on Social Media: A two-case study using automatic surrogate features. In *Proceedings of the 2017 IEEE/ACM International Conference on Advances in Social Networks Analysis and Mining 2017* (pp. 1136-1139). http://doi.org/dx79

[28] Fornari, J.; Abeille, C.; Ferrero, M.; Pérez, G. and Boglione, S. (2019). El potencial de la minería de opinión para la gestión de la reputación online corporativa. *Braziliam Journal of Development*, *5*(10), 22310-22320. http://doi.org/dx8b

[29] Foucault, M. (1973). *El orden del discurso*. Buenos Aires, Argentina: Tusquets editores.

[30] García, J.G.; Henríquez, P. and Herrera, J. (2019). Analítica de Twitter para el estudio de las emociones primarias durante el terremoto de México 2017. *Revista Ibérica de Sistemas e Tecnologías de Informção*, (E19), 479-492. https://bit.ly/3fc1z9c

[31] Gomes, F. (2013). *El uso periodístico de las redes sociales por los cibermedios: análisis de perfiles y estudios de caso*. Universidad Rey Juan Carlos, Móstoles, España. https://bit.ly/3cRZOfy

[32] Gómez, M.E. (2014). *Manual imperfecto para políticos en tiempo irreal*. Universidad de La Habana, Cuba.

[33] González, J.A; Hurtado, Ll. F. and Pla, F. (2018). ELiRF-UPV en TASS 2018: Análisis de Sentimientos en Twitter basado en Aprendizaje Profundo. *Proceedings of TASS*, 37-44.

[34] Gruzd, A.; Wellman, B. and Takhteyev, Y. (2011). Imagining Twitter as an imagined community. *American Behavioral Scientist*, *55*(10), 1294-1318. http://doi.org/dhbc34

[35] Hermida, A. and Young, M. L. (2017). Finding the data unicorn: A hierarchy of hybridity in data and computational journalism. *Digital Journalism*, *5*(2), 159-176. http://doi.org/dx8c

[36] Hernández, D. I. (2017). *Irony and sarcasm detection in Twitter: The role of affective content*. Universidad de Valencia, España. https://bit.ly/2MO5Mno

[37] Hootsuite and We are Social. (2020). *Informe 2020 sobre uso de redes sociales*. https://bit.ly/32FJKKI

[38] Iedema, R. (2003). Multimodality, Resemioticization: Extending the Analysis of Discourse as a Multisemiotic Practice. *Visual Communication*, *2*(1), 29-57. http://doi.org/cgtwwk

[39] Kwabla, S.; Kwame, N. and Katsriku, F. (2017). Sentiment Analysis of Twitter Feeds using Machine Learning, Effect of Feature Hash Bit Size. *Communications on Applied Electronics (CAE)*, *6*(9), 16-21. http://doi.org/dx8d

[40] Leavitt, A.; Burchard, E.; Fisher, D. and Gilbert, S. (2009). The Influentials: New Approaches for Analyzing Influence on Twitter. *Web ecology Project, 4*(2), 1-18. https://bit.ly/3c4zlXq

[41] Lotman, I. (1996). *La semiosfera*. Madrid: Ediciones Cátedra.

[42] Mancera, A. and Pano, A. (2013a). *El discurso político en Twitter*. Madrid: Editorial Anthropos.

[43] Mancera, A. and Pano, A. (2013 b). Nuevas dinámicas discursivas en la comunicación política en Twitter. *Círculo de Lingüística Aplicada a la Comunicación*, *56*, 53-80. http://doi.org/dx8f

[44] Marín, A.F and Quintero, J.M. (2018). Confianza en el proceso de paz en Colombia en Twitter. *Revista Mexicana de Sociología*, *80*(1), 115-137. https://bit.ly/2zpz3Se

[45] Martínez, F. (2013). *Los nuevos medios y el periodismo de medios sociales. Análisis y aplicación de las redes sociales en los medios digitales españoles*. Universidad Complutense de Madrid, España. https://bit.ly/2MLtArN

[46] Martínez, J.J. (2017). *Minería de opiniones mediante análisis de sentimientos y extracción de conceptos en Twitter*. Universidad Complutense de Madrid, España. https://bit.ly/2UtlWqc

[47] Moya, M. and Herrera, S. (2015) Cómo puede contribuir Twitter a una comunicación política más avanzada. *Arbor*, *191*(774). http://doi.org/dx8g

[48] O´Halloran, K.L (2012). Análisis del discurso multimodal. *ALED*, *12*(1), 75- 97. http://doi.org/dx8h

[49] Orihuela, J. L. (23/06/2011). Los trend topics se han convertido en una agenda social. *La Vanguardia*. (C. Buades, Entrevistador). https://bit.ly/2PCWso2

[50] Ortega, A.; Gutiérrez, E. and Montoyo, A. (2013). SSA-UO: Unsupervised Twitter Sentiment Analysis Second Joint Conference on Lexical and Computational Semantics (*SEM). *Seventh International Workshop on Semantic Evaluation*, (2), 501–507. https://bit.ly/2SrWHUw

[51] Oviedo, E.A.; Oviedo A. I. and Velez, G.L. (2016). Minería multimedia: hacia la construcción de una metodología y una herramienta de analítica de datos no estructurado. *Revista Ingenierías, 16*(31), 125-142. http://doi.org/dx8j

[52] Pérez, E.; Haber, Y. and Duvergel, M. (2016). Herramientas para el estudio del Nuevo periodismo electrónico o de medios sociales. La perspectiva de framing y análisis crítico discursivo de las etiquetas #The5Free, #Los5Libres y #TodosSomosAmericanos. *Estudios sobre el mensaje periodístico*, *22*(2). http://doi.org/dx8k

[53] Pérez, E.; Haber, Y.; Díaz, J. and Zamora, L. (2017). Un modelo periodístico. Interrelación teórica del Nuevo periodismo electrónico y la plataforma Twitter, en el escenario mediático latinoamericano. *Perspectivas de la Comunicación*, *10*(1), 157-186. https://bit.ly/3aYBIyW

[54] Pierce, C. S. (1974). *La ciencia de la semiótica*. Buenos Aires: Nueva Visión.

[55] Portillo, M. (2004). *Culturas juveniles y cultura política: la construcción de la opinión pública de los jóvenes de la Ciudad de México*. Universidad Autónoma de Barcelona, España. https://bit.ly/2YoKjqi

[56] Ramírez, A.C. (2017*). Polaridad en redes sociales y principales indicadores financieros del sector asegurador*. Escuela Colombiana de Ingeniería Julio Garavito, Bogotá, Colombia. https://bit.ly/3hcrEXj

[57] Reyes, A. (2012). *Linguistic-based Patterns for Figurative Language Processing: The Case of Humor Recognition and Irony Detection*. Universidad de Valencia, España. https://bit.ly/3dRDZ1c

[58] Reyes, J.A.; Paniagua, F. and Sánchez, L. (2017). Minería de opiniones centrada en tópicos usando textos cortos en español. *Research in Computing Science*, *134*, 151-162. http://doi.org/dx8m

[59] Rheingold, H. (1996). *La comunidad virtual*. Barcelona: Gedisa.

[60] Rodríguez, K. and Haber, Y. (2017). La influencia social de los medios de comunicación en Twitter. *Enunciación, 22*(1), 97-108. http://doi.org/dx8n

[61] Saussure, F. (1961). *Curso de Lingüística General*. Buenos Aires: Losada.

[62] Scolari, C. A. (2018). *Las leyes de la interfaz: Diseño, ecología, evolución, tecnología.* Barcelona: Gedisa.

[63] Suau, G; Percastre, S.J.; Palá. G. and Pont, C. (2017). Análisis de la comunicación de emergencias en Twitter. El caso del ébola en España. En *Uso y aplicación de las redes sociales en el mundo audiovisual y publicitaria* (pp. 119-130). McGrawhill Education. https://bit.ly/2SvpPtN

[64] Vallejo, A. M. (2018). Comunicar emociones en el discurso metapolítico de Twitter: el caso de #Maduro versus @NICOLASMADURO. *Observatorio (OBS)*, *12*(3), 175-194. http://doi.org/dx8p

[65] Vállez, M., and Codina, L. (2018). Periodismo computacional: evolución, casos y herramientas. *El profesional de la información (EPI)*, *27*(4), 759-768. http://doi.org/dx8q

[66] Van Hee, C. (2017). *Can machines sense irony?* Universidad de Gante, Bélgica.

[67] Vásquez, L.C. (2019). Semiosis entre palabras e imágenes en Twitter. ¿Cómo construyen significado los candidatos presidenciales de las elecciones 2017 de Chile? *Galaxia*, 42, 40-56. http://doi.org/dx8r

[68] Ventura, A.S. (2016). Argumentación y discurso político en Twitter. Análisis de la campaña presidencial argentina 2011. *Aposta: Revista de Ciencias Sociales*, *69*, 39-66. https://bit.ly/2xwQ4Je

[69] Verbeke, M.; Berendt, B.; D'Haenens, L. and Opgenhaffen, M. (2017). Critical news reading with Twitter? Exploring data-mining practices and their impact on societal discourse. *Communications, 42*(2), 127-149. http://doi.org/dx8s

[70] Verón, E. (1987*). La semiosis social.* Buenos Aires: Gedisa.

[71] Vilares, D.; Alonso, M.A. and Gómez, C. (2013). Una aproximación supervisada para la minería de opiniones sobre tweets en español en base a conocimiento lingüístico. *Procesamiento del Lenguaje natural*, 51, 127-134. https://bit.ly/3c1Et3Y

[72] Vilariño, D.; Zepeda, C.; Sanzón, Y.M.; Carballido, J.L.; Medina, C. and Flores, G. (2015) Aplicación del modelo Léxico-Sintáctico para detectar la polaridad de opiniones sobre profesores. *Pistas Educativas*, *36*(113), 228-239. https://bit.ly/2KUtJs1

[73] Welbers, K. and Opgenhaffen, M. (2019). News through a social media filter. Different perspectives on immigration in news on website and social media formats. *Journalism and Traslation in the Era of Convergence*, 85-105. http://doi.org/dnvh

[74] Wodak, R. and Meyer, M. (2003). *Método de análisis crítico del discurso*. Barcelona: Gedisa.

[75] Wu, S.; Hofman, J.; Mason, W. and Watts, D. (2011). Who Says What to Whom on Twitter. In *Proceedings of the 20th International Conference on World Wide Web* (pp. 705-714). http://doi.org/cxcjbn

[76] Zappavigna, M. (2011). Ambient affiliation: A linguistic Perspective on Twitter. *New Media & Society*, *13*(5), 788-806. http://doi.org/dx3px2

Notes

1. Virtual communities: Social groups formed through computers (Rheingold, 1996: 15). The Virtual communities are social aggregates that emerge from the Web when enough people lead to. These public discussions were held for a long time enough, with enough human feeling to create networks of personal relationships in cyberspace. Virtual communities are the result of the natural sociability characteristic of human beings. When they find the right site for the creation of links through the Internet, use communication mediated by computer to carry out social aggregates cohesive by a relationship of trust based on interests common, feeling of identification with others and constant exchange of experiences (Gomes, 2013).

2. Data obtained through the 2020 report provided by the Hootsuite and We are social platforms on the use of social networks at a global level. Available at: https://bit.ly/32FJKKI

3. BBC News World. (31th October 2019). Why Twitter banned all political advertising on its platform. Available at: https://bbc.in/2werCeC

4. Big Data: A data set or combination of data sets whose size (volume), complexity (variability) and growth rate (speed), make it difficult to capture, manage, process or analyze them by conventional technologies and tools, such as conventional relational databases and statistics or display packages, within the time necessary to make them useful. Their complex nature is part of the great amount of unstructured data generated by modern technologies such as social networks. What makes it useful to big data analysis is that it provides benchmarks in the gathering of large amounts of data that are filtered and qualified quickly and efficiently.

5. Algorithm: An algorithm (from Greek and Latin, dixit algorithmus and this in turn from the Persian mathematician Al-Juarismi) is a prescribed set of well-defined, ordered and finite instructions or rules that allow an activity to be carried out throughout successive steps that do not generate doubts to whoever must develop such activity. Given an initial state and an input, following steps is reached and a final solution is obtained (Colle, 2017: 6). Usually a data mining program uses several algorithms to filter information and give results about a sample given.

6. Clustering: is a data mining technique that automatically identifies groupings according to a similarity between them (concordance). This is called pattern finding or clustering techniques.

7. Polarity: It is a value assigned to a term that expresses opinion depending on the linguistic meaning of the word and it is based on sentiment analysis or opinion mining methodologies. The value of polarity can vary between different ranges including negative, neutral and positive. (Ramirez, 2017: 47)

8. A network is a group unit that relates nodes connected by edges. In the particular case it refers to the social relations that take place between the users of the virtual community and their ways of linking.

9. The analysis of symbolic representation is based on the interpretation of elements that generate meaning. The manifestation of this sense is limited to a dimension of significance associated with a socio-cultural dimension. See Saussure, (1961); Barthes, (1971); Pierce, (1974); Verón, (1987); Lotman, (1996); Eco, (2000).

10. The social influence is valued as the reactions of retweet, likes and replies of certain content or user about others in the group. That is, the potential of a user's action (e.g., tweet) to initiate future action (by e.g. retweet or reply) from others on the network, which is evidenced by the action of creating a message (tweet) and encourage its subsequent re-tweeting through its network of contacts and the networks of those who receive it and decide to give it currency (Leavitt, Burchard, Fisher and Gilbert, 2009: 5). Two types of influence persist on Twitter: one based on conversation and another based on content.