Social networks and food crises: the case of red meat and its carcinogenic effects according to WHO

Redes sociales y crisis alimentarias: el caso de la carne roja y sus efectos cancerígenos según la OMS

Joan Francesc Fondevila Gascón. Universitat Pompeu Fabra (joanfrancesc.fondevila@upf.edu)

Javier Sierra Sánchez. Universidad Camilo José Cela (javier.sierra@ucjc.edu)

Sheila Liberal Ormaechea. Universidad Francisco de Vitoria (s.liberal.prof@ufv.es)

Received: 11/12/2016 - Accepted el: 04/03/2017

Abstract:

The impact of comments on social networks in relation to food crises is an object of recent study, whose effects must be followed. This research analyzes the case of red meat and its presumed carcinogenic effects with the objective of determining the follow-up on Facebook and Twitter and the nature of the comments (positive, neutral or negative). The International Agency for Cancer Research (WHO dependent) submitted on 26 October 2015 a report evaluating the carcinogenicity of meat consumption by including it in group 2A. This news was present in social networks, mainly on Facebook and Twitter, the most used in Spain at that time. Therefore, the qualitative methodology of sentiment analysis in these social networks and the publications related to the news is used. Publications were searched from October 26 to November 26, 2015 and specific hashtags ("red meat" and "WHO") were used to focus the search. The publications were classified as positive, neutral or negative. The reactions of the other users in each publication were also evaluated by means of a quantitative statistical analysis. We analyzed 391 publications on Twitter and 33 on Facebook. In the first social network, publications were found throughout the time frame; on Facebook, only half the time. 57.6% of Facebook publications have a neutral intentionality, 27.3% negative and 15.1% positive. On Twitter, 47.6% are neutral, 39.6% negative and 12.8% positive. The follow-up of other users to the publications studied was not statistically significant. User appreciation changes over time, going from negative and neutral to neutral and positive comments until only objective information is left. There is also a lack of communication between the official entities and the users of these networks.

Key words:

Social networking, sentiment analysis, communication, WHO, red meat.

Resumen:

El impacto de los comentarios en redes sociales con relación a crisis alimentarias es un objeto de estudio reciente, cuyos efectos hay que seguir. En esta investigación se analiza el caso de la carne roja y sus presuntos efectos cancerígenos con el objetivo de determinar el seguimiento en Facebook y Twitter y el carácter de los comentarios (positivos, neutros o negativos). La agencia internacional para la investigación del cáncer (dependiente de la OMS) presentó el 26 de octubre de 2015 un informe donde se evaluaba la carcinogenicidad del consumo de carne incluyéndola en el grupo 2A. Esta noticia estuvo presente en las redes sociales, principalmente en Facebook y Twitter, las más utilizadas en España en ese momento. Por ello, se utiliza la metodología cualitativa de análisis de sentimiento en esas redes sociales y las publicaciones relacionadas con la noticia. Se buscaron las publicaciones del 26 de octubre al 26 de noviembre de 2015 y se utilizaron *hashtags* específicos ("carne roja" y "OMS") para centrar la búsqueda. Las publicaciones se clasificaron en positivas, neutras o negativas. También se valoraron las reacciones de los otros usuarios en cada publicación mediante un análisis estadístico cuantitativo. Se analizaron 391 publicaciones en Twitter y 33 en Facebook. En la primera red social, se encontraron publicaciones en todo el periodo de tiempo acotado; en Facebook, solo en la mitad del tiempo. El 57,6% de las publicaciones de Facebook presentan una intencionalidad neutra, el 27,3% negativa y el 15,1% positiva. En Twitter, 47,6% son neutras, 39,6% negativas y 12,8% positivas. El seguimiento de otros usuarios a las publicaciones estudiadas no fue estadísticamente significativo. La apreciación de los usuarios cambia con el paso del tiempo, pasando de comentarios negativos y neutros a neutros y positivos hasta que solo quedaron información objetiva. También se observa una falta de comunicación entre las entidades oficiales y los usuarios de estas redes.

Palabras clave:

Redes sociales, análisis de sentimiento, comunicación, OMS, carne roja.

1. Theoretical framework

1.1. Social networks and crisis communication

Red meat is consumed daily and is one of many foods that form the basis of our diet. It has great nutritional value, and its consumption is necessary in childhood, adolescence, and adulthood. Along the same lines, cancer is a disease that has become more prevalent in our society in the last few decades with an increase in the number of cases.

The statement from the World Health Organization (WHO) joined both issues by linking diets high in red meat with the appearance of colorectal cancer. However, that news did not receive enough attention in the media. In connection with this situation, a question arose: did that news have the same influence or recognition in social networks?

To answer the question, this research was carried out with two objectives. The first was to discover the relevance that the news obtained in social networks and the opinions related to it. The second was to examine the influence of these comments on the rest of the users of these networks through the options offered there, such as Sharing or Likes.

Social networks can be defined as forms of social interaction, or in other words, they offer dynamic exchanges between people, groups and institutions within complex contexts, an open system in permanent development that involves groups that identify themselves according to similar needs and problems, and who organize themselves in order to enhance their resources (Caldevilla Domínguez, 2010).

A total of 82% of Internet users aged 18-55 in Spain use social networks, which is approximately 14 million users, most of whom are looking for information. Facebook and Twitter (Fondevila et al., 2014) have become the most widely used social networks in the last decade: 96% of users visit Facebook daily, and up to 56% of those users participate in Twitter as well (IAB, 2015).

These networks allow users to quickly comment on news, special reports, or posts from other users. These comments can be posted using other resources like images, videos or computer graphics. The multimedia factor is the key, together with hypertextuality and interactivity (Fondevila, Beriain and Del Olmo, 2013). In fact, there is an empirical model of analysis of these variables (Fondevila, 2014).

Opinions on social networks can be decisive in resolving critical communicative issues. In the case of Twitter, diverse ways to automate content and summaries have been studied. Real-time updates result in a constant flow of content in the form of so-called 'cloud journalism' (Fondevila Gascón, 2010), an essential feature of the Broadband Society that leads to social journalism commerce (Fondevila Gascon, 2013). The monitoring feature of Twitter allows for the identification in real time of events occurring at the moment (Sayyadi, Hurst and Maykov, 2009, Popescu and Pennachiotti, 2010, Watanabe et al., 2011). Events such as earthquakes (Sakaki, Okazaki and Matsuo 2010), or even the launch of products to market (Chua and Asur, 2013), can also be foreseen. Twitter is becoming an outstanding source of information to the point of bringing about the Automatic Microblog Summary, which allows for the identification of main trends and the summarization of topics (Sharifi, Huttyon and Kalita, 2010). In the case of food, it is both prone to and sensitive to communicative crises, which can be generic (Cascante Hernández, 2011, Russell Brown, 2012), as well as specific, and are focused on the digital environment (Farré i Coma et al., 2012; Martín-Ruiz et al., 2014; Pang et al. Al., 2015).

The International Agency for Cancer Research (IARC) is the WHO agency that specializes in cancer. Its aim is to promote international collaboration in cancer research. This agency is interdisciplinary and brings together professionals skilled in epidemiology, laboratory research and biostatistics for the purpose of identifying the causes of cancer. In this way, it can adopt possible preventive measures and reduce the burden of disease and its associated suffering. An important feature of the IARC is its experience in coordinating research in all countries and organizations.

Thanks to the IARC, possible carcinogens that are known and need to be evaluated are classified into 4 large groups: Group 1 corresponds to carcinogenic substances; Group 2 is divided into 2 subgroups; A (*probably* carcinogenic) and B (*possibly* carcinogenic). Group 3 cannot be classified with regard to its carcinogenicity to humans, or in other words, there is no evidence that these substances cause cancer in humans; Group 4 refers to products that are not considered carcinogenic to humans.

On October 26, 2015, the agency submitted a report (IARC Monographs, Press release No. 240) assessing the carcinogenicity of red meat consumption and including it in group 2A as "probably carcinogenic". Included in this group are products such as refined petroleum, combustion gases and tanning lamps. This classification is based on limited evidence from epidemiological studies showing a positive association between red meat consumption and the development of colorectal cancer. The strongest evidence points to colorectal cancer, although red meat consumption has been linked to pancreatic and prostate cancer as well.

All muscle meat of mammals is considered to be red meat, including beef, veal, pork, lamb, horse and goat. This food was chosen for evaluation because in recent epidemiological studies researchers have found evidence of small increases in the risk of several types of cancer that could be associated with high consumption of red meat. Although risks are low, they may be pertinent for public health, as many people in the world eat meat. Moreover, its consumption is increasing in low and middle-income countries. To reach these conclusions, we took into account over 800 different studies on cancer in humans (in total, more than 700 epidemiological studies provided data on red meat). The work group consisted of 22 experts from 10 countries.

The word cancer is a very broad term that covers more than 200 types of diseases. Each of them can have completely different characteristics, but all have a common denominator: cancer cells acquire the ability to multiply and spread throughout the body without control.

The majority of colorectal cancer appears as a polyp on the lining of the colon, which develops into a malignant tumour. This tumour can be local, or in other words, it can increase by growing in depth and penetrating all the layers that form the wall of the digestive tract. That is to say, it grows from the mucosa to the serosa through the submucosal and muscular layers. Once the tumour passes through the entire intestinal wall it can invade any organ through two types of dissemination; lymphatic and haematogenous.

1.2. Sentiment analysis in social networks

The benefits of obtaining qualitative information from the opinions of users, as well as the technical complexity of the analysis of these opinions, generate the demand for solutions capable of monitoring the flow of reviews. For this reason, in recent years the sample of opinions, more commonly known as sentiment analysis, has been developed as an area of research. According to Pang and Lee (2005 and 2008), sentiment analysis focuses on automatically processing information that contains opinions, which among other things allows for classification of the polarity of a given text (positive, negative, neutral or mixed).

Classification of the polarity of texts written on social networks usually requires the use of semantic dictionaries and the analysis of the syntactic structure of sentences in order to classify a text as subjective, positive or negative. With regard to sentiment analysis, much of the effort is now being directed at tasks related to the classification of polarity, which is a problem that has been approached from two perspectives:

1) The first approach assumes that the task is a generic classification process (Pang, Lee and Vaithyanathan, 2002): it starts with a training set in which texts are recorded along with their polarity. A classifier is built through automatic learning.

2) The second approach is based on the semantic orientation of words in which each term that expresses an opinion is recorded with a value representing its polarity (Turney, 2002). Most systems of sentiment analysis are focused on the treatment of texts written in English. In the case of texts written in Spanish, the most relevant system is TheSpanish SO Calculator (Taboada et al., 2011), developed by Simon Fraser University in Canada. This system resolves the sentiment and polarity that individually contain adjectives, adverbs, nouns and verbs, treating all of these linguistic constructions as lexicon. This tool allows us to know the information contained in a text that we intend to analyse in order to extract its correct polarity and its intention in the message made by the issuer.

Systems of sentiment analysis are faced with a condition that should be considered, which is the orthographic quality of the texts. When texts come from Internet, authors often omit accents or letters, or use unrecognized abbreviations. The solution is to use patterns to adapt the text.

A basic task in this type of analysis is to classify the polarity of a text, sentence or word, that is to say, to determine if the text, sentence or word contains sentiments that are positive, negative or neutral.

Being in the dominant position, even though TheSpanish SO Calculator treats all linguistic constructions at the lexical level, there are some authors who propose the use of the sentence syntactic structure to obtain the semantic orientation of a text. To this end, it is necessary to follow these steps:

a) Process the texts: for this purpose, an *ad hoc* processor has been developed that deals with the unification of compound expressions, such as "unless". As a result, these compounds will have a single unit of meaning. In this section, the standardization of punctuation marks is fundamental.

b) Segment the text into sentences and divide them into tokens to perform the morphosyntactic labelling of each of the words in the text.

c) Syntactic analysis of dependencies: Binary dependency relationships between the terms in a sentence are identified. The result of each of these binary links is a dependency. The structure obtained is called a dependency tree.

Another source of qualitative data related to content is the Automatic Text Summary. This formula was created for the library system in the United States, given the need to digitally index contents. Storage limitations of the twentieth century led to the creation of this system, capable of summarizing contents with precision (Das and Martins, 2007).

The first attempts at developing automatic summaries by computer came in the late 1950s (Luhn, 1958). The first investigations studied the frequency of certain words, and then the position of the sentence within a text (appearance of the main sentence at the beginning of the text). Keyword detection began with Edmunson (1969), and tag clouds (tags) started in 1975 (Salton, Wong and Yang, 1975; Salton, 1988).

The models applied to the analysis of content using statistics are diverse. Thus, we find Bayesian models (Kupiec, Pedersen and Chen, 1995), machine learning (Berger and Mittal, 2000; Barzilay and Lapata, 2008) and semantic clustering techniques (Radev, Jing and Budzikowska, 2000). In the more strictly linguistic approach, research focuses on positions within the text (Brandow, Mitze and Rau, 1995, Lin and Hovy, 1997), the structure of discourse (Marcu, 1997, Polanyi et al., 2004) and lexical chains (Barzilay, 1999). Here we find the history of qualitative analysis through "big data", which improved the techniques of automatic text summarization.

Automatic Text Summarization software obtains source information, extracts the content and presents it in a condensed and useful way. The methodology in this case starts from the interpretation or analysis (study of a source document and construction of the representation), transformation (summary abstraction and techniques of Neurolinguistic Programming, transforming the representation of the document into a new representation) and generation or synthesis (representation of the abstract and construction of a new abstract in natural language). This system promotes current knowledge, saves reading time, facilitates the selection of content that is considered relevant, permeates searches within texts, improves content indexing efficiency and helps to develop opinions.

The multimedia environment inherent to Internet is promoting Automatic Text Summarization, but not only in written format. There may be a single document or many; the language can be monolingual or multilingual; the category can be technical, scientific or news; the length may be short (one or two pages) or long (more than 50 pages); and the medium can be text, audio or video. The output formats are the document (abstract), the format (thematic categories or labelled press summaries) and informative style, indicative (pointing to the subject without giving further explanation), critical, or aggregate (incorporating different sources).

Automatic Text Summarization methods can be statistical, or in other words, centred on the number of times a term is repeated in a text. Twitter's *trending topic* is a good example. The Bayesian classifier indicates the probability that a sentence in the document will be included in the abstract. In fact, you can indicate the number of capitals, the length of the sentence, or the structure and position of the sentence in relation to the entire text. There is also the question of measuring the meaning of a word instead of the frequency of occurrence using the WordNet tool, adding the number of times the word appears with the occurrence of synonyms and other associated terms.

The written language and its connotations can make it difficult to detect the sentences that might be the main topic. Methods have arisen as a result, such as lexical chains, used to search for cohesive relationships among terms, frequency of use, synonymy, antonymy, hypernymy or homonymy. Chains can be established by association (Barzilay, 1999). Similarly, the content co-reference method identifies whether or not natural language expressions refer to the same topic.

As for the iterative graph methods, algorithms such as HITS (Kleinberg et al., 1999) or Google PageRank (Brin and Page, 1999) are used. These allow for the study of the structure of the links and the hierarchy of the websites.

When you take into account all of the characteristics of a text (frequency of words, similarity between keywords, similarity between contents and titles, position within the sentence, various grammatical attributes), the 'fuzzy logic' method is used, where each sentence is scored from 0 to 1. The values obtained determine the degree of relevance of the sentence to the final summary (Kyoomarsi et al., 2008).

Intrinsic evaluation is carried out using intrinsic metrics. These metrics evaluate the content of the information and compare it to other Automatic Text Summaries carried out by humans, or to the original source of information (Mollá, 2003).

In the linguistic field, these factors are taken into account as criteria for evaluating the quality of a text: grammaticality (the text should not contain non-textual elements such as markers, incorrect words or punctuation errors), non-redundancy, clarity of references (names and pronouns frequently referenced), as well as structure and coherence (summary with a solid, defined structure and coherent, non-contradictory, comprehensible sentences) (Steinberger and Jezek, 2009: 13).

The criteria to be taken into account when evaluating content (content evaluation) are the following: co-selection (topics linked by similarities), position, recall and F-score results. The most common Automatic Text Summarization techniques are binary classification systems. They are called 'supervised' because training is required for those responsible for the information system. The binary word refers to the existence of two categories (positive or negative, for example, even though in sentiment analysis the neutral value is added). The precision formula is Precision = TP / (TP + FP). TP are the true positives and FP the false positives. The Recall is calculated using the formula Recall = TP / (TP + FN). FN are false negatives. The perfect classification implies precision and recall equal to one, and FP and FN would be equal to zero.

As for the F-Score, this is a combination of precision and recall, with the perfect F-Score being equal to one (Carenini, Murray and Ng, 2011: 28). Thus, the formula of the F-score is F-score = (B2 + 1) PR / B2P + R.

Another way of analysing a text is by using the Receiver Operating Characteristic (ROC). This is a graphic representation for a binary classification system with an area between 0.5 and 1, with 1 being the perfect score. The True Positive Rate (TPR) is obtained with the equation TPR = TP / (TP + FN), and the false positive rate (FPR) uses FPR = TN / (FN + TN).

In order to avoid sentences with discrepancies, the concept of Relative Utility (RU) was created in which each sentence obtains a utility, which allows for the inclusion of sentences according to the criterion of the evaluator, the original document and the length of the abstract (Ding, 2005). When the content of two sentences have the same meaning, content-based evaluation systems are applied (Steinberger and Jezek, 2009). A software widely used in content-based evaluation is ROUGE, which evolved from BLEU (Carenini, Murray and Ng, 2011), and which uses n-grams (sequences of n words; the unigrams are individual words). It is based on Natural Language Processing (NLP) and uses metrics to compare automatic summaries. The Pyramids method is a semiautomatic evaluation system that identifies summary semantic units from the comparison of several abstracts (Nenkova and Passonneau, 2004).

Latent Semantic Analysis (LS) is used to extract and represent the contextual use of the meaning of words through statistical programs that work on the content of large texts. The idea is that the sum of the contexts of the appearance or non-appearance of a word provides a set of relationships that determine the similarity of meaning between words or groups of words. Its purpose is to detect hidden dimensions, such as irony (Ding, 2005).

Extrinsic evaluation is based on a specific task and is called 'task-based measures'. It can be developed using full original texts, human-created summaries or Automatic Text Summaries. Qualitatively, humans are asked about the automatic summaries they have read, and the responses are compared to those of other people who have read the original. The comparison of the results of the two surveys determines the quality of each summary. These metrics also allow us to evaluate the usefulness of a summary with regard to the ability to substitute a real-world task. The creation of the abstract must have an implied objective (Sparck Jones, 1998).

There are several extrinsic evaluation systems. According to 'document categorization', categories are assigned to documents according to a characteristic (author or edition, for example). 'Information retrieval' is a system for evaluating the quality of a summary.

According to relevance or pertinence criteria, the relevance assessment task is used. A topic is assigned to an individual and the individual has to determine the relevance of the theme in the summary. Valuation accuracy and time invested are indicators of effectiveness.

Responsiveness (*question answering*) evaluates the number of correct answers made by a group of people when reading several documents (originals, automatic summaries or human summaries). The individuals answer questions and the responses are compared (Steinberger and Jezek, 2009). Reading comprehension also asks questions about the document to be read. Evaluation is based on the comprehension of the document and whether or not the responses are similar to those of Automatic Text Summaries (Morris, Kasper and Adams, 1992).

2. Methodology

This research focuses on red meat and its possible carcinogenic effects in order to evaluate the follow up by users on Facebook and Twitter and the nature of the comments made by them (positive, neutral or negative). The methodology used in this research was qualitative (sentiment analysis) and quantitative (quantitative data related to comments by users).

Sentiment analysis, (also known as "opinion mining", "brand monitoring", "buzz monitoring", "conversation mining", "online consumer intelligence" and "user- generated content"), is the field of study that analyses opinions, feelings, evaluations, aptitudes and emotions of people with regard to services, products or events, among other things. It is intended to conceptualize feelings as perceptible human reactions, or in other words, traceable and identifiable with a specific valence (Clore, Ortony and Foss, 1987).

The term sentiment analysis was first applied in the study carried out by Nasakawa and Yi (2003), and is linked to data processing and databases (Minqing and Bing 2004, Minqing, Bing and Junsheng, 2005; Martinez Cámara, Marín Valdivia and Ureña, 2011; Aguado-De-Cea et al., 2012). Studies of feelings have been present in different fields, such as psychology and linguistics, for purposes other than those used in this work (Fondevila et al., 2016).

Currently, we can find an ample number of social networks on Internet, some examples of which are Facebook, YouTube, Twitter, Google+, Snapchat and Instagram. Of all these networks, Facebook and Twitter were chosen because they are the most widely used networks among Spanish users in recent years (IAB, 2015), and because they are the quality reference in terms of giving and monitoring opinions through the Internet (Montesino García, 2014).

In order to find publications related to the topic in question, or in other words, news items in which meat products are included in the groups of substances that carry a risk of producing cancer (specifically, publications on red meat) the hashtag for #carneroja (#redmeat) was used in the search engine on the Facebook main page, and after the results were obtained, advertisements of butchers and brasseries that used this hashtag just to advertise their services were removed from the list. To find the publications on Twitter, the advanced search engine was accessed, and in the tags section the WHO and red meat (#WHO/#redmeat) hashtags were included. Given the large number of posts, they were filtered by date.

On the other hand, in order to verify the dissemination of the news among the population and possible changes of opinion over time, being that each day new information related to the subject of study appeared, a one month time period was used from the moment the report was presented by the WHO. In other words, the posts on the social networks were observed from October 26, 2015 to November 26, 2015.

For the posts, assessments were made regarding the date of publication, the category of the issuer, the text described, attachments, the positive appraisals of the remaining "participants" (to unify the language, "likes" on the social networks studied are called "ratings"), and links to the publication ("links" is the same as "retweets" in Twitter jargon and "shared" on Facebook). Additionally, comments were also assessed on Facebook.

The category was subdivided according to possible publication sources. "News" referred to publications released by the media, or to newspapers and online newspapers of a specific area or not well known. "Personal" referred to publications issued by private accounts that corresponded to an individual, either the account that used the person's real name or accounts with false or sarcastic names. "Official" referred to publications circulated by verified accounts of public and official entities (a Ministry, the WHO, city council, etc.).

"Nutritional" refers to blog publications, websites or specialists dedicated exclusively to aspects of nutrition if they are in this group and are not classified elsewhere. "Association" refers to publications released by unofficial organizations and non-profit institutions. "Ads" or "Advertising" refers to publications related to the subject matter, but with the intention of advertising and attracting attention to a restaurant or shop related to food. "Company", unlike the previous example, is a category for businesses not related to restaurants or food in general. Lastly, publications of web pages with these characteristics were also classified as "Blogs".

After posts were collected from the two networks named above, they were classified according to whether they had a "positive", "negative" or "neutral" assessment (Table 1). In order to carry out a sentiment analysis in the most empirical way possible, a list was made of words that signified positive feelings and another with words that could be considered negative in terms of feelings. Posts that did not contain any of the words or groups of words above were considered to be in the "neutral" category. On both Facebook and Twitter, an assessment was also carried out regarding the links and multimedia components that accompanied the publication, due to the fact that this type of supplement can reinforce the sentiment of a publication.

The presence of irony or sarcasm was also evaluated in the posts. If such aspects were detected, the post was classified as "negative", even if the words used had a positive meaning.

Table 1. Positive and negative adjectives on Facebook and Twitter in the food crisis analysed

Adjectives of positive comments	Adjectives of negative comments
Abundant	Accused
Suitable	Alarmist
Admirable	Censored
Affirmative	Be careful
Benefit	Deficient
Bold	Disproportionate
Trust	Wrong
Delicious	Madness
Desirable	Kill/destroy
Excellent	Fearful
Fantastic	Shit
Happy	Miserable
Funny	Death
Honest	Dangerous
Interesting	Laziness
Hurray	Suspicious
Common sense	Foolishness
Courageous	Clumsy

Source: Created by the authors with the collaboration of Victoria Villegas Jorquera

On the other hand, monitoring of the posts was carried out among other users of these networks to determine what type of opinion or publication had the greatest follow up. This was verified by performing a descriptive statistical analysis with the help of Excel options. Evaluations were carried out regarding mean and mode, standard error, variance, maximum and minimum values, asymmetry coefficient, and confidence level for each characteristic previously mentioned (comments, links and ratings for Facebook, as well as ratings and links for Twitter).

The research hypotheses are the following:

H1. The food crisis follow-up on Twitter is higher than on Facebook.

H2. Most posts related to the food crisis on Facebook are categorized as neutral.

H3. Facebook users have the tendency to assess the posts they find on this social network regarding the food crisis.

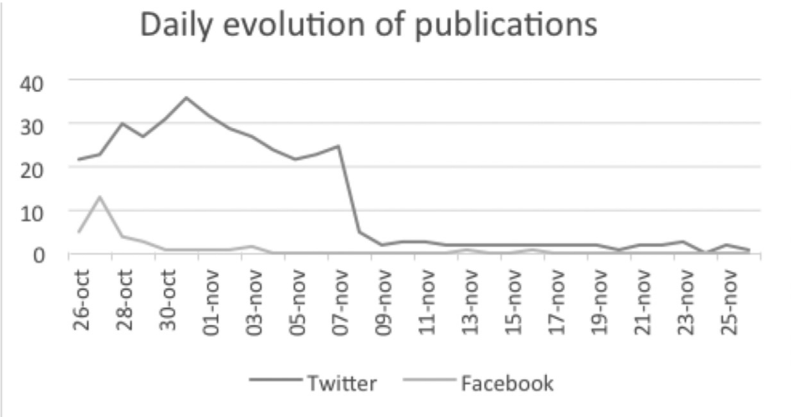
H4. The sum of positive and negative posts of the food crisis on Twitter surpasses neutral posts.

H5. Communication between official bodies and users of social networks can be improved.

3. Results

The time needed for publication is different between Facebook and Twitter: the time span of the former is shorter than that of the latter.

Graph 1. Daily publications



Source: created by the authors

It can be observed on Twitter that for two weeks the number of posts remained stable. As of November 8, the number of posts found was significantly reduced. Not a single post was found on November 24th, although some were encountered on the days that followed.

On Facebook, most of the posts were found in the first 4 days (25 posts). After that, for two days one post appeared each day, then there were no posts until the last day when one publication appeared (Graph 1). This confirms H1 (states that follow-up to the food crisis on Twitter is higher than on Facebook).

3.1. Facebook

Thirty-three publications on Facebook covering the period from October 26, 2015 to November 26, 2015 were analysed. Of these publications, 8 belonged to news pages, 10 to private users, 1 was part of an advertisement for a brasserie, 4 were from official pages, 5 from pages about nutrition, 1 from general information blogs, and 4 were from information or mothers associations (Table 2).

Table 2. Sentiment analysis on Facebook regarding red meat

Nº	Assessment	Category	No.	Assessment	Category	No.	Assessment	Category
1	Neutral	News	12	Negative	Nutritional	23	Neutral	Personal
2	Neutral	Advertising	13	Positive	News	24	Neutral	Association
3	Neutral	News	14	Neutral	News	25	Negative	Personal
4	Neutral	Association	15	Neutral	Nutritional	26	Neutral	Blog
5	Neutral	News	16	Neutral	Official	27	Positive	Personal
6	Neutral	News	17	Positive	Nutritional	28	Negative	Personal
7	Negative	Association	18	Neutral	Nutritional	29	Negative	Personal
8	Negative	Association	19	Positive	Personal	30	Negative	Personal
9	Neutral	Official	20	Neutral	Personal	31	Neutral	Official
10	Negative	News	21	Negative	Personal	32	Neutral	Official
11	Neutral	News	22	Positive	Personal	33	Neutral	Nutritional

Source: created by the authors

3.1.1. Sentinment analysis

A figure of 19/33 (57.6%) was considered to be from neutral posts. These were informative messages about the news, references to their own posts, or direct questions to the readers about the situation. The publications belonging to official sources were all neutral. Media publications were mostly neutral (31%). The rest were divided among the various categories.

We classified 9/33 (27.3%) as negative posts. To determine this category, we considered whether or not there was irony and/or sarcasm involved, and words with negative feelings, such as "harmful" or "bad." Images or videos accompanying the publications were taken into account, especially involving personal anecdotes.

We considered that 5/33 (15.1%) of the posts were positive. To determine this category, we took into account words with a positive meaning, such as "common sense" or "healthy". We also considered the images that accompanied the posts. This confirms H2, which states that most posts related to the food crisis on Facebook are classified as neutral, probably due to prudence on the part of the media analysed and by the conservative evaluation criteria inherent in the sentiment analysis methodology.

3.1.1. Statistics of evaluations, links and comments

A large data dispersion was observed. Fully 51.5% of the posts contained no comment. Regarding comments from the remaining publications, the mean is 2.12, with a standard error of 0.74. Mode is 0. Standard deviation, 4.24. Variance is 17.98. Coefficient of asymmetry, 2.69. Confidence level at 95% is 1.5. The post that contains the most comments is a post from a private account, which has 18.

In contrast, almost all posts have an evaluation of (84.8%), with an average of 20.45 "sentiments" per post with a standard error of 7.73. The one with the highest number of "Likes" (223) corresponds to the largest number of comments. Mode is 3 with a standard deviation of 44.42. Variance is 1,973.13. The asymmetry coefficient is 3.61. Confidence level at 95% is 15.75. As such, this confirms H3 (states that Facebook users have the tendency to assess posts related to the food crisis found on this social network).

There are very few shared posts, 5 of 33 (15.1%), and these are widely separated, between 1 and 100 times (with the 100 times belonging mostly to an official account). Statistical analysis was not performed due to the data dispersion.

3.2. Twitter

A total of 391 posts were analysed during the pre-determined time period. Of these publications, 57 were from news sources, 255 from private accounts, 2 were advertisements, 24 from associations, 15 from blogs, 13 from companies, 20 were from nutritional sources, and 5 were from official sources (Table 3).

Table 3. Sentiment analysis on Twitter regarding red meat

Date	Positive	Neutral	Negative	Total	Date	Positive	Neutral	Negative	Total
26-0ct	0	8	14	22	11-nov	0	1	2	3
27-0ct	0	9	14	23	12-nov	0	0	2	2
28-0ct	0	9	21	30	13-nov	0	1	1	2
29-0ct	0	13	14	27	14-nov	2	0	0	2
30-0ct	1	12	18	31	15-nov	1	0	1	2
31-0ct	5	20	11	36	16-nov	0	2	0	2
01-nov	5	16	11	32	17-nov	0	2	0	2
02-nov	5	16	8	29	18-nov	0	2	0	2
03-nov	6	16	5	27	19-nov	0	2	0	2
04-nov	5	13	6	24	20-nov	0	1	0	1
05-nov	5	14	3	22	21-nov	0	0	0	2
06-nov	2	8	13	23	22-nov	1	1	1	2
07-nov	9	11	5	25	23-nov	0	2	0	3
08-nov	1	1	3	5	24-nov	0	0	0	0
09-nov	0	2	0	2	25-nov	0	2	0	2
10-nov	0	2	1	3	26-nov	1	0	1	1

Source: created by the authors

3.2.1. Sentiment analysis

A total of 186/391 (47.6%) of the posts were considered neutral. Of these, 75% contained factual information related to the news or its consequences. That being the case, neutral publications were distributed among all of the categories named above so that none in particular stands out to a large degree.

The figure of 155/391 (39.6%) of the posts were considered negative. Of these, 39% were ironic, which could be directed at not eating meat, or at the WHO, or even at the government. In this case, the category that accounted for the largest number of posts was highly differentiated: 70% were published from personal accounts.

Finally, 50/391 (12.8%) of the posts were positive. More than half (51%) corresponded to personal posts; the rest belonged to blogs, companies, and advertisements. As such, this validates H4 (states that the sum of positive and negative posts on the food crisis on Twitter surpasses neutral publications).

3.2.2. Statistics of valuation and links

The assessment mean is less than one per publication (0.98), with a standard error of 0.1. In this case, the mode is also 0. The one with the highest number of assessments accumulated is 15; it was a post from the association "La Jornada Online". 61.4% did not contain any assessment. The standard deviation is 2.13; Variance is 4.54. When the 95% point is reached, the confidence level is 0.21. Asymmetry coefficient is 3.94. The initial comments were ironic, such as *"according to the #WHO, everything produces cancer, so you might as well not bother and just go ahead and die".* There was also a lot of informative news (in this case considered to be composed of neutral comments: *"The latest health news from the #WHO on the relationship between #redmeat and #cancer*").

As the days passed, the trend gravitated toward neutral and informative comments, along with the emergence of positive messages. This is not to say the messages were in support of the WHO, but they were calls for common sense without reaching the level of sarcasm (*"#WHO I'm sorry, happy Saturday ... I promise never again!!!"*). In the end, after nearly a month had passed, emotional reactions from users ended and only objective information remained. Just one comment was found from the vegan community showing its agreement with the WHO study. However, there were user comments that connected the agreement with a possible alliance between vegans and this organization to increase adherents to the discipline (*"They are connected with the #WHO, this is a #vegan plot"*).

Despite the potential drama of the news, it was sometimes seen as relative. Some companies took advantage of the situation to promote themselves even in different areas not related to the meat industry (*"Don't stop enjoying the greatest pleasures of life. Pasteleria Dgusto - Dgusto Pastry Shop"*). On the other hand, the case is shared with and related to politics, and many comments used red meat as a metaphor for communists (red people) (*"La Falange [a far right party in Spain] of the WHO declares #redmeat bad for health; we already have #fascism, #EverythingRedIsBad"*). Comments were also observed regarding people saying that everything the WHO states is nonsense, which corresponds to statements made by politicians (*"the #WHO declares as much nonsense as you @PabloIglesias"*). With regard to poverty, it was implied that people in the third world do not have to worry about cancer because not everyone has access to the amount of meat necessary to produce the disease.

This attempt to make the news relative can be understood by the reaction of social network users. Judging by their comments, many users did not believe the information given by the WHO. The vast majority attached photos of plates and barbecue grills filled with meat, or even with themselves eating (*"I've never liked them, but tomorrow I'm going to start practicing high-risk sports"; "My rebellious spirit makes me respond to the #WHO by preparing a pot of pork loin. Would that be suicide?"; "Red meat causes cancer, according to the #WHO... this is like when they say the sun causes cancer. I'll continue eating meat from my barbecue"*

On the other hand, from November 3 onward it became trendy to announce that the consumption of beer would nullify or delay the onset of cancer (*"Faced with the #WHO, we still have an ally in #beer"*). An increase in negative messages was also detected on October 30 after the European Parliament approved the legalization of insects as a new food. Many tried to unite the two news reports by saying that the second message was a solution to the first conflict (*"The new diet according to the WHO"*). A final blow to the organization came when the comment began circulating among users that until 1990 the WHO considered homosexuality to be a mental illness. This occurrence greatly diminished the credibility of the organization even more, as well as the significance of the news.

After the presentation of the news and the avalanche of negative comments, the WHO changed the message three days later to soften the impact, and announced that they were not indicating that people should stop eating meat, but that it was necessary to moderate its consumption. Afterward, several publications appeared from organic food advocates refuting the information about meat and trying to clarify that the most detrimental aspects were preservatives and other additives used in the meat processing industry (cutting, packaging, preserving, etc.), yet defended meat itself (*"Meat is not bad, but additives are, #meat #WHO #gastronomy"*).

It was stated that if meat was organic, locally-produced, and consumed in moderation, there was no reason to worry about the WHO announcement. With regard to this point, another statement was then made, mentioning that although several reports appeared regarding other substances included in the carcinogenic groups, such as glyphosate (a broad-spectrum non-selective herbicide), not much commotion was generated regarding this issue because it was not such a delicate matter and did not affect the industry as much as the item about meat.

Among the posts, especially after the WHO rectification regarding meat consumption, some complained about institutional communication (*"Why is the #WHO so inept at communicating?"*). Correcting their own information indicated that they were inefficient with their own communications.

The average number of links is one per post. The maximum number is 34, pertaining to a news publication. The typical error is 0.15, although the mode is 0, since 70% of the publications do not contain any link. The standard deviation is 3.14; therefore, variance is 9.83. Asymmetry coefficient is 6.53. When the 95% point is reached, the confidence level is 0.31. In total, this validates H5 (states that communication between official entities and users of social networks can be improved).

4. Conclusions

As one can observe, on Facebook there was less follow-up than on the microblogging social network. This may be the result of more frequent use of Twitter to comment on current affairs, while Facebook is basically used to maintain social relationships, although it tends to include content related to current events as well. In any case, trivial issues and personal situations take precedence on Facebook.

In Facebook postings, a modest number of comments were found that came directly from Twitter by way of a direct link. Twitter attracted more follow-ups in the designated time period to the point that posts were found up to a month later. On Facebook, the news did not have as much impact and did not stay more than two weeks. After that period, the intensity of the news diminished. This is due to the fact that other news occurred in the media at that time: specifically, the end of the Ebola contagion in Sierra Leone. This happened to be the main news item at that moment with the hashtag #who. As such, Twitter is the most appropriate network to use in monitoring a food crisis.

From the moment the news appeared until the end of the sample, perception fluctuated. At first there was a greater frequency of negative comments, many of which were ironic with regard to the association between meat and cancer. In the end, objectivity prevailed.

It was noted that most posts related to the food crisis on Facebook were neutral, and that publications regarding the food crisis tended to be rated by users. On Twitter, the sum of positive and negative posts surpassed the number of neutral publications, although neutral posts were leaders of the classifications.

On both social networks, a large number of publications belong to private accounts of people who are largely devoted to giving their opinion rather than informing. In second position are publications from news sources, the messages of which are basically neutral comments, surveys, and links to national newspapers (*ABC, El Periódico de Catalunya, La Vanguardia*). In the first few days, publications from public institutions were scarce, especially regarding the red meat warning. It would have been indispensable to have further publications giving advice, data related to minimum requirements, or messages of support to industry or producers.

It was observed that communication between official entities and users of social networks clearly has room for improvement. Thus, more publications are required by the meat industry and primary producers in order to question the WHO, or to appeal for better judgment. There was a proliferation of comments from nutritionists, or from the nutrition sector, that reported the news and tried to give scientific explanations. On the other hand, there was no information (and almost no comments) from vegans.

On the two social networks analysed, no comment was seen from the meat industry, but there were posts with links to reports that responded to the controversial news. One of the major concerns in this area was whether this information would lead to a fall in the price of meat products, although according to comments made, the industry suffered no repercussions.

It was observed during the investigation that numerous publications appeared in the Italian language, even though searches were conducted using the Spanish language. This was due to the fact that the meat industry is highly important in Italy and this type of food is deeply ingrained in the culture. The majority of Italians had a negative opinion regarding comments made by the WHO.

Statistical analysis was complicated as a result of data dispersion. On Facebook, the number of samples needed was barely obtained. On the other hand, a large amount of data from both social networks had a value of 0. For that reason, the mode in most samples also had a value of 0. According to variance related to Facebook data, valuations are more dispersed than the comments, probably due to the fact that the data range is lower (18 for the second versus 223 for the first). However, according to variance on Twitter, valuations have less dispersion than the links (the range is 15 and 34, respectively). The asymmetry coefficient tells us that the distribution is not symmetric and that the values tend to gather to the left of the mean. Those with the highest number (Facebook ratings and Twitter links) once again indicate that their data is mostly scattered. The mean of the data in which the range is less than 35 is between 1 and 2. In the case of Facebook valuations, which had the broadest range, there was little follow-up by users of other social networks in posts from private users as well as those that disseminate information.

5. Discussion

After carrying out the sentiment analysis, it was ascertained that this method is suitable for analysing comments on social networks regarding food crises by adapting general methodology (Pang and Lee, 2008) to specific applications (Steinberger and Jezek, 2009; Ng, 2011). It was observed that Twitter is a social network that contributes to the monitoring of food crises, which is in line with the thinking of Popescu and Pennachiotti (2010) and Montesino García (2014).

Sentiment analysis can easily be applied to publications of public health interest because it allows people to follow comments on the news and to introduce information so that citizens are aware of the seriousness of a situation or the need for change. In addition, there is a need to improve and apply this technology to the field of public health. In the bibliography, automatic analysis programs in English can be found, but there have been few advances in this technology for the Spanish language (Taboada et al., 2011).

Regardless of the news, it is important to highlight the negative perception that Facebook and Twitter users have of the WHO. This organization was created for the purpose of managing prevention policies and promoting health. If the population does not take into account the opinions of this institution, either because its information is somewhat alarming every time it appears, is inaccurate or incomplete, or requires later explanation, then in the moment when there is a real food crisis, either specific or global, people might ignore statements made by the organization. The exaggeration about which users have complained can result in negative consequences for the WHO. Furthermore, judging by the results of the investigation, the WHO should strive to convey more strictness and prudence. If not, the organization will not fulfil the purpose for which it was created.

Finally, it would be advisable for the meat industry to have greater presence in digital communication and social networks in general, given the amount of public opinion generated by these sources of information. The fact that the meat industry has not sent any messages stating its opinion regarding its own sector (whether positive, neutral or negative), reflects a lack of communicative experience in a business environment that cannot turn its back on reality and the content that consumers receive and share, and upon which they comment.

6. Bibliographic references

Aguado-de-Cea, G.; Barrios, M. A.; Bernardos, M. S.; Campanella, I.; Montiel-Ponsoda, E; Muñoz-García, Ó. y Rodríguez, V. (2012): Análisis de sentimientos en un corpus de redes sociales. Tenerife: 31 Congreso Asociación Española de Lingüística Aplicada, Comunicación, Cognición y Cibernética.

Barzilay, R. (1999): "Using Lexical Chains for the Text Summarization". *Advances in Automatic Text Summarization*. Massachusetts: USA.

Barzilay, R. y Lapata, M. (2008): "Modeling Local Coherence: An entity-based approach". *Computational Linguistics*, vol. 34, n. 1.

Berger, A. y Mittal, V. O. (2000): "OCELOT: A System for Summarizing Web Pages". *Proceedings of the 23rd annual international ACM SIGIR conference on research and development in information retrieval*, pp. 144-151.

Brandow, R. M.; Mitze, K. y Rau, L. F. (1995): "Automatic condensation of electronic publications by sentence selection". *Information Processing & Management*, 31(5): 675-685.

Brin, S. y Page, L. (1999): "The anatomy of a large-scale hypertextual web search engine". *Computer Networks*, vol. 56, n. 8, pp. 3825-3833.

Caldevilla Domínguez, D. (2010): "Las Redes Sociales. Tipología, uso y consumo de las redes 2.0 en la sociedad digital actual." *Documentación de las Ciencias de la Información*, n. 33, pp. 45-68.

Carenini, G.; Murray, G. y Ng, R. T. (2011): *Methods for Mining and Summarizing Text Conversations*. London: Morgan & Claypool Publishers.

Cascante Hernández, K. (2011): "Crisis alimentaria: un consenso por la agricultura". *Política exterior*, vol. 25, n. 142, pp. 128-137.

Chua, F. C. T. y Asur, S. (2013): "Automatic Summarization of Events from Social Media". *International AAAI Conference on Web and Social Media.*

Clore, G. L.; Ortony, A. y Foss, M. (1987): "The psychological foundations of the affective lexicon" *Journal of personality and social psychology*, vol. 53, n. 4, pp. 751-66.

Das, D. y Martins, A. (2007): "A Survey on Automatic Text Summarization". *Computer Science and Applied Cognitive Science*, pp. 1-31.

Ding, C., H. Q. (2005): "A Probabilistic Model for Latent Semantic Indexing". *Journal of the American Society for Information Science and Technology*, vol. 56, n. 6, pp. 597-608.

Edmundson, H. P. (1969): "New Methods in Automatic Extracting". *Journal of the Association for Computing Machinery*, vol. 16, n. 2.

Farré i Coma, J.; Gonzalo Iglesia, J. L.; Lores García, M; Lozano Monterrubio, N. y Prades Tena, J. (2012): "Comunicación de riesgos y seguridad alimentaria en la era 2.0". *El profesional de la información*, vol. 21, n. 4, pp. 381-384.

Fondevila Gascón, J. F. (2010): "El *cloud journalism*: un nuevo concepto de producción para el periodismo del siglo XXI". *Observatorio (OBS*) Journal*, vol. 4, n. 1, pp. 19-35.

- (2013): "Periodismo ciudadano y cloud journalism: un flujo necesario en la Sociedad de la Banda Ancha". *Comunicación y Hombre*, n. 9, pp. 25-41.

- (2014): "El uso de hipertexto, multimedia e interactividad en periodismo digital: propuesta metodológica de ranking de calidad". *ZER, Revista de Estudios de Comunicación (Journal of Communication Studies)*, vol. 19, n. 36, pp. 55-76.

Fondevila Gascón, J. F.; Carreras Alcalde, M.; Del Olmo Arriaga, J. L. y Feliu Roé, L. (2014): Facebook and Twitter in education: a new factor to improve classes. Valencia: *INTED2014 (8th International Technology, Education and Development Conference)*.

Fondevila Gascón, J. F.; Mir Bernal, P.; Puiggròs Román, E.; Muñoz González, M.; Berbel Giménez, G.; Gutiérrez Aragón, Ó.; Feliu Roé, L.; Santana López, E.; Rom-Rodríguez, J.; Sorribas Morales, C. y Crespo, J. L. (2016): Sentiment analysis as a qualitative methodology to analyze social media: study case of tourism. Porto: *CIAIQ2015 y ISQR2016, 5er Congreso Iberoamericano de Investigación Cualitativa y 1st International Symposium on Qualitative Research*.

Kleinberg, J. K.; Kumar, R.; Raghavan, P.; Rajagopalan, S. y Tomkins, A. S. (1999): "The web as a graph: measurements, models, and methods". *Computing and Combinatories*, n. 1627, pp. 1-17.

Kupiec, J. P.; Pedersen, J. y Chen, F. (1995): "A trainable document summarizer". *Proceedings of the 18th annual international ACM SIGIR conference on Research and Development in Information Retrieval*.

Kyoomarsi, F.; Khosravi, H.; Eslami, E.; Khosravyan Dehkordy, P. y Tajoddin, A. (2008): "Optimizing text summarization based on fuzzy logic". *Seventh IEEE/ACIS International Conference on Computer and Information Science*.

Lin, C. Y. y Hovy, E. (1997): "Identifying Topics by Position". *Proceedings of the fifth Conference on Applied Natural Language*, pp. 283-290.

Luhn, H. P. (1958): “The automatic creation of literature abstracts”. *IBM Journal of Research Development*, pp. 159-165.

Marcu, D. (1997): *The rhetorical parsing, summarization, and generation of natural language texts*. Toronto: University of Toronto.

Martín-Ruiz, E.; Danet Danet, A.; Pérez Corral, O.; Cruz Piqueras, M. T. y March Cerdá, J. C. (2014): “Opiniones de menores residentes en Andalucía (España) sobre temas de salud: Aplicación de un cuestionario por Internet”. *Revista Salud Uis,* vol. 46, n. 2.

Martínez Cámara, E.; Martín Valdivia, M. T. y Ureña López, L. A. (2011): Análisis de sentimiento. Jaén: *IV Jornadas TIMM. Tratamiento de la Información Multilingüe y Multimodal.*

Minqing, H. y Bing, L. (2004): “Mining and Summarizing Customer Reviews”. *Proceedings of the ACM SIGKDD International Conference on Knowledge Discovery and Data Mining (KDD-2004).*

Minqing, H.; Bing, L. y Junsheng, C. (2005): “Opinion observer: Analyzing and comparing opinions on the web”. *Proceedings of the 14th International World Wide Web conference (WWW-2005).*

Mollá, D. H. (2003): “Intrinsic versus extrinsic evaluations of parsing systems”. *Proceedings of EACL03 workshop on Evaluation Initiatives in Natural Language Processing.*

Montesino García, L. (2014): “Análisis de sentimientos y predicción de eventos en Twitter”. Santiago, Chile: Universidad de Chile, Facultad de Ciencias Físicas y Matemáticas [Consultado el 03/12/2016]. Disponible en: http://www.repositorio.uchile.cl/handle/2250/130479

Morris, A. H.; Kasper, G. M. y Adams, D. A. (1992): The Effects and Limitations of Automated Text Condensing on Reading Comprehension Performance. *Information Systems Research*, vol. 17, n. 35.

Nenkova, A. P. y Passoneau, R. (2004): Evaluating Content Selection in Summarization: The Pyramid Method. Vancouver: Document Understanding Conference.

Pang, B. y Lee, L. (2005): “Seeing stars: Exploiting class relationships for sentiment categorization with respect to rating scales”. *Proceedings of the 43rd Annual Meeting ion Associations for Computational Linguistics.*

Pang, B. y Lee, L. (2008): “Opinion mining and sentiment analysis”. *Foundations and Trends in Information Retrieval,* vol. 2, n. 1-2, pp. 1-135.

Pang, B.; Lee, L. y Vaithyanathan, S. (2002): Thumbs up?: sentiment classification using machine learning techniques. *Proceedings of the ACL-02 conference on Empirical methods in natural language processing,* n. 10, pp. 79-86.

Pang, Z.; Chen, Q.; Han, W. y Zheng, L. (2015): “Value-centric design of the internet-of-things solution for food supply chain: Value creation, sensor portfolio and information fusion”. Information Systems Frontiers, vol. 17, n. 2, pp. 289-319.

Polanyi, L. C.; Culy, C.; Van den Berg, M.; Lorenzo Thione, G. y Ahn, D. (2004): “A rule-based approach to discourse parsing”. *Proceedings of the fifth SIGdial workshop on discourse and dialogue.*

Popescu, A. M. y Pennachiotti, M. (2010): "Detecting Controversial Events from Twitter". *Proceedings of the 19th ACM International conference on Information and Knowledge Management.*

Radev, D. R.; Jing, H. y Budzikowska, M. (2000): "Centroid-based summarization of multiple documents: sentence extraction, utility-based evaluation, and user studies". *Proceedings of the 2000 NAACL-ANLP workshop on automatic summarization.*

Russell Brown, L. (2012): "Para prevenir una nueva crisis alimentaria". *Política exterior*, vol. 26, n. 150, pp. 176-184.

Sakaki, T; Okazaki, M. y Matsuo, Y. (2010): "Earthquake shakes Twitter users: real-time event detection by social sensors". *Proceedings of the 19th International Conference on World Wide Web.*

Salton, G. (1988): *Automatic Text Processing. The Transformation Analysis and Retrieval of Information by Computer.* Boston: Addison-Wesley Longman Publishing CO.

Salton, G.; Wong, A. y Yang, C. S. (1975): "A vector space model for automatic indexing". *Communication of the ACM*, vol. 18, n. 11, pp. 613-620.

Sayyadi, H.; Hurst, M. y Maykov, A. (2009): "Event Detection and Tracking in Social Streams". *Proceedings of the Third International ICWSM*, pp. 311-314.

Sharifi, B.; Hutton, M.-A. y Kalita, J. (2010): "Summarizing Microblogs Automatically". *Human Language Technologies: The 2010 Annual Conference of the North American Chapter of the ACL*, pp. 685-688.

Steinberger, J. y Jezek, K. (2009): "Evaluation measures for text summarization". *Computing and Informatics*, n. 29, pp. 1001-1026.

Sparck Jones, K. (1998): *Automatic summarizing: factors and directions*. Cambridge: MIT Press.

Taboada, M.; Brooke, J.; Tofiloski, M.; Voll, K. y Stede, M. (2011): "Lexicon-based methods for sentiment analysis". *Computational linguistics*, vol. 37, n. 2, pp. 267-307.

Turney, P. D. (2002): "Thumbs up or thumbs down?: semantic orientation applied to unsupervised classification of reviews". *Proceedings of the 40th annual meeting on association for computational linguistics*. pp. 417-424.

Watanabe, K.; Ochi, M.; Okabe, M. y Onai, R. (2011): "Jasmine: a real-time local-event detection system based on geolocation information propagated to microblogs". *Proceedings of the 20th ACM International Conference on Information and Knowledge Management*, pp. 2541-2544.