10, 11, 12, 13 e 14 de novembro de 2025
ISSN 2177-3866

HOW FAKE NEWS AFFECTS BRAND ASSOCIATIONS AND CORPORATE REPUTATION: A QUANTITATIVE STUDY FROM THE BRAZILIAN MARKET

BRUNO PERES
FACULDADE DE ECONOMIA, ADMINISTRAÇÃO E CONTABILIDADE DA UNIVERSIDADE DE SÃO PAULO - FEA

EDSON CRESCITELLI
FACULDADE DE ECONOMIA, ADMINISTRAÇÃO E CONTABILIDADE DA UNIVERSIDADE DE SÃO PAULO - FEA

INÁ FUTINO BARRETO
CENTRO UNIVERSITÁRIO ÁLVARES PENTEADO (FECAP)

HOW FAKE NEWS AFFECTS BRAND ASSOCIATIONS AND CORPORATE REPUTATION: A QUANTITATIVE STUDY FROM THE BRAZILIAN MARKET

INTRODUCTION

The spread of disinformation is a historical phenomenon that has acquired new contours with the rise of digital platforms (Lazer et al., 2018; Tandoc et al., 2018a). In an increasingly decentralized information ecosystem, the affordances of social media have enabled any individual to become a content creator, accelerating the spread of false information at rates significantly higher than those of verified information (Vosoughi et al., 2018a). Within this context, disinformation—defined as the intentional generation of misleading content to deceive (Wardle & Derakhshan, 2017)—poses critical threats to institutions, individuals, and especially to brand perception, which reflects how consumers cognitively and emotionally evaluate brands in online environments (Wang, 2016).

Brand associations are a set of images and perceptions reflected in the consumer's memory about the brand (Aaker, 1996; Keller, 1993; Yoo et al., 2000) . The dynamic has drawn increasing concern from scholars and institutions, as disinformation not only threatens democratic discourse and public health but also poses direct risks to brands and their perception in digital contexts (Borges-Tiago et al., 2020), which includes affective and cognitive judgments formed through digital interactions, can be manipulated by malicious content and automated agents. Liu (2019), for example, analyzed over 29 million tweets across 24 brands and found that social bots distorted brand-related discourse, creating misleading narratives. Similarly, Santini et al. (2020) demonstrated that media bots amplified media-related content, artificially inflating visibility and altering perceived relevance for media brands. Tuten and Perotti (2019) further emphasized that false claims about brands contributed directly to increased negative sentiment, suggesting that fake news often damages brand reputation in nuanced ways. Together, these studies support the notion that misinformation, often spread via bot-driven mechanisms, may negatively affects brand perception on platforms like Twitter.

Research question

While previous research has highlighted the disruptive influence of disinformation on brand-related discourse (X. Liu, 2019), limited empirical evidence exists regarding its direct effect on consumer sentiment and perception of brands across different stages of misinformation exposure. Particularly in emerging digital markets like Brazil—where misinformation exposure is high (Newman et al., 2020)—this gap becomes more salient. Furthermore, scholars have paid less attention to the recovery phase, as Dutta and Pulling (2011), state that brand crises can affect aspects of brand equity, leaving a critical gap in understanding how brands might strategically restore trust and sentiment in the aftermath of disinformation. Given this context, the research question guiding this study is: How does exposure to misinformation influence user sentiment and brand perception on social media platforms?

Objectives

So, the primary objective of this study is to identify the impact of disinformation on user sentiment expressed toward brands on social media. Specifically, the study seeks to examine whether exposure to *fake news* related to brands on Twitter alters user sentiment and comment polarization. To address this, a quantitative methodology was employed, using sentiment analysis on user-generated content. Tweets mentioning brands affected by disinformation were collected before and after the circulation of misleading content. Sentiment polarity was assessed using a lexicon-based method (Kim & Lim, 2021), which has been widely validated in digital behavior studies (Anderson et al., 2019; Hasan et al., 2018a; Ram et al., 2024).

From a theoretical perspective, this study contributes to advancing the literature at the intersection of brand association, and misinformation on digital communication environment. By integrating brand equity, disinformation, and digital sentiment, the research offers new insights into how reputational narratives are shaped and distorted in algorithmically driven environments. From a managerial perspective, the findings provide strategic guidance for brand managers and communication professionals on how to monitor and mitigate the reputational damage caused by disinformation. As brands increasingly operate in polarized, data-rich ecosystems, understanding the dynamics of perception under misinformation pressure becomes essential for effective crisis communication and long-term brand equity.

LITERATURE REVIEW

Misinformation and brand crisis management on digital channels

Initially found on websites that intentionally posted fictional partisan content and news with impactful titles in search of traffic with a financial bias, called *clickbait* (Marwick & Lewis, 2017) misinformation actions have multiplied. False information, disguised as news, has already created concerns in several countries throughout history and, more recently, has been popularly named *fake news* and associated with misinformation. Despite the popularity of the term, Wardle and Derakhshan (2017) point out that *fake news* does not encompass the complex phenomenon of informational disorder currently experienced, with misinformation being the official term for information shared to cause harm and linked to information disclosed in the form of news, but which do not correspond to the truth (Dentith, 2016), with the producer intending to misinform through the manufacture of items that can be published on websites, blogs, social media or messaging apps. The content and format are not the only ones responsible for creating a real appearance, but the fact that they are widely disseminated through the internet falsely built to feed this purpose can give the reader the feeling that many people are also reading such an item, adding legitimacy to it (Kalsnes, 2018; Tandoc et al., 2018b)

The practice of public relations in communication starts building relationships between different audiences. This occurs through defined discussions on topics involving two or more groups, which raise issues to be debated by society and which often generate positive or negative consequences, including companies and organizations (Heath, 2006; Jahng et al., 2020; Kent et al., 2002). When it comes to misinformation, according to a survey conducted by Reuters in partnership with the University of Oxford, social media are the biggest sources of distribution of untruths (Newman et al., 2020).

In the crisis communication theory, the focus is on how a company should respond to a problem, depending on the type and how severe it is, or is becoming, in the face of brand equity (Pace et al., 2017). This becomes more challenging in times of shared creation and with more misinformation disseminated with the help of social media algorithms (UNESCO; & University of Oxford;, 2018), and the speed of sharing messages on such networks. Dutta and Pulling (2011, p. 1285), state that “brand crises negatively affect aspects of brand equity”.

In an era where social media has changed the way people communicate, public relations professionals spend part of their time online, due to the agility of information and reactions (Bashir & Aldaihani, 2017). The social platform X (former Twitter) has 585 million users worldwide and is also one of the oldest social platforms (DataReportal & Hootsuite, 2025). Its retweet mechanism — a form of replying and sharing messages — offers an unprecedented opportunity for computer scientists, sociologists, and other scholars to study human behavior (Kwak et al., 2009).With the productive intensity of content generated by digital communications serving as a ground for sharing misinformation, Jahng (2020) highlighted the scarcity of studies related to how public relations professionals understand and address the complex nature of untruths. The analysis by Brennen et al. (2020)indicated social media platforms as the environment with most posts classified as false by fact checkers, with the X having the highest percentage. For this research, X is defined as the source of the collection of misinformation.

Brand equity and Brand associations

The Brand equity is the group of assets and liabilities of a brand, adding or subtracting value from the products and services provided by a firm. The equity is based on consumers' perceptions and memories about the brand, from previous information and experiences, either positive or negative, this study considers the following dimensions of brand equity: perceived quality, brand loyalty, brand awareness, and brand associations (Aaker, 1996; Keller, 1993; Shocker & Aaker, 1993).

By definition, brand associations are customers' memories connecting attributes, benefits, prices, ideas, experiences, facts, etc. to a given brand (Keller, 1993). The associations derive not only from the experience or knowledge of a company's products and services, but also from its intrinsic social issues and how it deals with crises (Keller and Lehmann, 2006). Brand associations are largely shaped by brand awareness, being stronger when based on many consumer experiences or consistent exposure to marketing communication. The more consumers are familiar with a brand, and the more positive their associations are, the more favorable are the attitudes towards the brand – and the higher is the brand equity (Yoo, Donthu and Lee, 2000; Grohs et al., 2016).

Relationship between misinformation and brand association

In the contemporary digital ecosystem, exposure to misinformation has emerged as a significant variable influencing how individuals interact with and evaluate brands. This flow of disinformation may affect user sentiment and brand association, which are central components of how consumers evaluate brands in online contexts. User sentiment, in particular, has become increasingly measurable through computational techniques such as sentiment analysis, which detects emotional valence in user-generated content and reflects broader reputational patterns (Kim & Lim, 2021; Stewart & Arnold, 2018). When exposed to *fake news* about a brand, consumers may experience a shift in emotional response—typically a move toward negative sentiment—which in turn alters how the brand is cognitively stored and retrieved in memory.

This conceptual framework illustrates the hypothesized relationship between exposure to misinformation and brand perception, mediated by user sentiment on social media platforms. Grounded in theories of brand equity (Aaker, 1996; Keller, 1993) and digital communication (Wardle & Derakhshan, 2017; Vosoughi et al., 2018a), the model proposes that disinformation, when encountered in digital environments, shapes user-generated sentiment, which in turn influences how brands are cognitively and affectively evaluated by consumers.

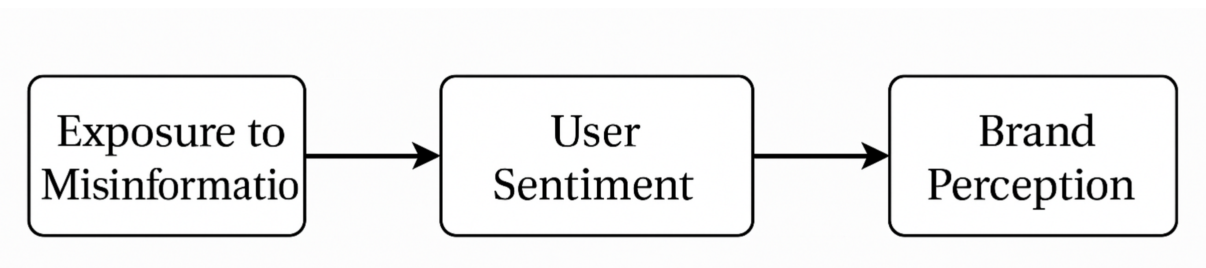

Figure 1: Conceptual framework illustrating the effect of misinformation exposure on brand perception through user sentiment.

Moreover, in emerging markets like Brazil, where concern about misinformation ranks among the highest globally (Newman et al., 2020), this relationship becomes even more critical to explore. The speed and volume of disinformation exposure increase the likelihood of sentiment volatility and reputational instability—factors that are particularly acute in platform-centric environments. As Dutta and Pullig (2011) note, brand crises can negatively affect core dimensions of brand equity, including brand perception, underscoring the need for robust frameworks to monitor and respond to misinformation-fueled crises. Based on the discussion, the study presents the following research question: How does exposure to misinformation influence user sentiment and brand perception on social media platforms?

Digital listening and sentiment analysis

The process of listening to the audience on social media is called *social listening* (Edwards, 2011; Stewart & Arnold, 2018).Online monitoring is carried out to identify and evaluate what is being said about a company, individual, product, or brand on digital channels. This practice was quickly perceived as a benefit, being able to listen to customers' opinions. This monitoring is defined by Stewart and Arnold (2018, p. 86) as “an active process of attending, observing, interpreting and responding to a variety of stimuli through electronic channels and social media” which, within its spectrum, includes sentiment analysis of user-generated content.

Sentiment analysis is defined as a computational study of opinions, feelings, and emotions expressed in texts (Kim & Lim, 2021a; B. Liu, 2010; Ortigosa et al., 2014) carried out through a generally unsupervised method, which uses a lexicon of words containing polarity values, such as neutral, positive or negative, to understand changes in loyalty to a particular brand (Neri et al., 2012; Vosoughi et al., 2018b) it can also be used to classify and quantify opinions about individual human behavior about a subject (Araque et al., 2017; Byun et al., 2012; Dey et al., 2016; Hasan et al., 2018b; Kim & Lim, 2021a)

RESEARCH METHOD

Research techniques

The techniques started with *tweets* collection through data mining using the Python language with direct requests through an application programming interface (API). Each of these contents was automatically translated from Brazilian Portuguese into English using the Google Translate API. After a sample verification to validate the method, each translation went through the sentiment analysis stage, when text mining technology was used to identify people's opinions, feelings, and evaluations about entities such as brands and products, classifying a given text through a lexicon-based approach that uses a dictionary containing information about the polarities of words related to different feelings (Kim & Lim, 2021a)Figure 2 summarizes the script of techniques used in this study and defines the research protocol.

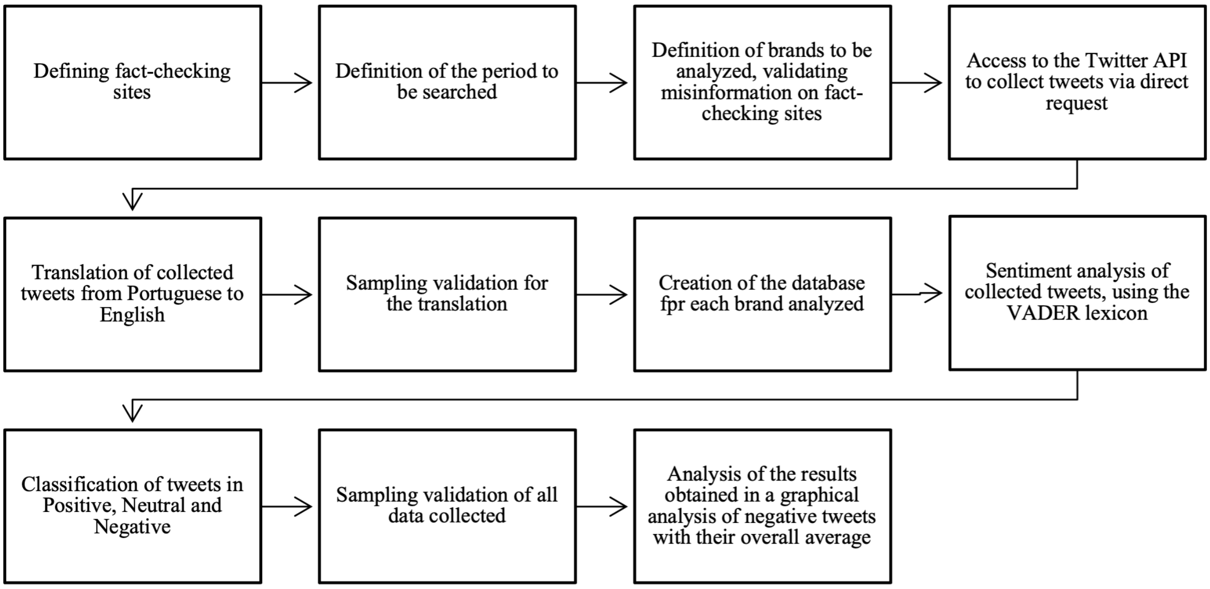

Figure 2: General methodology structure. Adopted and modified by the author based on Hasan (2018)

Definition of brands and periods analyzed

The criteria for choosing misinformation actions were through the analysis of the news that cited retail brands operating in Brazil and was recognized as misinformation by fact-checking sites so that such an act would be free of any intervention by researchers, and which already had some counter-information action. To define the fact-checking for this research, the independent Brazilian platforms in operation in 2022 were analyzed and considered Boatos.org, Aosfatos.org, and e-farsas.com.br. During this definition, project sites such as Agência Lupa, Fato ou Fake, Estadão Verify, and Projeto Comprova had little relevance in terms of visits, or belong to larger companies, not fitting in the correspondence of independent verifiers.

The chosen time frame to obtain the sample was between January 2016 to February 2022, using the dates in which a relevant amount of search for the term *"fake news"* was found in the *Google Trends tool,* which evaluates the popularity of specific terms used in the Google search tool, providing a time series index of the volume of searches performed by users to evaluated result showed a beginning of popularity in searches from 2016. It is important to note that this study chose to define time using the term "*fake news"* due to its popularity.

Data collection

For the collection of posts from the social media platform Twitter, a time frame of thirty days before and thirty days after the misinformation related to each brand was determined, considering the speed of emergence of information on social media, as well as the amount of content generated (Crawford, 2009; Macdonald, 2020)To get access to the Twitter *content* platform, it was necessary to request academic access to their central team, upon receipt of unique keys, given on February 2, 2021.

Sentiment analysis

Among the available methods for carrying out sentiment analysis of *tweets*, according to (Kim & Lim, 2021a) and (Park et al., 2015) the text classification approach is not appropriate for the format of this research, as general classifications may not be reliable in analyzes of short messages used in social networks, especially X (Twitter). Therefore, this study adopted a lexicon-based sentiment analysis of words contained in digital libraries. Research by (Hutto &

Gilbert, 2014) and (Araque et al., 2017) tested and compared several of these, as well as (Kim & Lim, 2021a) and (Park et al., 2015)VADER *(valence aware dictionary and sentiment reasoner)* lexicon for performing better even with short messages such as *tweets*, since VADER's sentiment lexicon is tuned to sentiment in microblogs such as Twitter and has over 4,200 *tweet-like* messages, originally inspired by actual comments from this social network. For the entire use of its library, it was decided to use the original English lexicon provided openly by VADER, which added a step to the method of this study, validating the translations of the collected tweets, through a connection *via* API of the Google Translate tool right after the collection stage, requiring manual validation of the process, detailed in the next topic.

The official VADER library has resources of natural language processing fundamentals that are used to calculate the polarity and subjectivity of *tweets*. Process prepared for this social network, and automatically trained on billions of entries, with reports of up to 70% agreement when compared to assignments performed manually (Byun et al., 2012; Hasan et al., 2018b)

The environment created in Python performed sentiment analysis of each *tweet*, then classified its content as negative, neutral, or positive, according to the VADER lexicon. This value is negative when the sentence has a negative connotation, close to zero for neutral connotations, and with a positive value for connotations given as positive; thus, being possible to verify if each word in the collected *tweet* is contained in the lexicon and assign a sentiment score. If the word is not contained in the VADER library, the environment performs a search in complementary lists, adding them and learning with each new action — *hashtags* and symbols are ignored by the analysis system, which did not generate changes in results in other studies due to the scope of the chosen library (Araque et al., 2017; Baccianella et al., 2010; Dey et al., 2016; Edwards, 2011; Loria et al., 2020; Parveen & Pandey, 2017; Ram et al., 2024; Vosoughi et al., 2018b)

Each collected *tweet* was scored with values that varied according to VADER lexicon ratings, indicating negative, neutral, and positive sentiment (Hutto & Gilbert, 2014; Park et al., 2015)When VADER analyzes a snippet of any text in English, it checks for the presence of any of the words in its lexicon and produces sentiment metrics from these word rankings that represent the proportion of the text that fall into the categories. It is worth noting the efficiency of VADER in the analysis by (Hutto & Gilbert, 2014)especially in the simple form in which the lexicon is updated automatically (*machine learning*), including current expressions, slang, symbols, *emojis*, punctuation, considering, also, words inserted within contexts depending on the way they are written, such as the capitalization of any word, which increases the intensity of a positive and negative score. VADER also considers words preceded by changes in pitch; for example, “extremely bad” increases the negative intensity of a sentence, but the sentence “kind of bad” receives a lower score. For the validation of the translation and the final classification of sentiments, an analysis was performed manually and by sampling during the initial test.

Data validation

A final sample of 30 tweets was used, based on (Hennink et al., 2017; Marshall et al., 2015; Suddaby & Suddaby, 2006) studies, in which theory and practice are evaluated on the saturation level within a sample universe in search of a solid and reliable validation. Thus, 30 (thirty) *tweets* were manually validated, randomly selected in Microsoft Excel software. In these tweets, both the quality of the translation of the contents from Portuguese into English through the Google Translate API, as well as the sentimentalization classifications carried out by the VADER library, were analyzed. As a result, 100% validation was obtained in the translations carried out automatically and in literal translation format of the contents, which collaborated with the classification process of sentiments. A result of 20% of errors in the content classification was obtained; perceived precisely in tweets involving irony or very specific expressions and therefore, literally translated and not yet learned by the learning intelligence of the VADER library. This result follows the (Hasan et al., 2018b)which validated the sentimentalization library learning method with a 70% assertiveness rate.

All content was acquired in groups, as per API limitation. Data storage was done using Microsoft Excel, which allowed representation by spreadsheets and dynamic graphics. The processing of the content of each tweet consisted of the translation, done using *Deep Google Translator*'s *Translator*, which did it in clusters of 200 sentences at a time, and sentence sentiment polarization through segregation and *tokenization,* which classified them according to the intensity level of the NLTK polarity classifier, which made use of the VADER lexicon. So, to analyze the effect of misinformation on the perception of the studied brands on Twitter, a graphical analysis was performed highlighting the number of negative tweets each day compared to a general mean line.

ANALYSIS

The presented protocol was followed, within a 30 (thirty) days window before and after the counterinformation date. The collection file included the specific keywords of *tweets* in Portuguese, translated into English, with general information about the misinformation, the fact-checking website, which the fake news action was identified, and the classification final sentiment of each collected content. Therefore, 230.569 tweets were collected involving four retail companies: Skol, AMBEV, McDonald’s and Coke.

Statistical Process Control

Given the growing need to monitor public reactions to misinformation on digital platforms, Statistical Process Control (SPC) emerges as a suitable methodological approach for detecting behavioral patterns and disruptions caused by specific events. Originally developed in the field of product quality management, SPC has since been adapted to other domains. Montgomery (2009) proposed a theory for the creation of control charts, which has been applied by various scholars, including Kim & Lim (2021), who monitored sentiment in mobile app user reviews using SPC and sentiment analysis.

SPC-based methods offer the potential to examine the dynamic aspects of service quality, particularly by tracking changes in quality over time. Common cause variation is indicated by a stable pattern, whereas special cause variation is evidenced by a disruption in the repetitive pattern, signaling that something has changed in the process. Based on the methodology presented by Montgomery (2009) and employed in Kim & Lim (2021), this study applied control charts to the dataset collected, generating three lines parallel to the x-axis: the center line (CL), derived from the mean value, the upper control limit (UCL), and the lower control limit (LCL).

In the construction of these limits, the use of 3σ (standard deviations) is common, supported by consistent results in practice. According to Anderson et al. (2019), the standard deviation represents the distance of each observation from the mean, and as emphasized by Bittencourt (Bittencourt, 2014, p. 39), “it provides us with an indication of what occurs between the two extremes. Therefore, the standard deviation is the measure of how much the observed values vary around the mean.”

To determine whether the collected reactions were within statistical control, this study followed the framework established by Werkema (1995) evaluating the presence of points outside the defined control limits. In this context, a sensitivity parameter was adopted — a constant representing the width of the control band, defined by the number of standard deviations required to ensure a given confidence interval (Kim & Lim, 2021b)

As Montgomery (2009) suggests in quality control applications, any data point that falls outside the control limits — either above the UCL or below the LCL — is interpreted as evidence that the process is out of control, requiring investigation and corrective action. While the UCL was calculated as:

$$UCL = \bar{x}+3\sigma$$

the LCL followed the standard formula:

$$UCL = \bar{x}-3\sigma$$

However, since the observed variable is a count of daily negative tweets, negative values would be nonsensical. Thus, the LCL was truncated at zero:

$$LCL=max(0,\bar{x}-3\sigma)$$

This truncation approach is consistent with standard SPC applications for count-based data, particularly in scenarios where negative values are not meaningful, such as real-time social media monitoring. For the analysis conducted, only the variance of negative tweets was evaluated. Neutral and positive tweets were excluded, resulting in a number of analyzed variables equal to one (n = 1). Consequently, the general mean used to define the center line was:

$$\bar{x}=average\ number\ of\ negative\ tweets$$

This led to the final formula used to construct the statistical process control chart reflecting the variation of negative tweets within a given observation period. To initiate the implementation of the selected software tools and validate the methodological protocol described above, an initial test was conducted, applying each of the steps previously outlined and presenting the first data collected.

Affected brands

SKOL

The case involving the SKOL beer brand revolved around misinformation alleging that the beverage had caused hospitalizations and cancer, as reported by supposed medical sources. According to the fact-checking website Boatos.org, which verified and debunked the claim on January 17, 2018, the message falsely stated: “Skol advertising stopped on TV because the beer is contaminated, causing tumors and kidney problems. Doctors have recommended you not to drink Skol.” The report clarified that this narrative had resurfaced on multiple occasions and had even mentioned other brands in previous years. Boatos.org confirmed the falsity of the information based on a note issued by the brand's press office.

To assess the public reaction to this misinformation, the period analyzed encompassed 30 days before and after the publication of the fact-checking article, covering the range from December 17, 2017, to February 17, 2018. The keywords used in the data collection included

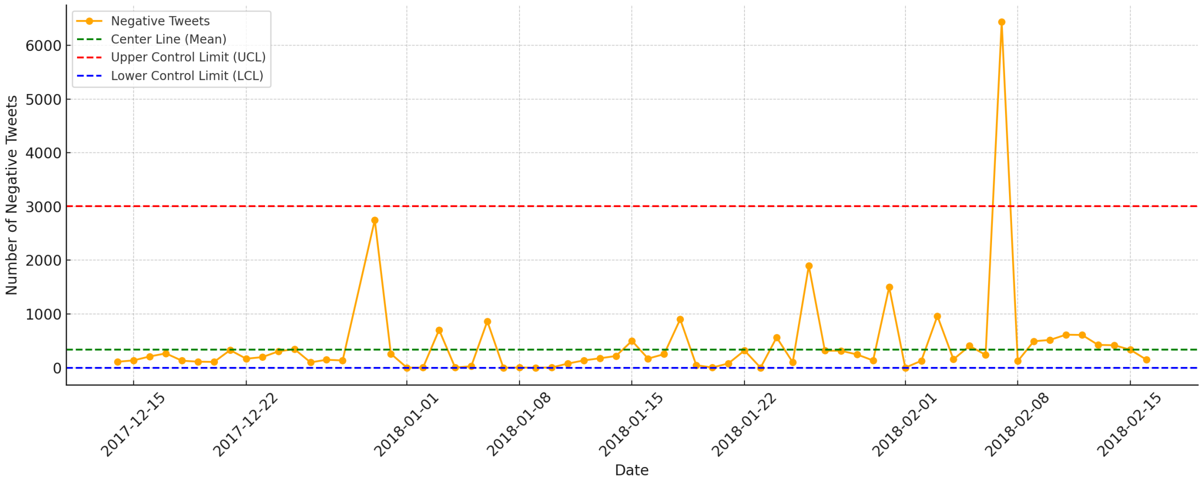

Figure 3: Chart of negative tweets related to Skol between December 17, 2017, and February 17, 2018. The chart shows the number of daily negative tweets referencing Skol during the monitoring period. The center line is set at 335.8, with an upper control limit of 3,007.71 and a lower limit truncated at 0.

variations of the brand name: Skol, "Skol beer", skol, SKOL, and "skol beer".

A total of 82,839 tweets were collected and categorized using sentiment analysis: 27,351 were classified as negative (33% of the total), 33,836 as positive (41%), and 21,652 as neutral (26%). The subset of negative tweets was selected for a more in-depth analysis using Statistical Process Control (SPC). The daily average number of negative *tweets* during the observed period was 335.8, with a calculated standard deviation of 890.64. Based on these values, the Upper Control Limit (UCL) was determined as 3,007.71, and the Lower Control Limit (LCL) was truncated at 0, as negative tweet counts are not semantically or statistically viable.

Figure 3 presents the resulting control chart, plotting the daily volume of negative *tweets* against the statistical thresholds. The chart highlights multiple peaks of negative sentiment, including at least one outlier well above the UCL, which indicates a special cause variation in public reaction. These peaks were noticed, although the volume of *tweets* referencing the specific misinformation incident may have been relatively small when manually reviewed, the overall surge in negative sentiment during the analyzed window was statistically significant but not related to the misinformation case.

MCDONALD'S

The case involving the McDonald's brand emerged from a misinformation claim circulated on July 19, 2016, and subsequently debunked by the fact-checking website Boatos.org. The counter-information was delivered through a note issued by the brand's press office and shared across other digital platforms. The misinformation alleged: "Federal Police and Procon intercept dog meat that would be used in McDonald's restaurants — the images shock everyone." Upon investigation, the images were proven to be unrelated to the claim, and no reputable media outlets reported any information supporting the allegation.

The monitoring period for this case covered the window from June 16 to August 18, 2016 — 30 days before and after the date of the counter-information. The keywords used to extract tweets associated with the brand included variations and misspellings, such as: Mc'Donalds, mc'donalds, méqui, "méqui donalds", "mequi donaldis", "mequi donalds", mcdonalds, "mc donalds", Mc'Donaldis, MCDONALDS, "MC DONALDS", and "MEQUI DONALDS".

The chart (Figure 3) illustrates the daily fluctuations in negative tweets during the observed period, benchmarked against the established control limits. While most values fall

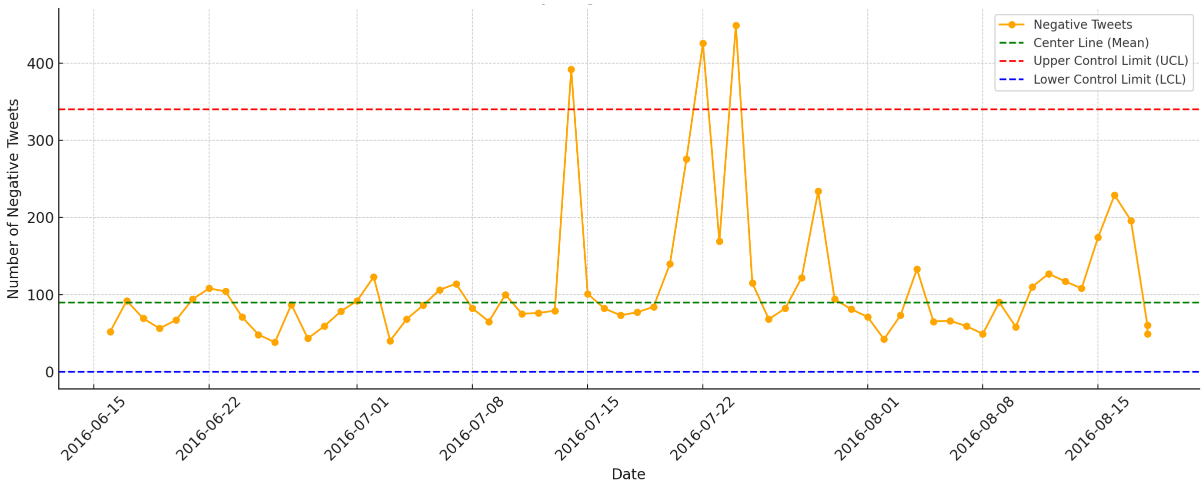

Figure 4: Chart of negative tweets related to McDonald's between June 16 and August 18, 2016. This figure displays the daily fluctuations in negative sentiment about McDonald's. The data maintain stability around a mean of 89.7, with no observations exceeding the control limits.

within the expected range of variation, a singular spike above the UCL was observed, indicating a special cause variation, related to the misinformation. This pattern is consistent with both the manual review of *tweet* content and the graphical interpretation, reinforcing that public reaction to the misinformation was delayed, and that significant sentiment peaks occurred only at moments temporally disconnected from the debunking event.

AMBEV

The case involving the AMBEV brand originated from the circulation of a video, allegedly recorded by an employee, showing pigeons being ground along with barley during the beer production process. The misinformation was debunked on March 18, 2017, by the fact-checking website Boatos.org, which identified the video as a recycled clip from 2016 depicting a grain silo in Russia, as reported by several international media outlets. AMBEV responded with a note issued via its press office, published on various websites. As a follow-up counter-information action, the brand released a satirical advertisement on April 1, 2017—April Fools' Day—featuring a video that addressed the impact of misinformation.

The case involving the AMBEV brand originated from the circulation of a video, allegedly recorded by an employee, showing pigeons being ground along with barley during the beer production process. The misinformation was debunked on March 18, 2017, by the fact-checking website Boatos.org, which identified the video as a recycled clip from 2016 depicting a grain silo in Russia, as reported by several international media outlets. AMBEV responded with a note issued via its press office, published on various websites. As a follow-up counter-information action, the brand released a satirical advertisement on April 1, 2017—April Fools' Day—featuring a video that addressed the impact of misinformation.

To evaluate the public's reaction, the analyzed period covered the 60 days between February 16 and April 17, 2017. The keywords selected for data collection included: AMBEV, ambev, "ambev brewery", and " ambev brewery".

A total of 9,756 tweets were collected: 3,663 were classified as negative (37%), 2,887

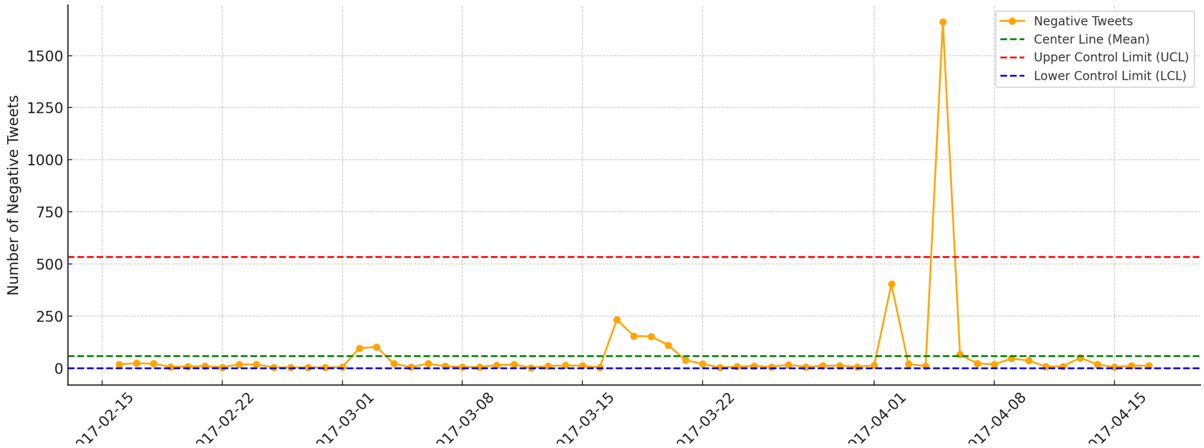

Figure 5: chart of negative tweets related to AMBEV between February 16 and April 17, 2017. The chart illustrates the number of negative tweets directed at AMBEV. With an average of 58.3 tweets per day and a UCL of 532.34, all data points remain within control.

as neutral, and 3,206 as positive. For deeper analysis, only the negative *tweets* were processed using the Statistical Process Control (SPC) methodology. The average number of daily negative tweets was 58.3, with a standard deviation of 158.01. The Upper Control Limit (UCL) was calculated as 532.34, while the Lower Control Limit (LCL) was truncated at 0, due to the nature of count-based data.

The chart (Figure 5) illustrates the daily distribution of negative *tweets* in relation to the calculated control limits. Throughout the monitored period, all values remained within the control boundaries, and no points surpassed the UCL, indicating that the reactions were statistically within expectations and did not constitute special cause variations. Nonetheless, small peaks were identified around key dates. A moderate increase in negative *tweets* occurred prior to the debunking of the misinformation, aligning with its initial circulation. A subsequent drop was observed following the publication of the press note and the brand's counter-campaign. Another small rise on April 2 coincided with renewed attention due to the brand's own video initiative, suggesting the counter-information may have inadvertently revived the topic. A final spike on April 6, however, was unrelated to the misinformation incident, instead referring to an older, unrelated controversy about corporate misconduct in 2016.

COCA-COLA

The case involving Coca-Cola centers on misinformation suggesting that the Brazilian government, under President Michel Temer, was negotiating the concession of the Guarani Aquifer to the two multinational companies. This piece of misinformation was debunked by the fact-checking platform Boatos.org on March 1, 2018. The website's investigation concluded that no official source confirmed such an agreement, and the Brazilian government itself denied the claim via its official Twitter account (Governo Federal do Brasil, 2018).

The monitored period was defined as 30 days before and after the publication of the counterinformation. Thus, the analyzed timeframe spanned from February 1 to April 1, 2018. The keywords used in this analysis were derived from variations of the Coca-Cola brand name, including coca-cola, cocacola, Coca-Cola, COCA-COLA, Coca-cola, coca-Cola, coca cola, Coca Cola.

A total of 84,514 tweets were collected and categorized using sentiment analysis into neutral, positive, and negative sentiment. From this total, 13,489 tweets were classified as

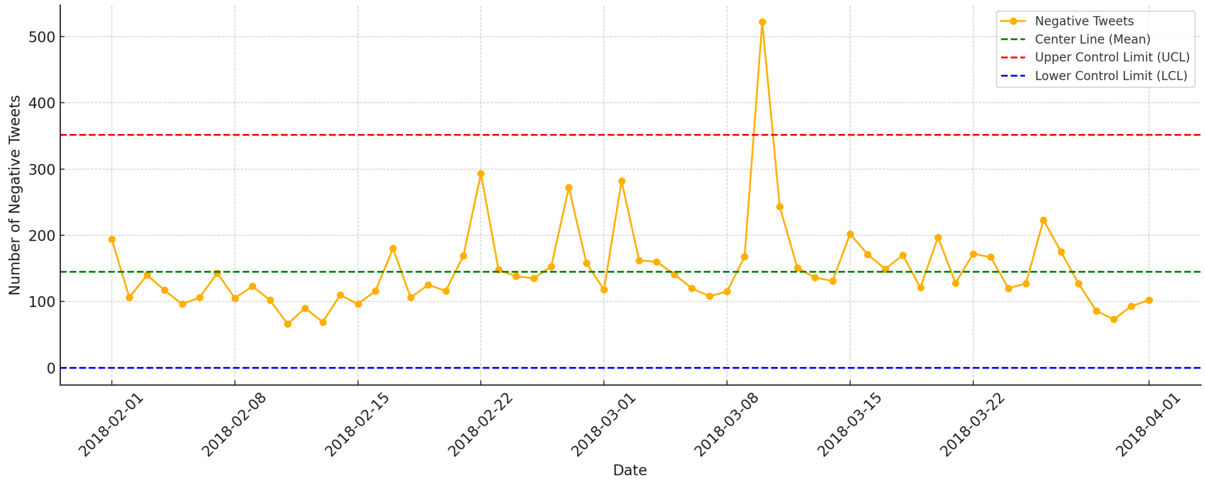

Figure 6: chart of negative tweets related to Coca-Cola between February 1 and April 1, 2018. This figure presents the daily volume of negative tweets mentioning Coca-Cola, monitored before and after the spread of misinformation about the Guarani Aquifer. The center line represents the average ($\bar{x}$ = 144.8), with upper and lower control limits set at ±3 standard deviations.

negative and selected for in-depth analysis. Following the SPC methodology, the average number of daily negative *tweets* during the monitored period was calculated as 144.8, with a standard deviation of 68.95. These values were used to construct a control chart, defining a Center Line (CL) at the mean, an Upper Control Limit (UCL) at 351.65, and a Lower Control Limit (LCL) at 0.

Figure 6 presents the resulting control chart, which visually tracks the daily volume of negative tweets over the analysis period. Spikes above the UCL indicate statistically significant anomalies that may be attributed to special cause variations, such as reactions to the misinformation event.

From a methodological standpoint, the analytical framework employed in this study—centered on Statistical Process Control (SPC) combined with sentiment analysis of social media data—demonstrates strong indicators of validity and reliability. In terms of content validity, the selection of tweets filtered by brand-specific keywords, aligned with each misinformation event, ensures that the data reflects public reactions to precisely defined informational contexts. Brand associations validity is supported by the operationalization of negative sentiment as a behavioral indicator of reputational impact, a concept reinforced in the literature by Kim & Lim (2021) and Ramanathan et al. (2018), who use similar metrics to assess perception shifts in digital environments. Which is strengthened through temporal alignment between misinformation events, fact-checking dates, and observable variations in tweet volumes, which are benchmarked against statistically robust control limits following Montgomery (2009).

CONTRIBUTION

This study set out to examine how exposure to misinformation affects user sentiment and brand perception within the context of social media platforms, particularly focusing on the Brazilian market. The analysis of over 230,000 tweets across four major retail brands — Skol, AMBEV, McDonald's, and Coca-Cola — revealed sentiment variations following misinformation events. Using a lexicon-based sentiment analysis technique (Kim & Lim, 2021) combined with Statistical Process Control (Montgomery, 2009), this research worked to empirically validated a conceptual framework that positions user sentiment as a mediating between misinformation exposure and shifts in brand perception.

The results demonstrate that misinformation can trigger an increase in negative sentiment toward brands, although the intensity and timing of the response may vary depending on the nature of the misinformation and the brand's counterinformation strategy. Notably, brands that engaged in structured counterinformation responses—such as AMBEV's satirical campaign or Coca-Cola's institutional rebuttal—were more effective in stabilizing sentiment trajectories. These findings align with Dutta and Pullig's (2011) assertion that crisis communication strategies are essential in mitigating damage to brand equity.

Moreover, the empirical evidence shows that fake news can distort consumer brand associations, echoing previous findings by Keller (1993), Wang (2016), and Borges-Tiago et al. (2020). As user sentiment becomes increasingly measurable through digital listening techniques (Stewart & Arnold, 2018), its role as an early indicator of reputational impact becomes more evident. The observed peaks in negative sentiment, when exceeding statistical control limits, suggest that misinformation acts as a special cause of reputational variation — not merely as background noise in public discourse.

Final considerations

This study contributes to the theoretical and practical understanding of how reputational narratives are shaped, distorted, and potentially recovered in algorithmically driven environments. It underscores the importance of active social listening, timely fact-checking, and strategic counterinformation in protecting brand equity in a context increasingly shaped by information disorder. Future research should explore longitudinal effects of repeated misinformation exposure and the comparative effectiveness of different crisis response strategies across markets and industries.

REFERENCES

Aaker, D. A. (1996). Measuring Brand Equity Across Products and Markets. *California Management Review*, *38*(3), 102–120. https://doi.org/10.2307/41165845

Anderson, D. R., Sweeney, D. J., Williams, T. A., Camm, J. D., & Cochran, J. J. (2019). *Estatística aplicada a adminstração e economia* (8a edição). Cengage.

Araque, O., Corcuera-Platas, I., Sánchez-Rada, J. F., & Iglesias, C. A. (2017). Enhancing deep learning sentiment analysis with ensemble techniques in social applications. *Expert Systems with Applications*, *77*, 236–246. https://doi.org/10.1016/j.eswa.2017.02.002

Baccianella, S., Esuli, A., & Sebastiani, F. (2010). SENTIWORDNET 3.0: An enhanced lexical resource for sentiment analysis and opinion mining. *Proceedings of the 7th International Conference on Language Resources and Evaluation, LREC 2010*, *0*, 2200–2204.

Bittencourt, H. R. (2014). Controle Estatístico da Qualidade. In *Pontifícia Universidade Católica do Rio Grande do Sul*. www.pucrs.br/famat/helio

Borges-Tiago, T., Tiago, F., Silva, O., Guaita Martínez, J. M., & Botella-Carrubi, D. (2020). Online users' attitudes toward fake news: Implications for brand management. *Psychology and Marketing*, *37*(9), 1171–1184. https://doi.org/10.1002/mar.21349

Brennen, J. S., Simon, F. M., Howard, P. N., & Nielsen, R. K. (2020). Types, Sources, and Claims of COVID-19 Misinformation. *Factsheet*, *April*(2020), 1–13.

Byun, C., Lee, H., & Kim, Y. (2012). Automated twitter data collecting tool for data mining in social network. *Proceeding of the 2012 ACM Research in Applied Computation Symposium, RACS 2012*, 76–79. https://doi.org/10.1145/2401603.2401621

Crawford, K. (2009). Following you: Disciplines of listening in social media. *Continuum*, *23*(4), 525–535. https://doi.org/10.1080/10304310903003270

DataReportal, & Hootsuite. (2025). *Global digital report 2025*. Https://Datareportal.Com.

Dentith, M. R. X. (2016). What makes 'fake news' 'fake' news. *Public Reason*, *8*, 65–79.

Dey, L., Chakraborty, S., Biswas, A., Bose, B., & Tiwari, S. (2016). Sentiment Analysis of Review Datasets Using Naïve Bayes' and K-NN Classifier. *International Journal of Information Engineering and Electronic Business*, *8*(4), 54–62. https://doi.org/10.5815/ijieeb.2016.04.07

Edwards, R. (2011). Listening and message interpretation. *International Journal of Listening*, *25*(1–2), 47–65. https://doi.org/10.1080/10904018.2011.536471

Hasan, A., Moin, S., Karim, A., & Shamshirband, S. (2018a). Machine Learning-Based Sentiment Analysis for Twitter Accounts. *Mathematical and Computational Applications*, *23*(11), 1–15. https://doi.org/10.3390/mca23010011

Hasan, A., Moin, S., Karim, A., & Shamshirband, S. (2018b). Machine Learning-Based Sentiment Analysis for Twitter Accounts. *Mathematical and Computational Applications*, *23*(11), 1–15. https://doi.org/10.3390/mca23010011

Heath, R. L. (2006). Onward into more fog: Thoughts on public relations' research directions. *International Journal of Phytoremediation*, *21*(1), 93–114. https://doi.org/10.1207/s1532754xjprr1802_2

Hennink, M. M., Kaiser, B. N., & Marconi, V. C. (2017). Code Saturation Versus Meaning Saturation: How Many Interviews Are Enough? *Qualitative Health Research*, *27*(4), 591–608. https://doi.org/10.1177/1049732316665344

Hutto, C. J., & Gilbert, E. (2014). VADER: A parsimonious rule-based model for sentiment analysis of social media text. *Proceedings of the 8th International Conference on Weblogs and Social Media, ICWSM 2014*, 216–225.

Jahng, M. R., Lee, H., & Rochadiat, A. (2020). Public relations practitioners' management of fake news: Exploring key elements and acts of information authentication. *Public Relations Review*, *46*(2). https://doi.org/10.1016/j.pubrev.2020.101907

Kalsnes, B. (2018). Fake News. *Oxford Research Encyclopedia of Communication*, *March*, 1–19. https://doi.org/10.1093/acrefore/9780190228613.013.809

Keller, K. L. (1993). Conceptualizing, Measuring, and Managing Customer-Based Brand Equity. *Jornal of Marketing*, *57*, 1–22. http://content.epnet.com/ContentServer.asp?T=P&P=AN&K=13046790&EbscoContent=dGJyMNLe80SeqK84yOvqOLCmr0mep7BSrq64SbKWxWXS&ContentCustomer=dGJyMPGusU6wp7ZIuePfgeyx%2BEu3q64A&D=bth%5Cnpapers3://publication/uuid/A5A4CEB4-6808-45B0-A2C8-1664208133CF

Kent, M. L., Taylor, M., & Taylor, M. (2002). Toward a dialogic theory of public relations. In *Public Relations Review* (Vol. 28).

Kim, J., & Lim, C. (2021a). Customer complaints monitoring with customer review data analytics: An integrated method of sentiment and statistical process control analyses. *Advanced Engineering Informatics*, *49*. https://doi.org/10.1016/j.aei.2021.101304

Kim, J., & Lim, C. (2021b). Customer complaints monitoring with customer review data analytics: An integrated method of sentiment and statistical process control analyses. *Advanced Engineering Informatics*, *49*, 101304. https://doi.org/10.1016/j.aei.2021.101304

Kwak, H., Lee, C., Park, H., & Moon, S. (2009). What is Twitter, a Social Network or a News Media? *Conference: Proceedings of the 19th International Conference on World Wide Web*, 26–30. https://doi.org/10.1145/1772690.1772751

Lazer, D. M. J., Baum, M. A., Benkler, Y., Berinsky, A. J., Greenhill, K. M., Menczer, F., Metzger, M. J., Nyhan, B., Pennycook, G., Rothschild, D., Schudson, M., Sloman, S. A., Sunstein, C. R., Thorson, E. A., Watts, D. J., & Zittrain, J. L. (2018). The science of fake news. *Science*, *359*(6380), 1094–1096. https://doi.org/10.1126/science.aao2998

Liu, B. (2010). Sentiment analysis and subjectivity. *Handbook of Natural Language Processing, Second Edition*, 627–666.

Liu, X. (2019). A big data approach to examining social bots on Twitter. *Journal of Services Marketing*, *33*, 369–379. https://doi.org/10.1108/JSM-02-2018-0049

Loria, S., Keen, P., Honnibal, M., Yankovsky, R., Karesh, D., & Dempsey. (2020). *Textblob: Simplified Text Processing*. https://textblob.readthedocs.io/en/dev/

Macdonald, K. M. (2020). *Listening in : Investigating Social Media Activity in the Streaming Service Industry*. University of New Hampshire.

Marshall, B., Cardon, P., Poddar, A., Fontenot, R., Marshall, B., Cardon, P., Poddar, A., & Does, R. F. (2015). *Does Sample Size Matter in Qualitative Research ?: A Review of Qualitative Interviews in is Research DOES SAMPLE SIZE MATTER IN QUALITATIVE RESEARCH ?: A REVIEW OF QUALITATIVE INTERVIEWS IN IS RESEARCH*. *4417*. https://doi.org/10.1080/08874417.2013.11645667

Marwick, A., & Lewis, R. (2017). Media manipulation and disinformation online. *Data & Society*, 1–104.

Montgomery, D. C. (2009). Introduction To Statistical Quality Control. In Inc. John Wiley & Sons (Ed.), *Plastics and rubber international* (Vol. 10, Issue 1). https://doi.org/10.2307/2988304

Neri, F., Aliprandi, C., Capeci, F., Cuadros, M., & By, T. (2012). Sentiment analysis on social media. *Proceedings of the 2012 IEEE/ACM International Conference on Advances in Social Networks Analysis and Mining, ASONAM 2012*, *August 2012*, 919–926. https://doi.org/10.1109/ASONAM.2012.164

Newman, N., Richard Fletcher, W., Schulz, A., Andı, S., & Kleis Nielsen, R. (2020). *Reuters Institute Digital News Report 2020*. https://reutersinstitute.politics.ox.ac.uk/sites/default/files/2020-06/DNR_2020_FINAL.pdf

Ortigosa, A., Martín, J. M., & Carro, R. M. (2014). Sentiment analysis in Facebook and its application to e-learning. *Computers in Human Behavior*, *31*(1), 527–541. https://doi.org/10.1016/j.chb.2013.05.024

Pace, S., Balboni, B., & Gistri, G. (2017). The effects of social media on brand attitude and WOM during a brand crisis: Evidences from the Barilla case. *Journal of Marketing Communications*, *23*(2), 135–148. https://doi.org/10.1080/13527266.2014.966478

Park, K., Kim, J., Park, J., Cha, M., Nam, J., Yoon, S., & Rhim, E. (2015). Mining the minds of customers from online chat logs. *International Conference on Information and Knowledge Management, Proceedings*, *19-23-Oct-*, 1879–1882. https://doi.org/10.1145/2806416.2806621

Parveen, H., & Pandey, S. (2017). Sentiment analysis on Twitter Data-set using Naive Bayes algorithm. *Proceedings of the 2016 2nd International Conference on Applied and*

Theoretical Computing and Communication Technology, ICATccT 2016, 416–419. https://doi.org/10.1109/ICATCCT.2016.7912034
Ram, N. R., Gautum, S., Jadeja, A., Joisar, H., & Rathore, N. (2024). Social Media Sentiment Analysis Using Twitter Dataset. *2024 1st International Conference on Cognitive, Green and Ubiquitous Computing, IC-CGU 2024*. https://doi.org/10.1109/IC-CGU58078.2024.10530694
Shocker, A. D., & Aaker, D. A. (1993). Managing Brand Equity. *Journal of Marketing Research*, *30*(2), 256. https://doi.org/10.2307/3172832
Stewart, M. C., & Arnold, C. L. (2018). Defining Social Listening: Recognizing an Emerging Dimension of Listening. *International Journal of Listening*, *32*(2), 85–100. https://doi.org/10.1080/10904018.2017.1330656
Suddaby, R., & Suddaby, R. O. Y. (2006). *From the Editors : What Grounded Theory Is Not FROM THE EDITORS : WHAT GROUNDED THEORY IS NOT*. *49*(4), 633–642.
Tandoc, E. C., Lim, Z. W., & Ling, R. (2018a). Defining "Fake News": A typology of scholarly definitions. *Digital Journalism*, *6*(2), 137–153. https://doi.org/10.1080/21670811.2017.1360143
Tandoc, E. C., Lim, Z. W., & Ling, R. (2018b). Defining "Fake News": A typology of scholarly definitions. *Digital Journalism*, *6*(2), 137–153. https://doi.org/10.1080/21670811.2017.1360143
UNESCO;, & University of Oxford; (2018). World Trends in Freedom of Expression and Media Development. Global Report 2017/2018. In *United Nations Educational,Scientific and Cultural Organization* (Vol. 7, Issue 17). https://doi.org/10.22201/ceiich.24485705e.2019.17.68086
Vosoughi, S., Roy, D., & Aral, S. (2018a). The spread of true and false news online. *Science*, *1151*(March), 1146–1151. https://science.sciencemag.org/content/359/6380/1146
Vosoughi, S., Roy, D., & Aral, S. (2018b). The spread of true and false news online. *Science*, *1151*(March), 1146–1151.
Wang, Y. (2016). Brand crisis communication through social media: A dialogue between brand competitors on Sina Weibo. *Corporate Communications*, *21*(1), 56–72. https://doi.org/10.1108/CCIJ-10-2014-0065
Wardle, C., & Derakhshan, H. (2017). Information Disorder: Toward an interdisciplinary framework for research and policy making. *Report to the Council of Europe*, 108. https://rm.coe.int/information-disorder-toward-an-interdisciplinary-framework-for-researc/168076277c
Werkema, M. C. C. (1995). *Ferramentas estatísticas básicas para o gerenciamento de processos* [Monography em Pt]. Universidade Federal de Minas Gerais.
Yoo, B., Donthu, N., & Lee, S. (2000). An examination of selected marketing mix elements and brand equity. *Journal of the Academy of Marketing Science*, *28*(2), 195–211. https://doi.org/10.1177/0092070300282002