Understanding the Dynamics of On-line Firestorms: Minimizing their Detrimental Impact on Firms' Brand Image and Market Performance

by

Louis LOISEL[1], Matthias NGUEND NJIKI[1], Karim ER-RACHDI[1], Thom TOUADERA[1], Sammy RABHINE[1], Matthieu SAJOT[2], Gábor NAGY[3], Jae Yun JUN KIM[4]

If you want to contact us: jae-yun.jun-kim@ece.fr, gnagy@inseec.com

[1]Master in quantitative and financial engineering, ECE Paris engineering school
[2]Master in information systems and big data, ECE Paris engineering school
[3]Ph.D. Marketing, Corvinus University of Budapest
[4]Ph.D. Applied Mathematics

Abstract

On-line firestorms can be defined as the sudden discharge of large quantities of messages containing negative WOM (word of mouth) and complaint behavior against a person, company, or group in social media networks. In these messages, intense indignation is often expressed, without pointing to an actual specific criticism. If not effectively managed, firestorms in firm-hosted on-line communities can be detrimental to organizations and their brands. Although organizations have been the target for criticism and complaints about, for example, their brands, products, practices, or actions, on-line technologies in social media have exponentiated the speed, range, and scale of diffusion, as well as the detrimental consequences for organizations. On-line communities represent complex social systems that require flexible simulation approaches to better understand the dynamics of conflict diffusion, conflict management, and the avoidance of on-line firestorms. If not effectively managed, such conflict in firm-hosted communities and resulting firestorms can be detrimental to organizations and their brands. This project aims to understand the dynamics of on-line firestorms and minimize their detrimental impact on firms' brand image and market or financial performance. More precisely, this research project aims to model the dynamics of on-line firestorms in firm-hosted on-line brand communities and of measuring their effects on brand image and market or financial performance.

Keywords: Word-of-mouth, social media, online firestorm, information diffusion, virality, identification and measurement of risks.

Introduction

The emergence of social networks has significantly changed the way companies communicate with their target audience. According to a study conducted by Hootsuite and We Are Social in 2020, approximately 3.6 billion people use social networks worldwide, which is almost 46 % of the world's population. Companies can therefore reach a huge audience using these platforms.

As a result, social networks have also become a key tool for companies in their marketing strategy. According to a study by Sprout Social in 2020, 92 % of companies use social networks to promote their business. Businesses can use social media to improve their brand awareness, increase engagement with consumers and increase their online visibility. The use of this new channel has a direct impact on sales. According to another Hootsuite report, businesses that use social media for sales have seen an average 18 % increase in revenue.

However, the use of social networks is thus a double-edged sword. Indeed, as explained previously, these platforms allow to reach massive amounts of potential customers but as a result consumers can now reach companies much more easily. In other words, as social media have massively increased both the reach and speed of dissemination of eWOM, they have accelerated and intensified customers' exposure to negative eWOM. It is now common for potential customers to seek out information about products and services, publicly available online opinions, ratings and reviews from other consumers (Chen, Y., & Xie, J. 2008). Today more than 85 % of online purchases are currently motivated by such reviews and ratings, the so-called electronic word-of-mouth (eWOM), which makes social media a particularly important (if not central) factor in corporate marketing communications (Forman et al. 2008).

It is interesting to use the example of Dell to illustrate our point, which gave rise to the term "Dell Hell", used to describe a situation where Dell customers expressed massive negative feedback on social networks and online forums about the quality of the company's products and services. It began in 2005, when a blogger named Jeff Jarvis began writing about his poor customer service experiences with Dell. Other customers then began sharing their own similar stories, and it quickly evolved into a public relations crisis for the company. This domino effect was cited as one of the main factors that led to a sharp drop in Dell's customer satisfaction rating, and even its share price. This example illustrates the considerable risks of eWOM: if the sentiment in eWOM becomes negative, it can spread like a firestorm, resulting in a significant loss of customers and damage to the company's reputation. The researchers coined the term ¨ online firestorm", which reflects the effect and impact on the affected companies. This term can be defined as follows: "Sudden discharge of large quantities of messages containing negative Word-of-Mouth and complaint behavior against a person, company, or group in social media networks. In these messages, intense indignation is often expressed, without pointing to an actual specific criticism" (Pfeffer et al. 2014, p.118). In recent years, companies in various sectors have suffered from these online firestorms and their negative consequences.

With the possibility of reaching thousands of potential customers in a very short time, it is very difficult to stop the spread of negative eWOM if the emergence of an online firestorm is detected too late. Therefore, the challenge in managing these events for a company is to be able to detect the possible transformation into an online firestorm before the consequences are irreparable, in order to calm the spread of negative eWOM. Due to the rapid nature and huge volume of eWOMs, this can only be achieved by automated, real-time detection approaches.

The research project's questions are as follows. (1) What is the dynamism between factors leading to the virality of negative customer posts and firm responsiveness in conflict diffusion (i.e., firestorms detection)? (2) What is the dynamism of conflict management in on-line brand communities with respect to negative WOM (i.e., firestorm mitigation)? (3) What is the dynamism of on-line firestorm avoidance within on-line brand communities (i.e., firestorm prevention)?

Methods

1.1 Social media

Today, social media are the predominant platforms for communication and interaction between companies and customers, as well as between customers themselves. We can classify social media into different distinct categories:

- Social networks: Facebook, Instagram, Twitter, LinkedIn, etc. These platforms allow users to create profiles, share content (text, images, videos) and connect with other users.
- Blogging platforms: WordPress, Blogger, etc. These platforms allow users to create and manage their own blogs.
- Professional networking platforms: LinkedIn, Viadeo, etc. These platforms allow professionals to connect with each other, share professional information and find jobs.
- Video sharing platforms: YouTube, TikTok, etc. These platforms allow users to upload and share videos with other users.

Social media can therefore be defined as follows: “a group of Internet-based applications that build on the ideological and technological foundations of Web 2.0 and that allow the creation and exchange of user-generated content” (Kaplan Haenlein, 2010, p. 61). The most popular online platforms for companies to carry out their marketing campaigns in recent times are Facebook, Instagram, Twitter, TikTok, YouTube and LinkedIn. What is particularly interesting about the use of social media is the creation of eWOM through their use, eWOM can be defined as follows: “any positive or negative statement made by potential, actual, or former customers about a product or company, which is made available to a multitude of people and institutions via the Internet” (Hennig-Thurau et al. 2004, p. 39).

In our research, we decided to focus on the use of Twitter for several reasons. Firstly, because the means of communication on this social network is mainly through text. Secondly, users on this social network do not hesitate to take sides, express themselves, criticise or support views, opinions and messages. Thirdly, worldwide, Twitter has 237.8 million ”monetisable active daily users”, in 2022 (August). Therefore, this social network is the best target to base our research on.

1.2 The data used

1.2.1 Twitter API and the data collected

This project aims to perform full-archive search of tweets related to vocalized criticism and complaints about firms’ products and actions and firms’ responses via Twitter. The microblogging service Twitter can play a critical role in the dynamics that lead to cascades of negative WOM effects within a short period of time (Pfeffer, Zorbach, and Carley 2014). Our research uses search tweets (i.e., full-archive data) for longitudinal studies via analyzing past topics and/ or events that mirror the ways firms had been targeted for criticism and complaints about, for example, their brands, products, practices or actions. Moreover, tweets related to criticism and complaints about firms’ products and actions will be investigated for intensity of high arousal, intensity of low arousal, linguistic style match, intensity of empathy, intensity of explanation, variation in firm responses, among others. In addition, tweets related to respective firm-hosted brand communities will be examined to derive measures such as brand community size, brand community attentiveness, brand community expressiveness, average tie strengths, brand reputation, brand familiarity, network clusters, unrestrained information flow, network-triggered decision processes, among others (i.e., independent variables). Our study aims to understand dynamics of firestorm conflict diffusion, conflict management, and avoidance, not Twitter itself as a subject.

In this study, we used the Twitter API to collect and analyze data on online customer engagement and conflict management in brand communities. The choice of this method was based on several factors, including its simplicity and effectiveness in obtaining a large amount of relevant data.

Twitter API (Application Programming Interface) is a set of tools and protocols that allow developers to interact with Twitter data, such as tweets, users and media.

Twitter is a widely-used social media platform with a vast amount of user-generated content, making it an ideal source of data for this research. The Twitter API provides access to this data in a structured and easy-to-use format, allowing for efficient data collection and analysis. Additionally, the API includes query parameters that enable us to filter and retrieve only the most relevant tweets, based on various criteria such as context annotation, keywords, location, and more. For example, we used the context annotation "47.10045225402"to retrieve tweets specifically related to the brand "Twitter".
The data were collected over a period of one year, from January 1st to December 31st, 2022. In total, we collected and analyzed over 218215 tweets related to the selected companies. This approach allows us to have a comprehensive understanding of the online customer engagement and conflict management patterns over a longer period of time. This can be beneficial for identifying trends, patterns, and seasonal variations that may not be as visible in real-time data collection.

Furthermore, by collecting data for a full year, we were able to have a large sample size and increase the robustness of our analysis. A large sample size can help to reduce sampling error and increase the generalizability of the results.

In addition, the API also has advanced features such as pagination, which enables us to retrieve a large amount of tweets even if it exceeds the rate limit. This is particularly useful for our research as we needed to collect a large sample of tweets for a period of one year.
Another benefit of using the Twitter API is the ability to easily access a wide range of data on each tweet, such as the number of retweets, quotes, likes, followers, and following, as well as the date and hour of the tweets, which were collected as metrics for this study. This allows us to conduct a comprehensive analysis of the data and answer the research questions.

In conclusion, the use of the Twitter API in this study has allowed us to efficiently collect and analyze a large and relevant sample of data on online customer engagement and conflict management in brand communities, for the full 2022 year, which is crucial to answer our research questions. The API's various features and query parameters have enabled us to filter and retrieve only the most relevant tweets, and access a wide range of data on each tweet, which has greatly facilitated our research.

1.2.2 Preprocessing Data

In order to prepare the data for analysis, we performed several preprocessing steps on the collected tweets.

First, we normalized the metrics (retweets, quotes, likes, followers, and following) using z-score normalization and then min-max normalization. This allowed us to standardize the data and make it comparable across different tweets and companies.

Z-score normalisation is a statistical technique used to standardise data by transforming them into a scale of standard values. This means that the data is expressed in terms of its standard deviation from the mean of the data distribution. It is calculated using the following formula:

$$z = \frac{X - \mu}{\delta}$$

X: the value of a specific data item
μ : average of the data distribution
δ : standard deviation of the data distribution

Min-Max normalisation is a statistical technique used to standardise data by transforming them into a specific range of values. This means that the data is expressed in terms of its proportion to the minimum and maximum value range of the data distribution.

$$mMnorm = \frac{X - min}{max - min}$$

X: value of a specific data item
min: minimum value of the data distribution
max: maximum value of the data distribution

Next, we calculated Jaccard similarities for each tweet grouped by date and company using context annotation. This allowed us to measure the similarity of the tweets based on their context annotation, and to identify patterns and trends in the data. We then calculated the average Jaccard score for each tweet, which was used in the subsequent analysis.

Jaccard similarity, also known as Jaccard index, is a measure of the similarity between two sets of data. It is defined as the size of the intersection of the sets divided by the size of the union of the sets.
In the context of text analysis, Jaccard similarity can be used to measure the similarity between different context annotations of tweets by comparing the sets of context annotations that they contain. The Jaccard similarity score between two context annotations can range from 0 (no similarity) to 1 (identical context annotations).
In our code, we are using the Jaccard similarity to measure the similarity between the context annotations of tweets. So, let's say we have two tweets t_1 and t_2 with context annotation sets A_1 and A_2 respectively. The Jaccard similarity between these two tweets can be computed as:

$$J(A1, A2) = \frac{|A1 \cap A2|}{|A1 \cup A2|}$$

To calculate Jaccard similarity for each context annotation of a tweet, we first convert the list of context annotations for each tweet into a set of unique annotations. Next, we create a sparse matrix of these sets of unique annotations, where each row represents a tweet and each column represents a unique context annotation. We then compute the dot product of the sparse matrix with its transpose, which gives us the number of co-occurrences of each unique context annotation in the tweets.

Next, we create an array of the lengths of the sets of context annotations (i.e., the number of unique context annotations in each tweet). Finally, we compute the Jaccard similarity scores by dividing the dot product by the outer product of the lengths of the sets of context annotations, which gives us the ratio of the number of co-occurrences of each unique context annotation in the tweets to the total number of unique context annotations in the tweets.

The formula for calculating the average Jaccard similarity score for each ID is :

$$Average\_score = \frac{\sum_{i=1}^{n} Jaccard\_similarity_i}{n}$$

Where $Jaccard\_similarity_i$ is the Jaccard similarity score for the i-th tweet, and n is the total number of tweets that have the same ID.

Finally, we transformed emojis to text, erased emoticons and URLs, and tokenized the text. This allowed us to clean the text data and make it more suitable for text mining and sentiment analysis.

1.2.3 Sentiment analysis

Sentiment analysis is the automatic analysis of emotions, opinions and attitudes expressed in texts, such as social media posts, blog comments or product reviews. It uses automatic natural language processing techniques to identify and extract information about the sentiments in a given text. Sentiment analysis is used in many areas, such as marketing, market research, brand monitoring and online reputation management. It can help companies understand their customers' opinions and improve their relationship with them, as well as identify trends and important themes in online conversations. This analysis allowed us to sort and filter our data and focus specifically on negative tweets. Indeed, we decided to restrict our study to the propagation of negative tweets, given that they are the most likely to result in online firestorms and have a negative impact on brands.

For the sentiment analysis of the text data, we used the AFINN library. AFINN is a lexicon-based approach for sentiment analysis that uses a pre-existing list of words and their associated sentiment scores to assign a sentiment score to each word in a given text. The lexicon is based on a list of words and their associated sentiment scores, which range from -5 (strongly negative) to +5 (strongly positive). The overall sentiment of the text is then calculated by summing the sentiment scores of all the words in the text. The AFINN lexicon includes over 2,500 words and their associated sentiment scores, and it was created using a crowdsourcing approach. The words were collected from various sources, such as online blogs and forums, and were then rated by human annotators on a scale of -5 to +5 according to their sentiment. The lexicon covers a wide range of words and phrases, including both positive and negative words, as well as neutral words. It also includes words and phrases that are commonly used in social media, such as emoticons and slang.The overall sentiment of the text is then calculated by summing the sentiment scores of all the words in the text. This library is simple and efficient, it has been widely used in many studies. The sentiment can be positive, negative or neutral.

The results of the data analysis were then interpreted and discussed in relation to the research questions, and the findings were presented in the results section of the paper.

It is worth noting that, as with any research involving social media data, there are limitations to the generalizability of the findings. The sample of tweets analyzed in this study may not be representative of all tweets related to the selected companies, and the results may not apply to other companies or industries. Additionally, the results may be affected by biases in the data, such as selection bias and self-selection bias.

Feature name	Description
normalized_retweet	The number of retweets normalized using z-score and min-max normalization.
normalized_like	The number of likes normalized using z-score and min-max normalization.
normalized_quote	The number of quotes normalized using z-score and min-max normalization.
normalized_followers_count	The number of followers normalized using z-score and min-max normalization.
normalized_following_count	The number of accounts followed by the tweet's author normalized using z-score and min-max normalization.
normalized_jaccard	The Jaccard similarity coefficient normalized using z-score and min-max normalization.
sentiment	The sentiment analysis of the tweet's text, either "positive", "neutral", or "negative".
jaccard	The Jaccard similarity coefficient between the text of the tweet and the text of the replies.
virality	A metric calculated based on the number of retweets, quotes, and likes.
id	The unique identifier of the tweet.
author_id	The unique identifier of the tweet's author.
text	The text of the tweet.
followers_count	The number of followers the tweet's author has.
following_count	The number of accounts the tweet's author is following.
tag	Hashtags or keywords associated with the tweet.
like_count	The number of likes the tweet has received.
retweet_count	The number of retweets the tweet has received.
quote_count	The number of quotes the tweet has received.
tweet_date	The date and time the tweet was posted.

Figure 1 - Presentation of all features data

1.3 Virality

Virality is a result that derives from an unexpected event that expands in a very small-time interval. In our case, the virality is the sudden discharge of large quantities of messages containing negative word of mouth 'eWOM' and complaint behavior against a person, company, or group in social media networks. In these messages, intense indignation is often expressed, without pointing to an actual specific criticism" (Pfeffer, Zorbach and Carley 2014: 118). In our case, this event is considered as an online-firestorm that may be harmful to organizations and their brands if they are not properly controlled.
Throughout our research, we have defined virality based on two research papers that we have studied [1][2]. This allowed us to delimit two types of virality.

1.3.1 Hoang et al. Virality [1]

In this paper, the researchers' problem was the Identification of the most viral messages during the 2011 Singapore General Election 'GE2011' and the users behind them. The aim of the research is to measure viral content (Popularity & new metrics), measure viral behavior analysis and analyze viral content (Tweets, politicians & topics).

Researchers offer a few models for calculating the virality of tweets, topics, and users. Each model gives a numerical virality value. To indicate the virality of a tweet (m), a topic (t), and a user (u), we use the notation mscore(m), tscore(t), and uscore(u).

M/T/U	set of tweets/topics/users
u(m)	original author of the tweet m
$U_R(m)$	set of users retweeting the tweet m
W(m)	set of non-stop words in tweet m
M(t)	set of (original) tweets in topic t
M(u)	set of (original) tweets tweeted by user u
$M_R(u)$	set of (original) tweets retweeted by user u
Follow(u)	set of followers of user u
mscore(m)	viral score of tweet m
tscore(t)	viral score of topic t
uscore(u)	viral score of user u

Figure 2 - Virality notations

To do so, the researchers defined virality based on three elements: Viral tweets, topic virality and viral users.

1.3.1.1 Viral tweets

- Virality by retweet count

The retweet count model tells us the popularity of a tweet, but does not tell us how good the tweet is propagated through the social network structure.

$$mscore_{rtc}(m) = |U_R(m)|$$

$U_R(m)$: set of users retweeting the tweet m
$mscore_{rtc}(m)$: viral score of tweet m

- Virality by retweet likelihood

Consider the two original tweets, m1 by user u1 and m2 by user u2. Suppose that only 4 out of 20 followers of u1 retweeted m1, and all 3 followers of u2 retweeted m2. By retweet count model, m1 is more viral than m2. However, one may consider m2 to be more viral since all of its receivers have retweeted while only a small fraction of users receiving m1 actually retweet m1.
Considering the supposition, researches proposed the formula below:

$$mscore_{rtl}(m) = \frac{|U_R(m)|}{|\bigcup_{u \in U_R(m) \cup \{u(m)\}} Follow(u)|}$$

$u_(m)$: origial author of the tweet m
$U_R(m)$: set of users retweeting the tweet m

The dataset lacks complete follower information for all second users. Given these tenuous connections, it was decided to use the set of users interested in u rather than the set of u's followers.

$$mscore_{rtl}(m) = \frac{|U_R(m)|}{|\bigcup_{u \in U_R(m) \cup \{u(m)\}} \{u' : u' \ interested \ in \ u\}|}$$

user u1 is interested in user u2 if u1 mentions "@u2" in at least k of u1's tweets.

Researchers set k = 1. This means u1 mentions u2 at least once or u1 retweets at least one tweet from u2.

The virality by retweet likelihood model has a flaw in that it favors tweets published by users with fewer followers while disregarding tweet popularity among retweeting users.

- Virality by a combined model

The paper proposes a combined model incorporating both retweet count and retweet likelihood.

$$mscorec(m) = mscorertc(m) * mscorertl(m)$$

1.3.1.2 Topic virality

A Topic is a collection of tweets with a common theme. We use each tweet as a topic to build a topic. Next, we apply a modularity-based clustering algorithm to a tweet graph, (tweets=nodes, pairs of tweets with overlapping terms=edges). To compute similarity, we use the "bag-of-words" representation of each tweet after removing tweet-encoding terms. The similarity between tweet m1 and tweet m2 is computed by Jaccard coefficient.

$$similarity(m_1, m_2) = \frac{|W(m_1) \cap W(m_2)|}{|W(m_1) \cup W(m_2)|}$$

Where W(m) is the set of non-stop words in tweet m.
This simililarity function '$similarity(m_1, m_2)$' returns a value between 0 and 1.

$$tscore(t) = \sum_{m \in M(t)} mscore(m)$$

Finally, the topic with a lot of viral tweets is calculated using this formula, where tscore(t) is the viral score of topict and mscore(m) is the viral score of tweet m.

1.3.1.3 Viral users

There are two basic modeling approaches for viral users, namely the author-take- all and the shared contribution models.

- 'Author-take-all' model

According to this model, a tweet's virality can be entirely attributed to its original author but not to the users who retweeted it. According to this supposition, a user's virality is equal to the sum of all of their original tweets u.

$$uscore(u) = \sum_{m \in M(u)} mscore(m)$$

That is, where the mscore(m) is the viral score of tweet m.

- 'Shared contribution' model

$$uscore(u) = \sum_{m \in M(u) \cup M_R(u)} \left(\frac{mscore(m)}{1 + |U_R(m)|} * (1 + |U_R(m) \cap Follow(u)|) \right)$$

This model accounts for a user's (u) contribution to the overall virality of all tweets (including retweets) when calculating the user's (u's) virality. In the case of an original tweet m and a user u, the latter's contribution to the virality of the former is proportional to the number of retweets on either the original tweet (if u is the author of m) or on the u's retweet of m (if u retweets m).

Given the complexity of the model as well as the complexity of importing data from API twitter, this research paper has been dedicated only to the theoretical part of our research to allow to understand more the elements that we will have to take into consideration as we go along the technical part.

1.3.2 Herhausen et al. Virality [2]

In this research paper, the concept of virality was defined by the researchers according to elements and theory that depends on the social network Facebook. It was defined as the combined sum of likes and comments post i receives from other customers in community c any time after it was posted.

$$virality_{ic}^{t-1} = \gamma_{00} + \gamma_{01-04} Firm\ Controls_c + \gamma_{05-10} Brand\ Community\ Controls_c$$
$$+ \gamma_{11-17} Post\ Controls_{ic} + \gamma_{18} Dum\ No\ Firm\ Response_{ic} + \gamma_{19-23} Post\ Predictors_{ic} + \gamma_{24-43} Dum\ Timing_{ic}$$
$$+ u_{0c} + r_{ic}$$

In the case of our research work, we redefined the notion of virality to be consistent with the social network Twitter. For this purpose, the definition of virality is still exchanged and is represented as follows:

$$virality = \sum received\_likes_{ic} + received\_comments_{ic}$$

1.4 Strength Structural Ties (SST)

Strength Structural Ties* known also as "frequency of communication" is a measure to define the ties between the user that posted the negative eWOM (word of mouth) and the community c before the user wrote the negative eWOM the general formula is shown below(Herhausen et all p-12).This formula was defined by the researchers according to elements and theory that depends on the social network Facebook.

$$SST_{ic} = \sum_{\tau=0}^{t-1} received\_likes_{i_c}^{\tau} + received\_comments_{i_c}^{\tau} + received\_shares_{i_c}^{\tau} + \sum_{\tau=0}^{t-1} likes\_given_{i_c}^{\tau}$$
$$+comments\_given_{i_c}^{\tau}$$

In the case of our research work (Twitter), we simplified this formula to have the Structural Ties between i and c for the received Likes, Retweets.

$$SST_{ic} = \sum_{\tau=0}^{t-1} received\_likes_{i_c}^{\tau} + received\_retweets_{i_c}^{\tau}$$

i: user that triggered the negative EWOM
c: community that have liked retweeted and quoted the negative tweet that triggered the firestorm

In our case, we choose 2 days before the user triggers the nEWOM because of the limitations of Twitter API (takes a lot of time 15 requests/15 minutes) to collect the data.For twitter API which we used we didn't have requests API for the Given Likes and Replies so we have concentrated only on the received.

The purpose of the SST is to determine how the community was related before the user i posted the negative eWom. The more the results are higher the more the community is connected with the user i and as a result the more there is a chance that the negative eWOM will be a big firestorm (we suppose).

Because of the limitations of twitter API, we had to limit ourselves with the received likes and comments. We choose a 2 day interval before the nEwom.

1.5 Linguistic Style Match (LSM) [5][10]

The networks created through the use of social media rely on technical features that allow users to establish online relationships with many other users and to communicate with each other. As a result, users form dense network groups. Within these groups, the flow of information is usually relatively constant and unlimited. The time until the next piece of information is usually very short, which promotes the rapid dissemination of information and fuels online firestorms. Therefore, a huge number of people can be reached by eWOMs in a short period of time.

In the context of our research, it is interesting to work on the LSM "Linguistic Style Matches"which will allow us to better understand the diffusion of information within user communities and at the same time the creation of online firestorms.

The "linguistic style matchïs a concept related to the analysis of language used in verbal or written communication. It describes the correspondence between the linguistic styles of different speakers or authors in a conversation or correspondence. Studies of linguistic style matches have shown that people are more likely to connect and communicate effectively with those who have styles similar to their own.

LSM elicits similarity perceptions, the increasing match between the review and the product interest group's linguistic styles should make that review more appealing, as well as grant greater importance to changes in its content.

So to make that happen we used this formula to proceed it:

$$LSMj_{ic} = 1 - \frac{|FWj_i - \overline{FWj_{ic}}|}{FWj_i + \overline{FWj_{ic}} + 0{,}0001}$$

i: user that triggered the negative EWOM
c: community that have liked retweeted and quoted the negative tweet that triggered the firestorm

We derived the degree of LSM between customer i posting negative eWOM at time t with the receiving brand community c in three steps.

First, we mined the use intensity of each of the nine function word categories j separately in focal customer i's message and across all customer messages (negative and positive) in the brand community c posted in the previous 2 days in response to the negative eWOM post.
Second, the degree of similar use intensity LSM of each function word category (FWj) by customer i posting the negative eWOM into community c.
Third, by aggregating all nine LSM scores with equal weights, we obtain an LSM score bound between 0 and 1, and scores closer to 1 reflect greater degree of communication style matching between customer i and the online brand community c.

For the FWC we relied on our on dictionary by FWC(Personal pronouns,Impersonal pronouns, Articles, the Conjunctions Prepositions, Auxiliary verbs, High-frequency adverbs Negations, Quantifiers).

1.6 Poisson regression

In statistics, Poisson regression belongs to the family of generalized linear models. It is used when the variable to be analyzed results from a counting process (such as a number of deaths, a number of adverse events, or a number of new cases). This method is widely used in epidemiology. Indeed, it can be observed that there is a link between the spread of diseases and viruses and the dissemination of information in social media. The characteristics of the disease, such as the probability of transmission or the complexity and structures of human contact networks play a crucial role in their transmission. Therefore, the methods used in epidemiology are adaptable to our research.

The spread of negative eWOM can be compared to the spread behavior of pathogenic organisms: Negative eWOM (like pathogenic organisms) can be spread via social (in our case virtual) links between social media users (i.e. humans). The rate of infection is therefore highly dependent on the characteristics of the disseminated information (such as negativity, emotion), the structure of the underlying network (the population density and connection between users) and the process of obtaining the information (the means of transmission in the network).

Therefore, we base the design of our new algorithm for the detection of emerging online firestorms on techniques for the early detection of infectious disease outbreaks. We aim to develop a model for automated, real-time detection of online firestorms.

In this section, we have mainly relied on the work done in the scientific article by Benedict Drasch, Johannes Huber, Sven Panz and Florian Probst "Detecting Online Firestorms in Social Media".

Through the data collected via Twitter scraping, the calculation of virality using the method cited [2] and the sentiment analysis cited, there is still a filtering to be done at the data level. Unlike the paper on which we were inspired [3], we do not base our model on the set of tweets that gather the total sentiments, we only filter on negative sentiments. As explained earlier, we aim to predict the emergence of firestorms that can have detrimental effects on a company on the Twitter social network, it would be distorted to take into account those who praise it.

In order to have interesting data and a qualitative method, it was crucial to choose the entity that had the most important emergence and a large number of tweets. The data was therefore filtered in a first step on the companies that have the most negative tweets about them.

The data being spread over a full year, it is difficult through a sample of this size to make estimates of viral tweet volumes in the future. Trends on social networks can change quickly, which means that tweets on a subject can vary considerably from one year to another, from one month to another and from one day to another. In addition, seasons and important events can have a significant impact on the number of tweets on a given subject, which can make estimates based on a full year of tweets inaccurate. It is therefore preferable to use more recent data and to take into account seasonal trends and important events to make more accurate estimates.

The most logical trend and the one most likely to correspond to the Twitter environment, different from other social networks, seems to be a week. Indeed, we do not consume content in the same way on each of the social networks. The life of a tweet is relatively low, 15 to 20 minutes on average, but its spread is much more important and consequential.

Through the date of each tweet, another time indicator, the number of the week, was added. We group by week and calculate an average of the virality factor defined "hereïn order to highlight the week that is the most viral over the full year.

When the week is identified, we set a threshold that serves to separate into two (viral, non-viral) the variable that defines the viral score calculated according to the method mentioned [3]. To do this, we use the quantile method. A measure of the distribution of a data set that allows to determine the level of value lower to a certain percentage of data. More precisely, a quantile is a value that separates a data set into two equal parts: a lower part containing a certain percentage of data and an upper part containing the remaining percentage of data. This distribution thus allows us to identify extreme or unusual values. It is important to define it with precision to obtain distributions close to the definition of virality.

Once the date has been identified, a set of days before and after the date with the most tweets exceeding the viral threshold is collected.

Two options are possible once the interval that encompasses the most viral day on the average most viral week is collected.

The same quantile defined above can be kept or another one can be determined. Before continuing the selection of data and the most likely interval, the data is visualized to confirm that one day stands out the most from the others and that it follows the behavior of a firestorm on the social network.

Here is the notation that we will use to define the model we used to define the volume of negative and viral negative tweets. We set $v \in \{negativ, viral\}$ and the variable $C_i^v \in \mathbb{N}_0 \; \forall i, v$ encompasses negative and viral tweets (which are negative but one of extreme virality). C_0^v is the volume for a future period noted t_0.

As referenced in the paper [3], we hypothesize that our data follows a Poisson distribution. To do this, we retrieve the fixed data sample set previously and apply different statistical tests to confirm if it is comparable to a Poisson distribution.

The Kolmogorov-Smirnov (K-test) is a statistical test used to verify if a data series comes from a particular distribution.

When used to test a Poisson distribution, the K test measures the distance between the empirical distribution of the data and the theoretical Poisson distribution. The result of this test is a p-value, which indicates the probability that the data conforms to the Poisson distribution.

If the p-value is greater than a given significance threshold (usually 0.05), we can conclude that the data probably comes from a Poisson distribution. However, if the p-value is lower than this threshold, we can reject the hypothesis that the data comes from a Poisson distribution and consider that the data comes from a different distribution.

The K test verifies if a data series corresponds to a Poisson distribution and determines if this hypothesis can be accepted or rejected based on the p-value obtained.

Following the paper [3], we apply the following regression equation:

$$log(E[C_i^v]) = log(\mu^v) = \theta^v + \beta^v * t_i \tag{1}$$

Where θ^v is a constant, β^v denotes the trend, and i is a control variable for all parameters of the last n periods.

This Poisson regression is a model used to predict the frequency of occurrence of an event over a given period or in each space. We therefore perform a regression to have a correct estimation of the regression coefficients.

Results and discussion

2.1 Data Processing

The data collected and analyzed in this study provides a comprehensive understanding of the online customer engagement and conflict management patterns related to selected companies over a one-year period. The use of the Twitter API enabled us to efficiently collect and analyze over 218,215 tweets, providing a vast amount of data to be analyzed. By normalizing the metrics such as retweets, quotes, likes, followers, and following, we were able to standardize the data and make it comparable across different tweets and companies.

The findings of this study can provide valuable insights for companies to better understand their online reputation and engage with their customers effectively. By analyzing the intensity of criticism and complaints, firms can identify areas where they need to improve and address customer dissatisfaction. Additionally, the examination of the firm-hosted brand communities allows companies to understand the dynamics of the network and how to better manage their brand reputation.

Future studies could explore these limitations by incorporating additional data sources, such as surveys or interviews, to gain a more comprehensive understanding of the factors that influence customer engagement and conflict management. Additionally, time series analysis could be applied to the data to examine the evolution of customer engagement and conflict management over time, and to identify seasonal variations and long-term trends. The potential applications of this research are numerous, and we believe that these insights will be valuable for organizations, researchers, and practitioners alike.

However, it is important to note that the findings of this study should be interpreted with caution, as it is based on data collected from a single social media platform, Twitter. While Twitter is widely used and provides a rich source of data, it is still subject to limitations such as selection bias and self-selection, which may affect the representativeness of the data. Future studies can expand upon this research by collecting data from other platforms and exploring other variables that may impact customer engagement and conflict management.

2.2 SST

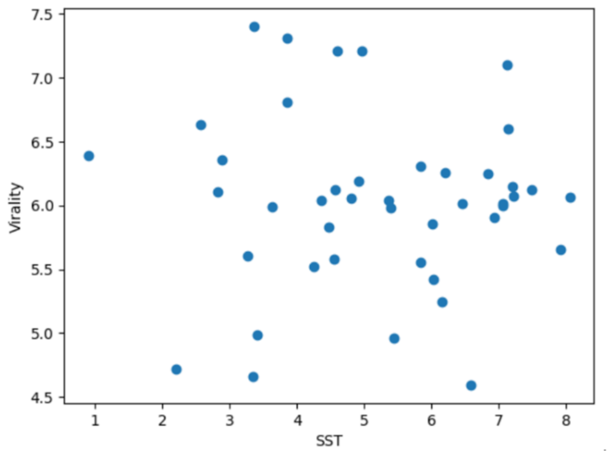

Figure 3 - Relationship between Strength Structural Ties (SST) and virality.

The relationship between strength of structural ties and virality is represented by the graph below, which was generated using data from Walmart and Amazon, with a total of 44 viral cases. The graph does not show a clear trend, with the points appearing scattered, which led us to perform a test of independence. The test showed that there was no significant dependence between strength of structural ties and virality, as indicated by the Pearson correlation with a p-value exceeding 0.05. To further understand the relationship, the data was subjected to a logarithmic transformation. The transformation provided a clearer representation of the relationship, but did not change the conclusion of independence between the two variables. These findings suggest that the strength of structural ties may not be a determining factor in the spread of information and virality in social networks.

As we can see from the analysis of Strength Structural Ties and Virality graph, there is no apparent correlation between the two variables. Despite initial assumptions that a relationship might exist, our research findings indicate otherwise. This was an important discovery, as it highlights the need to approach social media analysis with a critical eye, and to question any preconceived notions or hypotheses. Herhausen et al. conducted a similar study using Facebook data, and their findings were deemed relevant to our research goals. Based on this, we decided to apply the same methodology to Twitter data. However, during the technical implementation phase, we encountered several challenges.

Firstly, importing data from the Twitter API proved to be a major obstacle. The API has strict limits on the number of requests that can be made within a given time frame, which limited the amount of data that could be collected. Despite our best efforts, we were unable to find a viable solution to this problem.

Additionally, the library that we were using had limitations of its own, with a strict limit of 15 requests per 15 minutes. This further exacerbated the issue of collecting sufficient data in a timely manner, and made it challenging to import the data required for our analysis within the project deadline.

In conclusion, our findings demonstrate the importance of being mindful of the limitations and challenges that can arise when working with social media data. Despite our best efforts, the technical hurdles associated with data import ultimately hindered our ability to fully realize the potential of our research. Nevertheless, we learned valuable lessons that will inform future work in this area, and we are confident that our findings will be useful for other researchers and practitioners in the field.

2.3 Poisson regression

As we presented in previous sections, we were able to collect more than 200k tweets from 11 companies after various treatments. Each company has different and non-comparable virality rates. Taking the example of the Twitter company, the company has been at the center of several controversies, which contributed to the instability of its data on negative tweets and viral negative tweets.

One of the main reasons for these controversies is related to the actions of Elon Musk, the entrepreneur and CEO of Tesla and SpaceX, who often used Twitter to post provocative and bewildering tweets, which often created a buzz on social media and had an impact on the virality level on the platform.

All of this made it more difficult to generalize the statistical and regression coefficients for these data grouping all companies.

As a result, we made the choice to choose an 'ordinary company' that brings together a significant number of tweets but does not have a Twitter user who can vary the virality like Elon Musk and his new company. We chose the case of Disney, its streaming platform and its recurring releases of new series allow for a fairly constant and interesting flow of tweets to study.

Through the various methods we have defined, we were able to determine its most viral week and find the day that causes this high average. In Figure 4, we can see the distribution of the number of negative tweets over an 8-day interval. In Figure 5, we have a distribution of the number of negative viral tweets, fixed with a quantile of 0.9 on the virality variable of all our tweets listed on the Disney company.

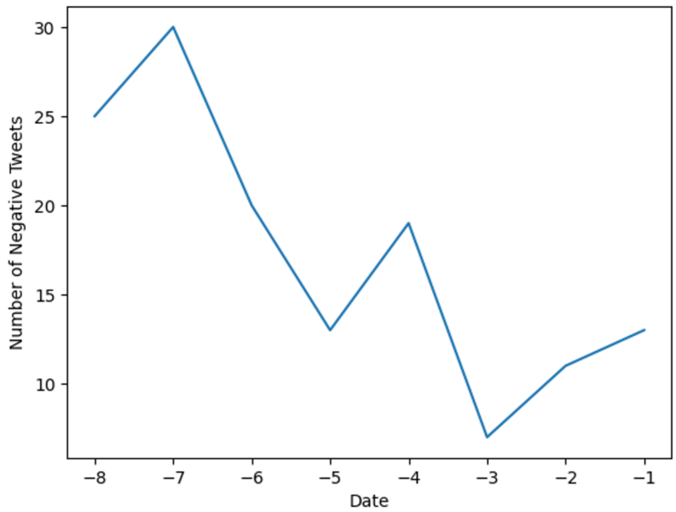

Figure 4 - Distribution of the number of negative tweets per day on average of the most viral week.

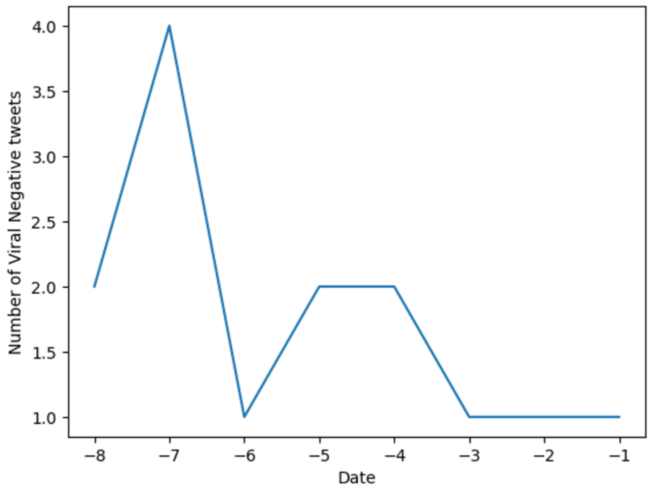

Figure 5 - Distribution of the number of viral negative tweets per day on average of the most viral week.

These two distributions seem to be in line with a firestorm, with a sudden peak on day -7 (the most viral day in the chosen interval) and the volume decreases day by day like any 'bad buzz' on social media, it fades over time.

We also carried out the statistics in order to be in accordance with the assumptions of the model and the type of distribution we should have. In addition to the K-test, we generated probabilistic distribution graphs, also known as Q-Q plots. This type of graph is used to check if a data series follows a Poisson distribution.

Here are the two graphs representing the verification of the desired distribution.

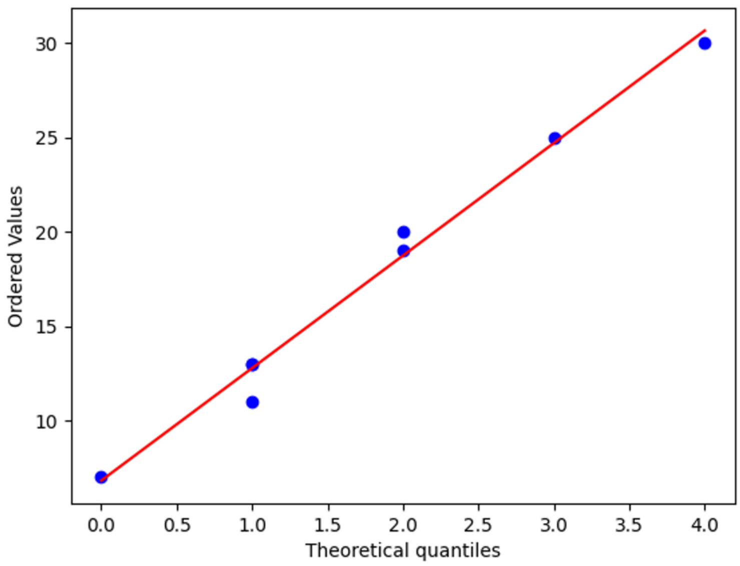

Figure 6 - Probability plot for the negative tweets.

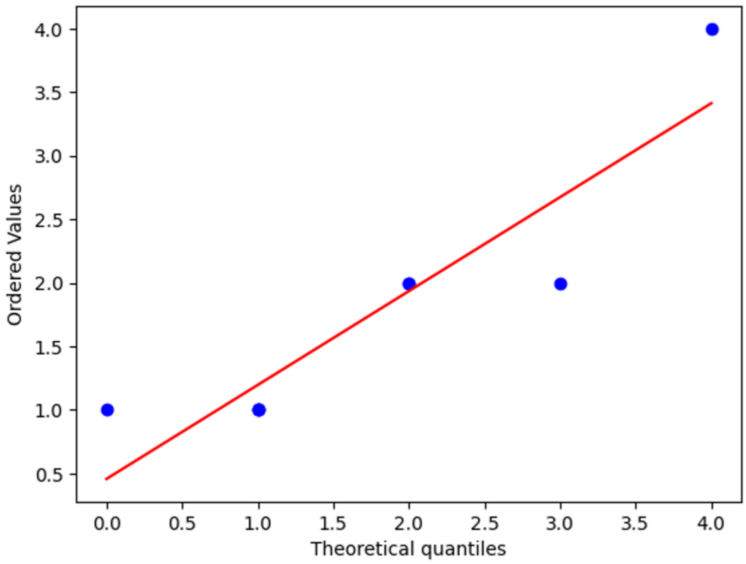

Figure 7 - Probability plot for the viral tweets.

Figure 6. seems to really approximate a Poisson distribution. Figure 7. a little less, still consistent with the result of the k-test which gives us a p-value allowing us to reject the null hypothesis.

After making these graphical representations and statistical validity tests, it is now possible to apply a Poisson regression to these data and derive the coefficients of equation (1).

Let us remember that the current coefs are affiliated with the Disney company for one week in the year. It is difficult to generalize these coefficients to the entire set of companies stored in our database.
For the number of negative tweets we have :

$$\theta^{negative} = 2{,}11$$

$$\beta^{negative} = -0{,}149$$

And below the graph associated with the expected volume prediction by applying our equation (1) raised to the exponential:

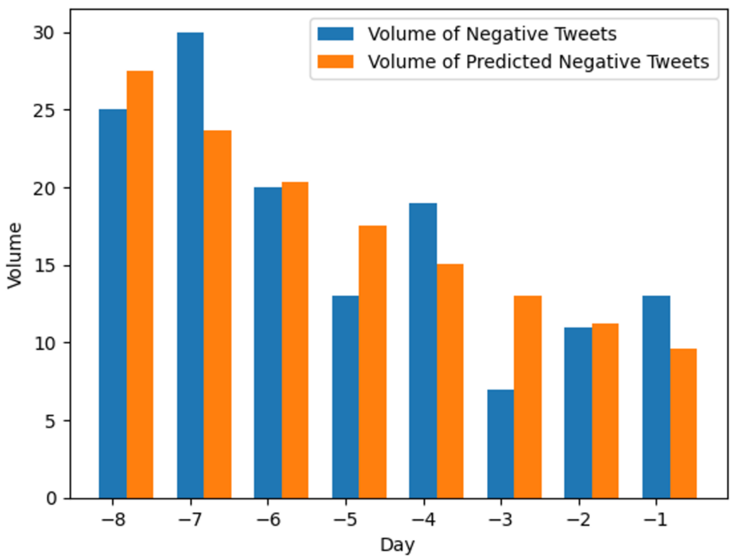

Figure 8 - Visualization of the actual and predicted volume of negative tweets.

For the number of viral tweets we have:

$$\theta^{viral} = -0{,}192$$

$$\beta^{viral} = -0{,}153$$

And below the graph associated with the expected volume prediction by applying our equation (1) raised to the exponential:

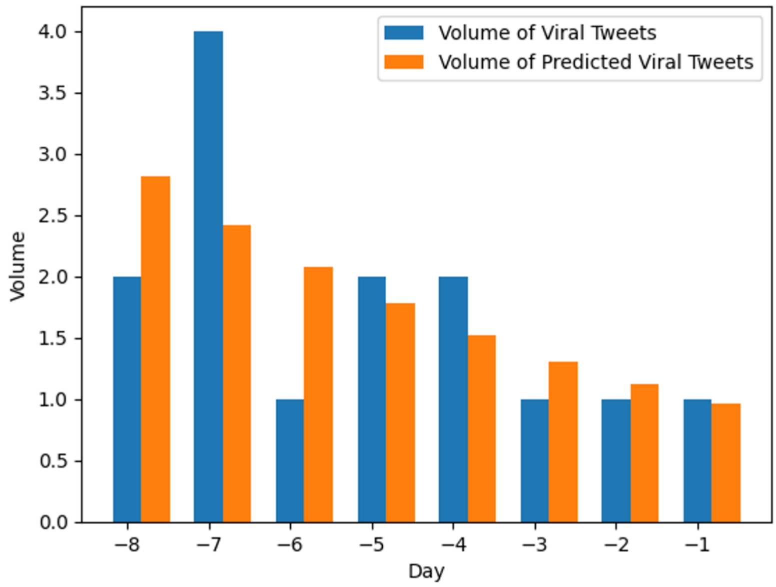

Figure 9 - Visualization of the actual and predicted volume of viral tweets.

The graphs express rather well the expected volumes on our two variables. On average between the two graphs, the two graphs have an accuracy of about 75 %.

Conclusion and future recommandations

The API has limitations in terms of the number of requests that can be made per minute. This means that there are restrictions on the frequency at which requests can be made to the API. To mitigate this, there is a server that hosts requests to retrieve any missing parameters. This will help to ensure that all the necessary information is available for the calculations. The calculation engine runs on the cloud and is designed to recalculate the indicators periodically in real-time. This ensures that the results are up-to-date and accurate. Finally, the recovery of Bloomberg data is critical for pushing the study to assess the financial impact generated. This information is used to understand the financial implications of the calculations and to make informed decisions.

In order to enhance the predictions of virality in the future, additional variables will be added to the regression model. This will help to increase the accuracy of the predictions by taking into account additional factors that could impact virality. Similarly, variables will also be added to the modified SST in the future in order to obtain a more meaningful result. These modifications will help to ensure that the results of the analysis will be comprehensive and take into account all relevant factors. The addition of these variables will allow for a more robust and nuanced understanding of the relationship between the various factors and virality in the future.

Finally, to gain a better understanding of the different degrees of virality across different sectors in the future, more data on companies from various sectors will be collected. This will allow for a more comprehensive analysis and a deeper understanding of the factors that will contribute to virality. What's more, In order to have a stable Poisson model, more tests will be conducted in the future to determine the optimal period size or "sliding days."This will help to ensure that the model will be robust and accurate in all cases. To further improve the analysis, efforts will be made in the future to estimate the quantile that will allow us to determine whether one tweet will be more viral than another on the fixed period. This will help to better understand the relationship between different variables and virality. In addition, each entity will be benchmarked in the future to determine which coefficients will be the best predictors of our volumes. Finally, proposals from research papers on the setting of the threshold for determining when to trigger an alarm will be pursued in the future to ensure that the results of the analysis will be meaningful and relevant.

References

[1] *On Modeling Virality of Twitter Content.* HOANG, Tuan Anh; LIM, Ee Peng; ACHANANUPARP, Palakorn; JIANG, Jing; and ZHU, Feida (2011). Digital Libraries: 13th International Conference on Asia-Pacific Digital Libraries, ICADL 2011, Beijing, China, October 24-27: Proceedings. 7008, 212-221. Research Collection School Of Information Systems.

[2] *Detecting, preventing, and mitigating online firestorms in brand communities.* Herhausen, D., Ludwig, S., Grewal, D., Wulf, J., Schoegel, M. (2019). Journal of Marketing, 83(1), 1-21.

[3] *Detecting Online Firestorms in Social Media.* Benedict Drasch, Johannes Huber, Sven Panz, Florian Probst (2015). Thirty Sixth International Conference on Information Systems, Fort Worth.

[4] *Disaster Dynamics: Understanding the Role of Quantity in Organizational Collapse.* Jenny W. Rudolph, Nelson P. Repenning (2002). Cornell University.

[5] *Take their word for it: The symbolic role of linguistic style matches in user communities.* Stephan Ludwig, Ko de Ruyter, Dominik Mahr, Martin Wetzels, and Elisabeth Brüggen, Tom de Ruyck (2014).

[6] *Unsupervised Detection of Sub-events in Large Scale Disasters.* Chidubem Arachie, Manas Gaur, Sam Anzaroot, William Groves, Ke Zhang, Alejandro Jaimes (2019).

[7] *Complaint Deescalation Strategies on Social Media.* Dennis Herhausen, Lauren Grewal, Krista Hill Cummings, Anne L. Roggeveen, Francisco Villarroel Ordenes, Dhruv Grewal (2022).

[8] *Brand Buzz in the Echoverse.* Kelly Hewett, William Rand, Roland T. Rust, Harald J. van Heerde (2016).

[9] *Financial development and dynamic investment behavior: Evidence from panel VAR.* Inessa Love, Lea Zicchino (2006). Board of Trustees of the University of Illinois.

[10] *More Than Words: The Influence of Affective Content and Linguistic Style Matches in Online Reviews on Conversion Rates.* Stephan Ludwig, Ko de Ruyter, Mike Friedman, Elisabeth C. Brüggen, Martin Wetzels, Gerard Pfann (2013). City, University of London.

[11] *Against the Others! Detecting Moral Outrage in Social Media Networks.* Wienke Strathern, Mirco Schoenfeld, Raji Ghawi, Juergen Pfeffer (2020).