USING BIG DATA TO EVALUATE CORPORATE SOCIAL RESPONSIBILITY AND SUSTAINABLE DEVELOPMENT PRACTICES

Adriana M. Barbeito-Caamaño, Ricardo Chalmeta
Grupo de Investigación en Integración y Re-Ingeniería de Sistemas (IRIS)
Departamento de Lenguajes y Sistemas Informáticos
Universitat Jaume I, 12006 Castellón, Spain.
Tel: + 34 964 728329 Fax: + 34 964 728435
adrianabcaa@gmail.com, rchalmet@uji.es

Corresponding author: Ricardo Chalmeta

Abstract

This paper shows an alternative method to evaluate sustainable development and CSR practices based on the opinions of companies' stakeholders expressed on Twitter. An application of the method is performed with the Inditex textile group.
This paper shows two important findings: (1) Knowledge about companies' CSR practices and stakeholders opinions can be obtained using Big Data to analyse CSR information about the company on online social networks; and (2) There are important differences between the contents of the sustainability report of a company and the CSR opinions of the its stakeholders on the internet. These findings have benefits for the company's stakeholders, who will be able to know the CSR practices of a company in a more objective way, and for the company, which will be able to improve its CSR performance and communication strategy, and the stakeholder engagement.

Keywords:

Corporate Social Responsibility, Twitter, Sentiment Analysis, Big Data, Sustainability reports, Sustainable development, Stakeholder engagement

1. INTRODUCTION

The idea that the main objective of a company is to obtain an exclusively economic benefit (Friedman, M., 1970) has been replaced by the stakeholder theory (Freeman, 1984), which sees other groups as being valuable to the company and as having interests that should also be taken into account. As a consequence, the concept of Corporate Social Responsibility (CSR) appears. Although, there is no consistency in the CSR concept (Huang et al., 2019), CSR can be defined as *the responsibility of enterprises for their impacts on society* (European Union Commission, 2011). In other to meet this responsibility, corporates need to involve their stakeholders to integrate economic, social, environmental, ethical human rights, and consumer concerns into their business operations and core strategy (Song & Jing, 2019).

Environmental, social or corporate governance concerns may impact firm value and managers can no longer ignore this (Capelle-Blancard and Petit, 2019). CSR has now become a strategic objective rather than a voluntary activity, since it is a differentiating factor and a tool for competitiveness (Ruff et al., 2001). CSR practices have great benefits for the company. They allow the company to improve its performance, to increase sales

and productivity, to achieve greater customer and employee loyalty (Maignan, 2001) and to enhance the image and reputation of the company (Smith, 2003).

Therefore, CSR commitments and actions of the company cannot remain internalised within it. They should be effectively communicated to its stakeholders. To communicate their CSR issues to their stakeholders, companies are using different ways such as sustainability reports, media advertisements, information disclosure on their corporate websites and posts on online social networks. In addition to the CSR information generated and disseminated by the company itself, stakeholders can find information about the CSR practices of the company in the reports carry out by Non-Governmental Organizations (NGOs), Environmental, Social and Governance (ESGs) rating agencies, and news outlets.

However, there is another source to obtain knowledge about the CSR practices of a company: the information, opinions, comments, suggestions, complaints, etc. posted on online social networks by company's stakeholders, in which they are expressing their own vision of the company's CSR practices. This information disseminated via online social networks can be collected and processed using Big Data techniques. Therefore, it is possible to assess companies' CSR practices according to the opinions of their stakeholders on the online social networks. It would allow to gain company CSR-relevant insights that can be used:

- On the part of the company, (1) to analyse if its CSR communication is being effective or not, and (2) to identify the strengths and weaknesses of its CSR practices according to the expectations and wishes of its stakeholders. If a company is engaged in CSR activities, it should identify stakeholders needs and expectations, as well as show them its good CSR practices, and thus improve the company's image, reputation and relations with its stakeholders (Etter, 2013). There are proven positive influences to be gained by companies that engage in CSR and communicate information about it, and stakeholders need to receive that information in a complete and updated way (Henna & Chalmeta, 2017).

- On the part of the stakeholders, they could compare the assessment of company's CSR practices according to that expressed by them with the CSR information of the company generated by the company itself. It will allow to determinate if the company is carrying out greenwashing practices.

However, until now, there are no studies that use Big Data technics to analyse the information about companies' CSR practices on the online social networks. Its use has been recommended by a few researchers but only in a conceptual way (Farache et al., 2017; Crawford et al., 2013). To solve this problem, a three-phase organised method called the CSR-IRIS method is proposed. In a first phase, the stakeholders' tweets about CSR are collected and stored. In a second phase, the relevant CSR tweets are cleaned, treated and selected to assess the economic, social, environmental and governance aspects of a company. Finally, in a third phase, the company's CSR performance is assessed by applying sentiment analysis techniques to those tweets. As a case study, the method is applying to a leading company in the textile market sector.

This paper is structured as follows. Section 2 reviews the theoretical and empirical literature related to CSR communication, sustainability reports, online social networks and Big Data. Section 3 describes the CSR-IRIS Big Data method, which allows the CSR

of a company to be evaluated based on the information that stakeholders post on online social networks. Section 4 shows the application of the method to a case study. Finally, section 5 discusses the CSR-IRIS method innovation and shows the conclusions, research limitations and future work.

2. LITERATURE REVIEW

2.1 CSR communication

Corporate Social Responsibility implies that the company must assume a corporate behaviour that is consistent with the standards, values and expectations of its stakeholders and not only seek an economic benefit (Fernández, R., 2005). This CSR behaviour of the company can be motivated mainly by three factors: one is the pressure of stakeholders (such as investors, shareholders, customers, employees and non-profit organisations) to alleviate the enormously harmful environmental and social impacts that are being generated, such as deterioration of the environment, scarcity of resources, increase in the amount of waste generated, increased pollution or poor working conditions (Eding & Scholtens, 2017); a second factor is to generate brand value and increase the company's reputation, which serves as a differentiating element against competitors (Bebbington et al., 2008); and finally, the third factor results from the increasingly restrictive regulations (Srivastava, 2007).

Companies' stakeholders not only expect this CRS behaviour but also demand transparency (Gray, 2006) and disclosure of information that enables them to evaluate companies' CSR impacts (Michelon & Rodrigue, 2015). Therefore, companies have to communicate their CSR issues to their stakeholders. To do so, companies are using mainly three ways: sustainability reports, media advertisements, and information on the Internet.

Sustainability reports aim to provide a balanced and reasonable image of a company's CSR performance. They include the results achieved, the improvements and the challenges in three areas: economic, social and environmental. At the same time, they are a public declaration of commitment that makes it possible to increase the company's credibility regarding good governance and information transparency (Global Reporting Initiative, 2016). Sustainability reports are developed by the companies themselves. So, they may have biases that tend to enhance the company's reputation by prioritising the dissemination of good CSR practices and hiding weaknesses, which favours the interests of the entity instead of providing objective data (Deegan, C., 2002). And thus what is known as *greenwashing* is produced. Another criticism that is made is that they provide long and excessive narrative information which makes them difficult to analyse and understand (Michelon et al., 2015), although companies with stronger CSR performance are more likely to have CSR reports with higher readability (Wang et al., 2018). Finally, the structure and organisation of sustainability reports is decided by each company, so it becomes very difficult to compare the CSR of one company with that of another.

Another means of CRS communication by the company is advertising. Research on the impact of CSR advertisements on stakeholders is scarce (Pomering et al., 2013). It is considered an aggressive tool (Morsing & Schultz, 2006), although it has a strong persuasive and informative capacity (Garcia-De los Salmones & Pérez, 2018). When

stakeholders are exposed to a CSR advertisement, they may feel both positive and negative emotions. This contradiction is due to scepticism and suspicion on the part of the audience, who may see the advertisements as an "image-washing" campaign or a mere intention to improve the corporate image, thus arousing negative feelings (Elving, 2013).

Finally, companies can carry out their CSR communication on their corporate websites and on online social networks. Nowadays, the internet is a significant tool for CSR communication and companies are creating a major CSR presence on their websites and their online social network profiles. The main difference between the two tools is that websites are the best way to disseminate information, and online social networks allow greater ease of access to dialogue and communication about CSR among users (Kemna, 2013; Capriotti, 2011; Castelló et al., 2013; Colleoni, 2013). Therefore, online social networks seem to be more appropriate for CSR communication because, as Morsing and Schultz (2006) state, there must be a solid bidirectional communication that guarantees interaction with the company's stakeholders and allows the company to know the stakeholders' necessities and the subsequent development of the actions that are needed to fulfil them.

Despite the different forms of disclosure of CSR issues by companies, several studies show that there is no effective CSR communication from companies due to the fact that it tends to lack relevance, credibility, neutrality, completeness and interactivity (Cortado & Chalmeta, 2016; Ettinger et al., 2018; Knebel & Seele, 2015; Cho et al., 2018; Garcia-Torea et al., 2019). There are deficiencies in the quantity and quality of the information and its materiality, the channels of communication with the stakeholders are sometimes neither fluid nor sufficiently operational, and bilaterality in the dialogue with stakeholders fails (Ancos, 2014). CSR communication seems to be a symbolic practice to improve company reputation (Michelon et al., 2015). Stakeholders need CSR information about the companies that they can interpret and use to help them make informed and responsible consumption, purchase or investment decisions, and the current forms are failing.

2.2 Big Data to analyse CSR communication in online social networks

At the same time that companies are using online social networks to communicate their CSR practices, their stakeholders are using online social networks to share their opinions and thoughts about companies. These contents related to CSR that stakeholders post on the online social networks can be answers to companies' CSR posts as well as information or opinions that stakeholders post about the company and that they want to share with others. These contents are freely accessible and can be collected, processed and analysed through Big Data analytics.

Big Data analytics are technologies and techniques used to analyse large-scale, complex data from various applications in order to acquire intelligence and to extract unknown, hidden, valid and useful relationships, patterns and information (Adams, 2010). The application of Big Data to the CSR information posted by company's stakeholders could be an alternative method to obtain knowledge about the CSR performance of a company. It will allow:

- On the one hand, companies can obtain information about what their stakeholders say, think and do; and, therefore, by using these stakeholder insights from online social networks, companies can make better strategic CSR decisions, increase the

speed of analysis, provide a more adapted and focused response to a target audience or problem and, ultimately, create more value for the community (Farache et al., 2017; Crawford et al., 2013). To monitor online opinions about the company should be a priority for companies because the attitudes of the stakeholders towards the company may be affected by being exposed to these opinions (Song & Jing, 2019). Stakeholders that use online social networks can become active subjects who take part in the companies' CSR-image formation process through the content they generate on online social media (Acuti et al., 2020).

• On the other hand, stakeholders have the opportunity to know information about CSR practices of companies beyond the information provided by the own companies and lobbies, thereby avoiding the tendency they have to offer their own information and interpretation. It also favours the development of collective solutions by the different stakeholders, once the CSR problems are known, and promotes transformative options such as the fight against corruption, climate change or poor social conditions.

However, in our knowledge, there are no studies showing the possibilities of Big Data tools to analyse the CSR information posted by the stakeholders of a company on the online social networks. In this context, six research questions are proposed, related with the Big Data analysis of the opinions and behaviour of company's stakeholders on the online social network Twitter:

RQ1: *Is it possible to assess the quality of the CSR of a company by analysing the information available on the internet and, more specifically, from the opinions expressed about a company by its stakeholders on the online social network Twitter?*

RQ2: W*hat dimension (Economic, Social, Environmental or Governance) is the one that generates the greatest number of stakeholders' tweets on the social network Twitter*?

RQ3: *Is there any relation between the appearance of news related to a company on the Internet, produced by other entities such as digital press or news agencies, and the number of CSR issues posted on Twitter by company's stakeholders*?

RQ4: *Is there any relation between the number of CSR tweets posted by the stakeholders of a company and the publication of its sustainability report*?

RQ5: *is Twitter used by a company's stakeholders more to inform than to provide opinions*?

RQ6: *Does the CSR information showed by the company in its sustainability report match the CSR information expressed by stakeholders on Twitter?*

3. PROPOSED METHOD TO ASSESS THE CSR OF A COMPANY USING INFORMATION AVAILABLE ON THE INTERNET

Although there are researchers who are proposing the use of Big Data to analyse the CSR of companies, no practical methods to guide this process have yet been put forward. For this reason, and based on the experience of the authors of this paper in the application of Big Data in business management and in the development of information systems for

sustainability management, the Integration and Re-engineering group of the Universitat Jaume I of Castellón, Spain, has developed the CSR-IRIS method, which allows the quality of the CSR performed by a company to be assessed based on information available on the internet, specifically on the basis of the opinions expressed about a company on the social network Twitter. The method is organised into three phases called data collection, information processing and sentiment analysis, and makes it possible to:

- Determine whether there is CSR information about a company on social media
- Identify which dimensions of the CSR are the most commented
- Evaluate if the comments made about a company's CSR are positive or negative

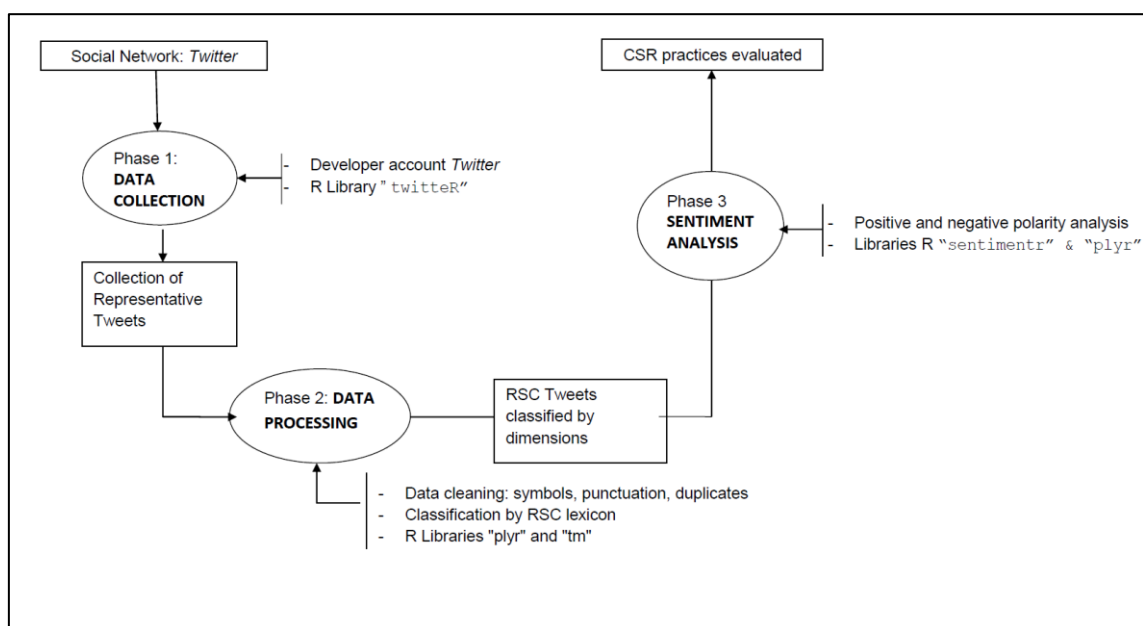The three phases of the method are described below (Figure 1):

Figure 1. Phases of the CSR-IRIS method

3.1. Data collection

The first step of the method is to collect data from the online social networks. In this study, Twitter data are used. Twitter is a well-known micro-blogging platform that helps users to share information (text, images, videos, GIFs, links, hashtags, headlines, etc.) in units of 140 characters, now upgraded to 280 characters called tweets. This social network acts passively, collecting data. However, a tweet is more than just a message; it contains a large number of metadata that help to understand and add meaning to each item of data (data-time, number of retweets, etc.). The reasons for using Twitter are that (1) tweets can be accessed freely; (2) tweets can be downloaded around themes, keywords or timelines; (3) Twitter is one of the fastest growing social media platforms with a very high number of users and reach; (4) users in Twitter follow others, post, like and share their updates publicly, more frequently than that on other platforms like Facebook; and (5) Twitter is the one of the most online social networks used to communicate CSR issues (Cortado & Chalmeta, 2016; Szumniak-Samolej, 2019).

From the technological point of view, there are different alternatives to connect and extract information from Twitter. On the one hand, there are specialised paid search engines such as Keyhole, TweetBinder and TweetReach. On the other hand, there are free tools as the programming language R. This second option is the one that was selected in this work because (1) the R language is free, freely available, can be used on the vast majority of operating systems and we did not have any funding for the project; (2) R has tools, frames and packages to process, analyse and visualise any type of data; and (3) R has the twitteR option prepared to obtain and process Twitter data very easily (Gentry, J., 2015). Annex 1 shows the technical procedure followed for the extraction of tweets using R, with the possibility of adjusting the query by filtering the user, the keyword, the language and the location. As a result, a database is obtained with the tweets that contain the message, the users, the entities that include information in URLs, hashtags and the geographical location where the message was posted.

3.2. Information processing

Once the sample has been obtained, data are pre-processed, manipulated, cleaned, formatted and filtered to eliminate content that is not relevant to the study. There is no single way to perform this information processing. In the CSR-IRIS method, two specific RStudio packages are used for pre-processing, cleaning and formatting, due to their suitability and ease of use. One of them is the "plyr" library by Wickham (2016), which can be used to divide a large data structure into homogeneous parts, apply a function to each part and then combine all the results again. The other corresponds to the library "tm" by Feinerer and Hornik (2017), which allows duplicate tweets to be eliminated through the unique() function, and the gsub() function to eliminate non-informative patterns such as urls from web pages (gsub ("http \\ w + "," ",)), exclamation marks (gsub (" (RT | via) ((?: \\ b \\ W * @ \\ w +) +))), and different characters or numbers (gsub ("[[: digit:]]", "",).

The data is filtered using the laply() function of the library (plyr). This function allows the tweets that include a certain word related to CSR to be selected and classified. Since there is no specific lexicon for the filtering of tweets with CSR content, one was developed for this purpose. This lexicon was elaborated in Spanish, leaving its development in other languages for future extensions. Table 1 shows the lexicon translated into English. This lexicon proposes a list of 133 different words related to CSR that were obtained from the analysis of the GRI Standards guides (2016). The 133 words are classified into five dimensions of CSR: General, Governance Standards, Economic Standards, Environmental Standards and Social Standards. Those words that can be classified in more than one dimension, such as "customers", "community" or "suppliers", have been included in the general dimension to avoid duplication.

Table 1. CSR lexicon keywords (translated into English)

Generals	Governance Standards	Economic Standards	Environmental Standards	Social Standards
Customers	Shareholders	Acquisition	Water	The right to be member of a union
Climate	Administration	Anti-corruption	Air	Hiring
Code	Authority	Benefits	Natural	Data
Community	Committee	Capital	Biodiversity	Rights

Conduct	Conflict	Competition	Conservation	Discrimination
Corporate	Board of directors	Competitiveness	Consumption	Diversity
Fulfilment	Corporation	Purchasing	Contamination	Employees
Performance	Corporate	Unfair	Eco-efficiency	Employment
Standards	Corruption	Economic	Effluents	Company
Ethics	Managers	Executives	Emissions	Teaching
GRI	Executive	Financial	Energetic	Labelled
Impact	Dividends	Expenses	Energy	Training
Report	Strategy	Infrastructures	Species	Equality
Legislation	Fiscal	Income	Gas	Indices
Materiality	Management	Investments	Habitats	Labour
Rules	Governance	Retirement	Materials	Marketing
Normative	Government	Suppliers	Environmental	Dangers
Opportunity	Interests	Salary	Environment	Permission
Values	Goals	Services	Ozone	Personal
Suppliers	President	Wages	products	Benefits
Reputation	Remuneration	Earnings	Recycling	Prevention
Responsibility	Responsible		Reprocess	Privacy
Risk	Bribery		Resources	Professional
CSE	Taxation		Waste	Health
CSR			Reused	Security
Sustainability			Supply	Social
Sustainable			Substances	Socioeconomic
Transparency			Waste	Workers
Values				Job
Verification				Volunteering

3.3. Sentiment Analysis

Once the data have been extracted and prepared (elimination of duplicates or symbols that do not add any value, such as punctuation or structural words like conjunctions or prepositions, etc.), in the third step of the CSR-IRIS method, sentiment analysis techniques are applied. Sentiment analysis is a compendium of computational methods from various fields, including linguistics, statistics, machine learning, text summarization, text mining, etc. which are used to categorize information from natural language text in the digital world, such as social networks, forums, websites, etc. (Liu, 2015). This information includes the opinions, feelings, assessments, attitudes and/or emotions of people towards certain entities, products, services, organisations, individuals, problems, events or topics.

With the application of sentiment analysis to the processed information obtained in the step 2 of the CSR-IRIS method, it is possible to extract a tangible and direct value for a tweet related to CSR and categorise the opinion expressed in it in three categories: positive, negative or neutral.

There are several techniques for estimating feelings in the opinions (Gong et al., 2019). The CSR-IRIS method is based in the *Naïve Bayes algorithm*, which is the most used machine learning algorithm for this purpose and has proven to give good results for sentiment analysis on Twitter (Ravindran & Garg, 2015). It provides a result derived from

the number of times a positive or negative word appears in a sentence (in this case in a tweet). In this way, the sentiment score is obtained as:

Sentiment = Number (positive words) - Number (negative words)

If the value of the feeling is less than zero, it is interpreted as a negative feeling in the final assessment of the tweet. If the value of the feeling exceeds zero, it is considered a positive feeling. Finally, if it is not possible to identify any of the words designated as positive or negative in a tweet, it is considered a neutral feeling. This situation can also arise, even though some words are designated as positive and negative, when compensation occurs between them. Therefore, the polarity is reflected as follows:

$$
\text{Polarity} = \begin{cases} \text{If Sentiment} < 0 & \text{- Negative} \\ \text{If Sentiment} = 0 & \text{- Neutral} \\ \text{If Sentiment} > = & \text{- Positive} \end{cases}
$$

With this method it is not possible to establish a maximum positive or negative margin of feeling. Twitter has a maximum number of characters per tweet. So, the number of positive or negative words that are identified in a tweet will depend on the length of the words and the way determiners, prepositions, conjunctions or spaces are used to prepare the message.

Although there are different programs than can help in the application of the sentiment analysis, such as Hootsuit insights, MeaningCloud or SentiStenght, the CSR-IRIS method uses R. R has been chosen because (1) R is free (some of the existing programs have free versions, but they are poorer than the commercial versions); (2) there are free libraries than can be used for the sentiment analysis in R, and (3) therefore, all the computer programs used in the CSR-IRIS method are integrated in a unique package developed with the same tool: R. The "plyr" library by Wickham (2016) and the laply() function are used to programme in R the feeling analysis:

feeling = laply (tweets, function (positive words, negative words, progress = "none")

The application of the sentiment analysis technique requires a lexicon of positive and negative words. There are different lexicons in the literature that can be used and the choice will depend on each specific case. For example, the library (tidytext) by De Queiroz et al. (2019) of the RStudio package contains three lexicons in English: AFINN proposed by Nielsen (2011), which assigns a numerical value to each word between -5 (extremely negative) and 5 (extremely positive); Bing, proposed by Liu Bing (2004), which categorises the words by assigning a positive or negative value; and NRC, proposed by Mohammad and Turney (2013), which establishes a word–feeling relationship: joy, trust, fear, negative, sadness, anger, positive, aversion and surprise. In the CSR-IRIS method, positive and negative words in each tweet are identified using the iSOL lexicon (Molina-González et al., 2013), which is based on Bing. This lexicon was used here because the tweets were in Spanish and iSOL contains a list of 2,509 positive and 5,626 negative words based on the lexicon of Bing Liu (2004).

4. CASE STUDY. APPLICATION TO INDITEX

To facilitate the understanding of the CSR-IRIS method, it was applied to Inditex, a group of companies in the textile market sector, as shown below. This sector was chosen due to it is one of the most polluting in addition to generating social and occupational risks (Slater, 2000), while at the same time it is also one of the sectors that is currently developing more sustainability initiatives (Arribas et al., 2013). In addition, it is a sector closely related to social networks such as Facebook (2004) or Twitter (2006), since the monitoring of any fashion event is experienced in real time, which gives brands greater accessibility and visibility. Inditex was chosen for being the world leader in the textile and fashion market sector who marks the step in the word of retail, with more than 120 million of followers on social networks in 2017, and a strong CSR compromise (Amor-Esteban et al., 2020). Inditex is a member of the Gold Community Pioneers Program of the GRI community and executes initiatives to achieve the sustainable development of the sector, such as the recently signed Fashion Pact (2019). The Inditex group was founded in Spain 1975 by Amancio Ortega and has eight global brands (Zara, Bershka, Pull & Bear, Massimo Dutti, Oysho, Uterqüe, Stradivarius and Zara Home) with a network of 7,490 stores in 96 countries and 156 markets with online stores.

The objectives that are intended to be achieved with the application of the CSR-IRIS method to Inditex are:

- Objective 1: To check whether users talk about the CSR of the Inditex group or not on the social network Twitter.
- Objective 2: If users do talk about the CSR of the Inditex group, what dimension (Economic, Social, Environmental or Governance) is the one that generates the greatest number of tweets?
- Objective 3: To determine whether there is a relationship between the appearance of news related to Inditex in digital media such as digital press or news agencies, and the number of Inditext CSR tweets.
- Objective 4: To check whether the CSR of Inditex is discussed more on the social networks when the Inditex group publishes its sustainability report.
- Objective 5: To confirm the fact that, regarding CSR issues of the Inditex group, Twitter is used more to describe some factual information than to express subjective opinions. In this work, according to Hu & Liu (2014), if a tweet contains one or more opinion words is considered an opinion tweet. Otherwise is an informational tweet.
- Objective 6: To check whether the opinions expressed in the social network coincide with what the Inditex group expresses in its sustainability report.

The following shows how the CSR-IRIS method has been applied to the Inditex group to achieve the six objectives outlined above.

4.1. Collection of data on Inditex

The first part of the method consists in obtaining the data to study, that is, a sample of tweets that is collected by synchronising the social network Twitter and the program RStudio. In addition, it is necessary to install the "twitteR" package, because it has the function searchTwitter(), which allows searches by keyword.

The function was defined as: searchTwitter (searchString = "search keyword", n = 10000, lang = "es"), where the search keywords were the name of the group and each of its brands (Inditex, Zara, Pull & Bear, Massimo Dutti, Bershka, Stradivarius, Oysho, Zara Home

and Uterqüe) and Spanish as the search language. We have selected only CSR tweets in Spanish because the sustainability report is written only in Spanish, and we are going to compare them. In addition, only stakeholders'' tweets were considered. Tweets from official company and brands accounts were removed to avoid that polarity analysis was influenced by positive messages from the company. However, retweets to official company and brands accounts were considered. The search for tweets was carried out over a period of 4 months (from 1 April 2019 to 31 July 2019) and 306,845 tweets were obtained. This period was chosen because it contains the publication of the Inditex sustainability report for 2018. Table 2 shows the total number of tweets in Spanish collected for each keyword.

Table 2. Total number of tweets in Spanish per keyword

Inditex	Zara	Pull& Bear	Massimo Dutti	Bershka	Stradivarius	Oysho	Zara Home	Uterqüe	TOTAL
74.174	160.160	16.413	977	31.679	17.919	3.694	502	1.327	306.845

4.2. Information processing

Once the sample had been obtained, it was necessary to carry out a data processing process, which first involved cleaning the sample to eliminate possible duplicates. For example, Twitter users often repeat a comment followed by a link to a web page (url), which can be interpreted initially as a different tweet without it really being a new one. Repeatedly using the unique() and gsub() functions, the symbols (urls, @, digits, exclamation marks, etc.) and duplicate tweets were eliminated. As a result, the sample size was reduced to 13,480 tweets that contained all the tweets, without duplications, referring to the Inditex group published worldwide in Spanish on the social network Twitter during the 4 months under study.

The next step was to select the CSR tweets and to classify them by categories using the lexicon shown in Table 1. The results were 4,649 tweets containing CSR words from the lexicon in Table 1. Table 3 shows the number of tweets (column No.T) for each word in the lexicon and the percentage of tweets for each dimension.

Table 3. Classification of tweets by keyword of the CSR lexicon

Generals			Governance			Economic			Environmental			Social		
	NºT	T.C		NºT	T.C		NºT	T.C		NºT	T.C		NºT	T.C
Customers	35	7	Share holders	96	4	Acquisition	-	-	Water	43	5	Association	1	-
Climate	-	-	Administration	5	-	Anti-corruption	-	-	Air	1	-	Hiring	-	-
Code	1	-	Authority	-	-	Benefits	54	4	Environmental	36	6	Data	22	5
Community	8	1	Committee	1	1	Capital	155	1	Biodiversity	-	-	Rights	137	15
Conduct	3	2	Conflict	-	-	Competition	8	4	Conservation	-	-	Discrimination	-	-
Corporate	11	4	Board of direction	36	2	Competitiveness	1	-	Consumption	2	-	Diversity	17	8

Fulfilment	3	2	Corporation	4	-	Purchasing	74	2	Contamination	15	1	Employees	217	28
Performance	-	-	Corporate	1	-	Unfair	1	2	Eco-efficiency	-	-	Employment	196	22
Standards	-	-	Corruption	-	-	Economic	1	2	Effluents	-	-	Company	426	32
Ethics	14	3	Managers	1	1	Executives	6	-	Emissions	3	3	Teaching	-	-
GRI	8	-	Executive	11	2	Financial	3	1	Energetic	-	-	Labelled	-	-
Impact	36	8	Dividends	29	3	Expenses	10	-	Energy	23	8	Training	28	13
Report	71	5	Strategy	131	21	Infrastructures	-	-	Species	-	-	Equality	35	10
Legislation	1	-	Fiscal	372	28	Income	87	-	Gas	-	-	Indices	-	-
Materiality	-	-	Management	1	2	Investments	16	1	Habitats	-	-	Labour	33	11
Rules	59	-	Governance	-	-	Retirement	-	-	Materials	21	5	Marketing	13	-
Normative	2	-	Government	20	2	Suppliers	186	8	Environmental	30	5	Dangers	2	-
Opportunity	2	-	Interests	2	-	Salary	-	-	Environment	8	2	Permission	-	-
Values	11	-	Goals	14	3	Services	84	11	Ozone	-	-	Personal	65	10
Suppliers	99	11	President	14	1	Wages	338	17	Products	13	7	Benefits	-	-
Reputation	11	11	Remuneration	-	-	Earnings			Recycling	3	3	Prevention	-	-
Responsibility	4	4	Responsible	16	1	Economic Standards			Reprocess	4	2	Privacy	-	-
Risk	1	-	Bribery	-	-	Acquisition			Resources	4	-	Professional	2	1
CSR	20	5	Taxation	-	-				Waste	3	2	Health	39	6
ESR	246	-							Reused	-	-	Security	46	12
Sustainability	168	64							Supply	2	2	Social	91	35
Sustainable	137	42							Substances	-	-	Socio-economi	-	-
Transparency	9	2							Waste	-	-	Workers	254	54
Values	2	-										Job	74	15
Verification	-	-										Volunteer.	-	-
N.º tweets by dimension	962			754			1024			211			1.698	
N.º tweets after manual check by dimension	171			71			53			51				277	
% tweets per dimension	**27.4**			**11.4**			**8.5**			**8.2**				**44.5**	

A subsequent content analysis revealed that although those 4,649 tweets contained words from the CSR lexicon in Table 1, not all of them speak about a CSR subject. Therefore, a manual filtering was carried out to eliminate those whose theme was not related to CSR,

leaving a final sample of 623 tweets. Table 3 shows the number of CSR tweets after a manual check (column T.C.) for each of the CSR lexicon words and the percentage of these tweets for each dimension.

In this way, an answer was given to:

- Objective 1, by verifying that on the social network Twitter users comment on the CSR of the Inditex group. If the sample of 13,480 tweets obtained after removing duplicate tweets is taken as the initial sample, it is seen that issues related to CSR are addressed in 623 of them, which is equivalent to 4.62% of the sample.

- Objective 2, by identifying that the social dimension generates the greatest number of tweets, with 277 tweets and 44% of the sample of CSR tweets.

Figure 2 shows the temporal evolution of the RSC tweets throughout the study period. There are four moments in time when the volume of tweets is higher.
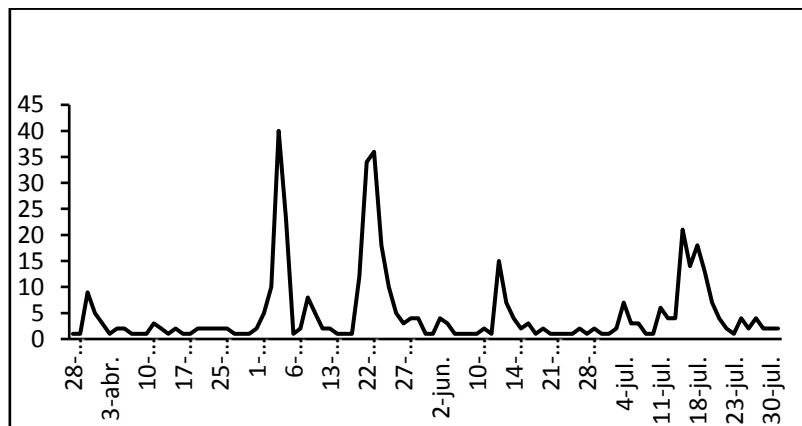
Figure 2. Temporal evolution of the CSR tweets throughout the study period

After analysing the articles published in the online and written press, it was discovered that these moments coincided with the publication in these media of relevant articles that refer to the Inditex group and its founder (Figure 3).

Figure 3. Press articles related to the Inditex group in the period studied. *Translation from left to right and from top to bottom: 2 May 2019, Amancio Ortega receives 813 million [euros] today in Inditex dividends; 20 May 2019, Inditex has paid 360 million in corporate tax; 21 May 2019, Amancio Ortega donations spark a debate between Podemos and the PP and Ciutadans*

[Citizens] [three Spanish political parties]; 16 July 2019, Inditex will strengthen its strategy to be a leader in sustainable transformation [This article refers to the Inditex sustainability report published on 16 July]; 16 June, Inditex earnings rise 10% in the first quarter of the year to 734 million euros.

In this way, an answer was given to:

- Objective 3, by verifying that there is a relationship between the appearance of issues related to CSR in other media and an increase in the number of comments on the social network Twitter.

- Objective 4, by identifying that the highest volume of comments on CSR does not occur when the company publishes its sustainability report. The sustainability report of the Inditex group for the tax year 2018 was published on 16 July 2019. Although there is a significant increase in publications on that date in relation to other days, it is not when more tweets were posted.

4.3. Sentiment analysis

The last step in the method is the sentiment analysis of each of the tweets. Table 3 shows several examples of the 623 tweets to which the corresponding sentiment analysis based on polarity was applied. Each positive word is assigned a value of 1 and each negative word is assigned a value of -1. Sentiment is the difference between the number of positive words and negative words.

Table 3. Examples of tweets to which sentiment analysis based on polarity was applied

Tweet	Positive	Negative	S	Polarity
Donations, tax **evasion,** labour **exploitation** ... Inditex **accusations** that shake the campaign	0	3	-3	Negative
Those who work in Spain in Inditex stores are **exploited** and with **miserable** salaries	0	2	-2	Negative
Super professionals ZARA online purchase customer service. **Fast** and **efficient**, that's the way to please customers.	2	0	2	Positive
Congratulations Inditex for the **ambitious** plan in which **sustainability** is the protagonist. A clear example of **commitment**	3	0	3	Positive
Inditex will be 100% **sustainable** in 2025. The strong commitment to **sustainability** of the world **leader** in textiles is **fantastic** news	4	0	4	Positive

The results obtained by applying sentiment analysis to the 623 tweets range between -3 and 4. The polarity interval depends on the number of positive and negative words included in the text. As Twitter limits the content of its message to 140 characters, the interval in this case is small. In longer texts, with a greater number of words, the interval could be longer. Figure 4 shows the number of tweets for each of the possible polarities.
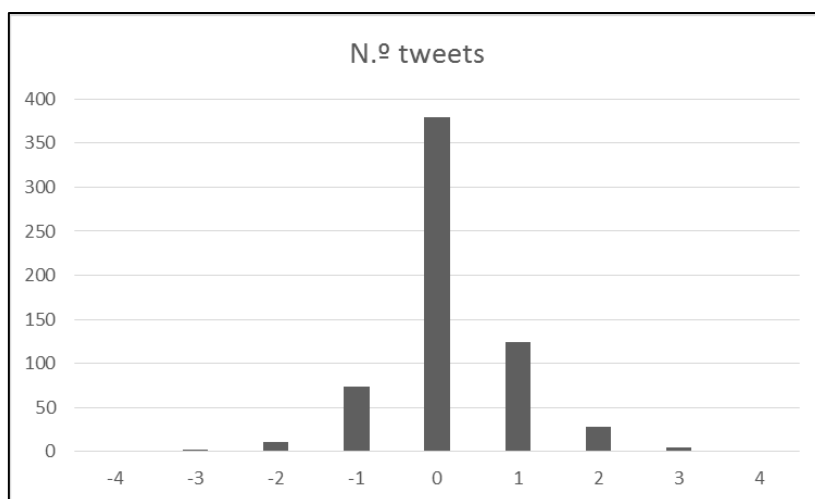
Figure 4. Number of tweets for each one of the possible polarities

In the sample of 623 tweets, the majority have a neutral polarity, that is, they report information or facts without any opinion. In addition, positive comments outnumber negative comments. So:

- Objective 5 is verified by confirming that, on Twitter, CSR tweets are used to inform more than to provide opinions, since there are more tweets of a neutral character.

4.4. Comparison between the tweets assessment and the sustainability report

An important consequence of the application of the method is that it is possible to compare the assessments on different CSR topics of a company obtained by analysing the comments on Twitter with the company sustainability report. This allows the company to generate a feedback mechanism to determine the CSR topics possibly in need of improvement, or to check whether the CSR practices carried out by the company are transmitted in a correct way or not.

The sustainability reports of the Inditex group for 2017 and 2018 were organised around ten material issues: Our clients, Our people, Innovation/Integral management of the supply chain, Socially responsible supply chain, Excellence of our products, Circularity and efficient use of resources, Fiscal transparency, Contribution to community welfare, Creating value for our shareholders, and Corporate governance and corporate ethical culture.

To carry out the comparison, it is necessary to classify the 623 tweets on the material issues of the Inditex sustainability report manually by performing a content analysis. Table 4 shows the number of tweets on each material issue and their percentage with respect to the total.

Table 4. Number of tweets per material issue

	Our clients	Our people	Innovation/Integral management of the supply chain	Socially responsible supply chain	Excellence of our products

Nº. Tweets	40	214	0	25	178
Percentage	6,42%	34,35%	0,00%	4,01%	28,57%
	Circularity and efficient use of resources	Fiscal transparency	Contribution to community welfare	Creating value for our shareholders	Corporate governance and corporate ethical culture
Nº. Tweets	15	58	48	31	14
Percentage	2,41%	9,31%	7,70%	4,98%	2,25%
TOTAL Nº. Tweets: 623					

As an example, the following shows the comparison between what was obtained after the analysis of sentiment by polarity and what is outlined in the 2017 Inditex sustainability report on these issues for the three material issues with the highest number of tweets (our people, excellence of our products, and fiscal transparency). The 2017 sustainability report was used because that of 2018 was published at the end of the sampling period, so it should have less effect on the content of the CSR tweets that were posted. In any case, the contents of the material issues in both reports are very similar.

Our people:

Sustainability Report: For the Inditex group the satisfaction of its employees is presented as one of its priorities. Therefore, the company highlights its commitment to offering safe, healthy, diverse and inclusive work environments. In addition, Inditex develops training plans for its employees and defends their quality employment. In fact, the company is considered "the best company to work for in Spain", according to Merco Talent ranking, a reputational assessment instrument launched in 2000, based on a multi-stakeholder methodology composed of six assessments and twenty-five sources of information (Merco Talent, 2020).

Internet assessment: This is the material issue that generates a greater number of comments on Twitter (214 tweets). Valuations are mostly neutral or negative. Twitter users associate the company with irregularities, exploitation, slave girls or tax havens.

The classification of tweets by polarity is shown in Figure 5 and examples of the comments are shown in Table 5.
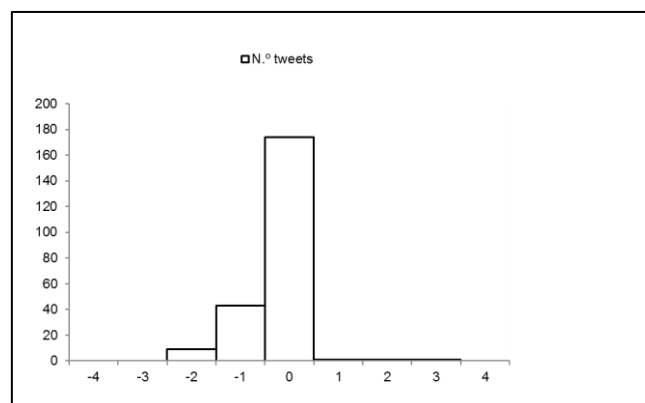

Figure 5. Distribution of Sentiment Analysis in tweets related to *Our People*

Table 5. Examples of tweets related to *Our People* and their polarity**Tweet**	**Positive**	**Negative**	**S**	**Polarity**

Inditex employment is not quality, **irregular** hours and **exploitation,** we have little class consciousness	0	2	-2	Negative
Inditex workers in Asia receive salaries above the average in their countries	0	0	0	Neutral
I have personally met several Inditex workers and they were quite **happy**	1	0	1	Positive

Excellence of our products:

Sustainability Report: The Inditex group strives every day to offer its customers products with the highest standards of quality, health and sustainability. Sustainable products, commitment to clean and responsible fashion, recycled raw materials or commitment to the environment are some of its objectives in this material issue.

Internet assessment: The actions and standards that the group has set in this material issue are appreciated since positive evaluations prevail. However, negative comments are also collected on aspects such as environmental pollution, which is significant because in the sustainability report Inditex cites specific actions to minimise this type of impact.

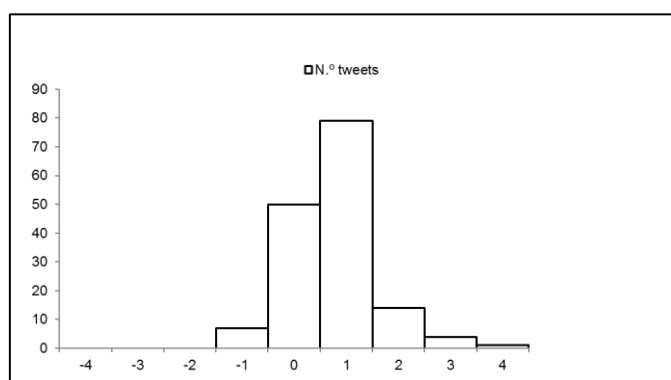The classification of tweets by polarity is shown in Figure 6 and examples of the comments are shown in Table 6.

Figure 6. Distribution of Sentiment Analysis in tweets related to *Excellence of our products*

Table 6. Examples of tweets related to *Excellence of our products* and their polarity

Tweet	Positive	Negative	S	Polarity
Things that make me laugh: people wearing Inditex clothes complaining about environmental **pollution**	0	1	-1	Negative
Companies that still do not meet the SDGs using corals and shells in their Zara, Mango, H&M products	0	0	0	Neutral
Inditex will be 100% **sustainable** in 2025. The strong commitment to **sustainability** of the world **leader** in textiles is **fantastic** news	4	0	4	Positive

Fiscal Transparency:

Sustainability Report: The Inditex group is committed to contributing to economic, social and industrial development through compliance with the tax legislation of the countries in which Inditex is present.

Internet assessment: In this CSR issue, which refers to information related to the group, neutral assessments prevail. Comments are also focused on possible tax evasion or fraud

by the company, aspects which are only mentioned in the sustainability report as issues to be fought against.

Figure 7 shows the classification of tweets by polarity and examples of the comments are shown in Table 7.
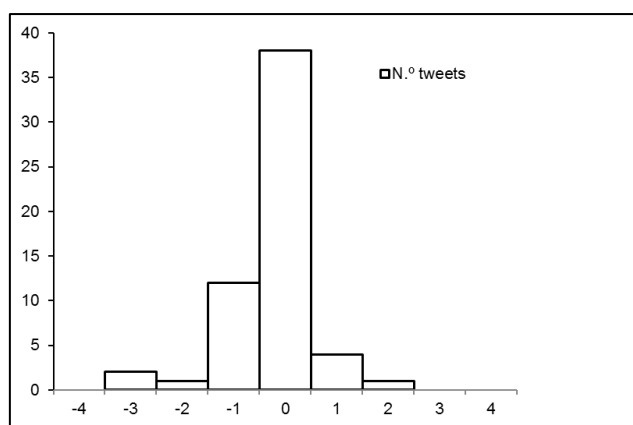

Figure 7. Distribution of Sentiment Analysis in tweets related to *Fiscal transparency*

Table 7. Examples of tweets related to *Fiscal transparency* and their polarity

Tweet	Positive	Negative	S	Polarity
Donations, tax **evasion,** labour **exploitation**... **Accusations** against Inditex that shake the campaign	0	3	-3	Negative
Inditex dodged 600 million in taxes thanks to tax engineering	0	0	0	Neutral

The above is an example of how to process the information to compare the opinions expressed by Twitter users with the CSR information provided by a company in its sustainability report. In this way, an answer was given to:

- Objective 6, by verifying that the opinions expressed on the social network Twitter mostly do not coincide with what the company expresses in its sustainability report. This may be due to the fact that Twitter users have a distorted vision in relation to the CSR actions carried out by the company and express their opinion without contrasting it with what is shown in the report, or that the company performs greenwashing in its sustainability report.

4.5. Comparison of the word clouds of the Inditex group sustainability report and opinions on the internet

In addition to sentiment analysis, another way to compare the contents of the sustainability report is through word clouds. Using the R wordcloud() function it is possible to obtain the most repeated words in the 623 tweets. The result is shown in Figure 8 and visually differentiates between what the iSOL lexicon considers positive (underlined) and negative (in a box). Word size also implies that it is repeated more times in the sample.

The five positive words that appear the most are sustainable, sustainability, leader, commitment and judiciary; the five negative words that appear most are evasion,

exploitation, precariousness, irregularities, and slaves; and the five neutral words that appear most are invest, taxes, support, labour and innovation.
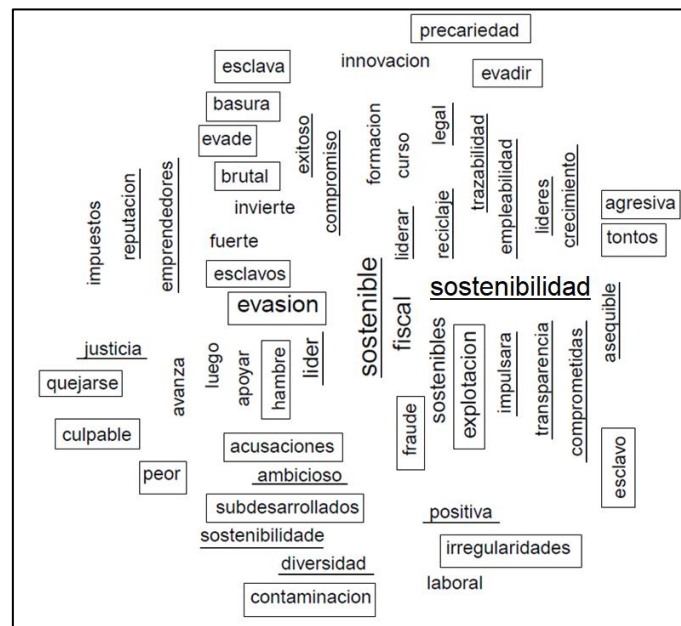

Figure 8. Word cloud of the 623 CSR tweets

On the other hand, Figure 9 shows different word clouds for each of the material issues of the 2017 Inditex group sustainability report. The size of each word is also associated with the frequency with which that word appears within the text. The fifteen words that appear most are: customers, chain, stores, data, social, programme, purchase, brands, investment, corporate, collaboration, good governance code, diversity, sustainable and development.

Figure 9. Word cloud of Inditex material issues

As can be seen by comparing the two figures, out of the fifteen words that are repeated most in the 623 tweets only *sustainability* appears among the most cited in the Inditex sustainability report. This could indicate that the keywords used by Inditex in the communication of its CSR through its sustainability report are not penetrating among the stakeholders, which would demonstrate that the communication of its CSR is not being carried out effectively.

5. DISCUSSION AND CONCLUSION

5.1 Contributions to theory

Enterprises are currently carrying out CSR practices and they need their stakeholders to know it because it can give companies a competitive advantage. Therefore, to communicate their CSR practices, companies are using different ways, being the online social network Twitter one of these ways. At the same time, company's stakeholders are tweeting about companies' CSR practices as well. These tweets can be answers to previous company's CSR tweets as well as opinions or comments that they post to express their personal opinions and views to others (Arora, Li, & Neville, 2015).

Therefore, there is valuable information in Twitter about a company CSR practice. This CSR messages have influence in company reputation, word of mouth, organizational identity, and stakeholder engagement (Grover et al., 2019; Zizka, 2017). However, until now, there are no studies that use Big Data technics to analyse the information about a company CSR practices on the online social networks, because its use has been proposed by a few researchers but only in a conceptual way (Farache et al., 2017; Crawford et al., 2013). Therefore, this article contributes to the existing literature by showing a method and a real example of how Big Data can be used to assess the information about companies' CSR practices expressed by their stakeholders on a social network. This information can be used both by companies to enhance their CSR communication and CSR performance, and by stakeholders to identify greenwashing.

The findings of the work shown in this paper have proved: *Firstly*, Knowledge about companies' CSR practices can be obtained using Big Data to analyse information about the company on online social networks (RQ1). To do this, we have proposed the CSR-IRIS method and we have shown an example of the application of this method to help other researchers to use it. The method makes it possible to obtain the sentiment analysis of online social network users regarding different CSR issues of a company, and therefore to identify its CSR weaknesses (CSR issues with negative assessments) and strengths (CSR issues with positive assessments). *Secondly*, the social dimension of CSR is the one that generates the greatest number of tweets among the stakeholders of the company (RQ2). *Thirdly*, there is a relation between the publication of CSR articles in the press and an increase in the number of CSR tweets by the stakeholders (RQ3). *Fourthly*, although there is an increase in the number of CSR tweets when the sustainability report is published, it is not the time when the most tweets are posted, even bearing in mind that articles about different subjects in the sustainability report are published in the press in that moment. So, the impact of the sustainability report is limited in those stakeholders that express their opinion on online social networks (RQ4). *Fifthly*, CSR tweets posted

by stakeholders are used to inform more than to provide opinions (RQ5). *Finally*, there are important differences between the contents of the sustainability report of a company and the CSR opinions of its stakeholders on the internet (RQ6). This could be due to the fact that CSR communication is not being carried out effectively and/or that the companies are engaging in greenwashing.

5.2 Managerial implications

These findings have important managerial implications. *Firstly*, regarding the CSR communication strategy of the company, the differences found between the sustainability reports and the stakeholders' opinions on the internet show that companies are not communicating their CSR practices in an efficient way, which has been stated by other authors such as Crawford et al. (2013) and Ancos (2014). Therefore, companies need to improve their CSR communication strategy, which must be based on interactivity with their stakeholders (Capriotti et al., 2011; Etter, 2014; Castelló et al., 2013). The CSR-IRIS method allows to identify in which CSR areas the company CSR communication on the Internet is failing and why. Therefore, companies can focus their communication efforts in these CSR areas, explaining their stakeholders what, why and how they are carrying out their CSR practices. For example, in the case of Inditex, the CSR practices related with the material issue *Our people* are mostly assessment as neutral or negative. Twitter users associate the company with irregularities, exploitation, slave girls or tax havens. However, Inditex has been considered "*the best company to work for in Spain*". This give information to Inditex managers about that in *Our people* material issue could be a bad communication of its CSR performance.

Secondly, the CSR-IRIS method helps to enhance the company CSR performance because it allows a continuous monitoring of the stakeholders' positive and negative opinions on the internet, which can be used to identify stakeholders' requirements and companies' strong and weak CSR points. Therefore, companies can establish their CSR strategy and CSR action plans to satisfy stakeholders' requirements in a more efficient and productive way. This will prove true interest of the company in taking into account the stakeholders' necessities and opinions, and will increase the engagement between the company and the stakeholders. For example, in the case of the Inditex *Our people* material issue, as Inditex is a multinational company, it could be that some of the bad CSR practices tweeted are really happened but the top managers of the company do not know it. With the use of the CSR-IRIS method, they can be conscious about it, and can make the suitable decisions to solve this problem.

Thirdly, companies' CSR solutions to stakeholders' requirements/suggestions/complaints etc. made through Twitter should be posted on Twitter because it favours the commitment of organisations with their stakeholders (Lovejoy et al., 2012) and companies that tweet more often about CSR are more likely to generate engagement among Twitter users (Araujo and Kollat, 2018). In addition, these company's CSR solutions should be also showed in the companies' annual sustainability reports, and other media such as digital press, because this work has proved that there is a positive relation between the publication of CSR articles in the digital press and an increase in the number of CSR tweets. This will allow to increase trust and long-term commitment with positively engaged stakeholders, as well as to reconcile with negatively engaged stakeholders, because stakeholders may move from positive to negative engagement or from negative to positive engagement (Luoma-aho, 2015).

Fourthly, findings are also useful for company's stakeholders because they can obtain information about company's CSR practices without the bias introduced by the company. CSR information about a company is generated by the company itself, whereas with the CSR-IRIS method, CSR information can be obtained in a more objective way from the stakeholders' opinions. For example, in the case of the Inditex *Our people* material issue, the mostly negative and neutral assessment give information to NGOs and other ESG rating agencies about that in this material issue Inditex could be doing greenwashing and bad CSR practices.

5.3 Limitations and future research

This work is an initial proposal to measure the sustainability of companies with the information that can be found on the internet. So, the CSR-IRIS method has several limitations that require future research:

a) There are no specific lexica on concepts related to CSR for the selection and classification of tweets. Therefore, for this work a lexicon has been proposed (Table 1) that could be reviewed and improved by other researchers. This lexicon was thought for applying in any kind of sector, but it could be improved adapting it to the special features of every type of business.

b) There is no specific CSR lexicon for the analysis of the sentiment of the content of tweets, so in this work a general purpose lexicon of positive and negative words (iSOL) has been used. The use of a specific lexicon could yield more accurate results.

c) Not all the followers on the online social networks are true stakeholders. There are fakeholders, which are opinions, socio-bots and stakeholders artificially generated by algorithms (Luoma-aho, 2015). They can pervert the results, so it is important to implement methods that detect fakeholders and to remove their opinions from the collected data.

d) Sentiment analysis helps to identify possible trends in users' opinions. However, the complexity of human language means that it is not an exact science and that it is not possible to perceive nuances such as sarcasm or the double intention that the users want to express in their comments. Table 8 shows three examples of errors in obtaining the sentiment, which implies that more research is needed in this field.

Table 8. Examples of errors in sentiment analysis

Tweet	Positive	Negative	S	Polarity
The news that Inditex will boost research in **sustainability** and textile **recycling** makes me laugh	2	0	2	Positive
Can we talk about **sustainability** by selling 1.6 billion items a year?	1	0	1	Positive
The economic impact of Inditex in the city of Coruña is **brutal,** directly or indirectly	0	1	-1	Negative

e) In this work, other agents who assess firm's CSR are ignored. Therefore, it could be interesting to compare the informative relevance of Twitter with that of other

  media that produce a CSR opinion, such as news outlets, NGO's, or environmental, social and corporate governance (ESG) rating agencies.

f) The lack of unanimity in the criteria for drafting sustainability reports makes analysis and automatic comparison of the results of other companies difficult, since the GRI Standards guidelines (2016) indicate the obligation to collect the GRI content index in the report but do not specify its structure.

g) It should be taking into account the limitation that maybe some stakeholders do not use Twitter, and therefore, their opinions have not been considered.

h) Finally, findings have been obtained applying the CSR-IRIS method to a textile market sector company. Application to other companies can give different results.

REFERENCES

Acuti, D., Grazzini, L., Mazzoli, V., Aiello, G. (2019). Stakeholder engagement in green place branding: A focus on user-generated content. Corporate Social Responsibility and Environmental Management, 26(2) pp. 492-501

Adams, M.N. (2010). Perspectives on Data Mining. International Journal of Market Research 52(1):11–19.

Amor-Esteban, V., Galindo-Villardon, M.P., Garcia-Sanchez, I.M. (2020). Bias in composite indexes of CSR practice: An analysis of CUR matrix decomposition. Corporate Social Responsibility and Environmental Management, In press.

Ancos, H. (2014). Big Data. Hacia una RSC de datos. RSC: Para superar la retórica. Economistas sin Fronteras. Nº14

Allianz Seguros (2010). Risk Management and Compliance, two sides of the same coin?. XXI Spanish Congress of Risk and Insurance Management, Madrid: AGERS, 26 May.

Araujo, T., Kollat, J. (2018) 'Communicating effectively about CSR on Twitter: The power of engaging strategies and storytelling elements', Internet Research, Vol. 28 Issue: 2, pp.419-431

Arribas, V., Díaz, T., Josa, M. E., Sanmiguel, P., Sádaba, T. (2013). Hoss Intropia: a Spanish brand that is born with sustainability as its main value proposition. 5th Global Conference Fashion Exploring Critical Issues, Oxford, 9-12 September 2013.

Bali, R., Sarkar, D., Sharma, T. (2017). Learning Social Media Analytics with R. Packt Publishing Ltd.,pp. 47-87.

Bebbington, J., Larrinaga, C., & Moneva, J. M. (2008). Corporate social reporting and reputation risk management. Accounting, Auditing & Accountability Journal, 21(3), 337–361. https://doi.org/10.1108/09513570810863932

Bernal, N. (2012). Open data, better than Big Data to generate profits for companies. TICBeat. http://innovacion.ticbeat.com/opendata-mejor-big-data

Blasco Vázquez, J.L. y ElortzArtega, I. (2009). Informes de responsabilidad corporativa: la historia de un éxito. Estrategia Financiera, 257 pp 26-29.

Capelle-Blancard, G., Petit, A. (2019). Every Little Helps? ESG News and Stock Market Reaction. Journal of Business Ethics 157, 543–565

Capriotti, P. (2011). Communicating Corporate Social Responsibility through the Internet and Social Media. The Handbook of Communication and Corporate Social Responsibility, NJ: Wiley

Castelló, I., Morsing, M., Schultz, F. (2013). Communicative Dynamics and the Polyphony of Corporate Social Responsibility in the Network Society. Journal of Business Ethics (118)4 683-694

Colleoni, E. (2013). CSR communication strategies for organizational legitimacy in social media. Corporate Communications: An International Journal, Vol. 18 No. 2, pp. 228-248. https://doi.org/10.1108/13563281311319508

Chalmeta, R., Viinikka, H. (2017). Corporate philanthropy communication on donor websites. Journal of Information, Communication and Ethics in Society 15 (1):53-73 (2017)

Cho, C. H., Laine, M., Roberts, R. W., & Rodrigue, M. (2018). The frontstage and backstage of corporate sustainability reporting: Evidence from the Arctic National Wildlife Refuge Bill. Journal of Business Ethics, 152(3), 865–886. https://doi.org/10.1007/s10551-016-3375-4

The European Union Commission (2011). Corporate social responsibility (CSR) in the EU. Retrieved at http://ec.europa.eu/social/main.jsp?catId=331 .

Crawford et al. (2013). Communities and Ethical Resilience: A framework for Action. Big Data, Communities and Ethical Resilience: White paper.

Deegan, C. (2002). The legitimising effect of social and environmental disclosures - A theoretical foundation. Accounting, Auditing and Accountability Journal, 15 (3) (2002), pp. 282-311.

De Queiroz, G. et al. (2019). tidytext: Text Mining using "dplyr", "ggplot2", and OtherTidy Tools. R package version 0.2.2 https://CRAN.R-project.org/package=tidytext

Eding, E., & Scholtens, B. (2017). Corporate social responsibility and shareholder proposals. Corporate Social Responsibility and Environmental Management, 24(6), 648–660. https://doi.org/10.1002/csr.1434

Elving, W. (2013). Scepticism and corporate social responsibility communications: the influence of fit and reputation. Journal of Marketing Communications, 19 (4), 277-292.

Etter, M. (2013). 'Reasons for low levels of interactivity (Non-) interactive CSR communication in twitter'. Public Relations Review 39 pp.606–608

Etter, M. (2014). 'Broadcasting, reacting, engaging-three strategies for CSR communication in Twitter'. Journal of Communication Management, 18(4), 322-342.

Ettinger, A., Grabner-Kräuter, S., Terlutter, R. (2018). 'Online CSR communication in the hotel industry: Evidence from small hotels'. International Journal of Hospitality Management. Vol 68, pp. 94-104

Farache F., Tetchner I., Kollat J. (2018). 'CSR Communications on Twitter: An Exploration into Stakeholder Reactions'. In: Grigore G., Stancu A., McQueen D. (eds) Corporate Responsibility and Digital Communities. Palgrave Studies in Governance, Leadership and Responsibility. Palgrave Macmillan, Cham

Feinerer, I. and Hornik, H. (2017). tm: Text Mining Package. R package version 0.7-6. https://CRAN.R-project.org/package=tm

Fernández, R. (2005). Administración de la responsabilidad social corporativa. Madrid - Thomson, Colección Negocios.

Freeman, R. Edward (1984). Strategic Management: A Stakeholder Approach. Boston: Pitman.

Friedman, M. (1970). The Social Responsibility of Business is to Increase its Profits. The New York Times Magazine, 13 September.

Garcia-De los Salmones, M., Pérez, A. (2018) Reactions to the corporate social responsibility advertising of financial entities. Vol 4 DOI: 10.3232/UBR.2018.V15.N4.01

Garcia-Torea, N., Fernandez-Feijoo, B., De La Cuesta, M. (2018). CSR reporting communication: Defective reporting models or misapplication?. Corporate Social Responsibility and Environmental Management, 1(17) DOI: 10.1002/csr.1858

Gentry, J. (2015). twitteR: R Based Twitter Client. R package version 1.1.9. https://CRAN.R-project.org/package=twitteR

Global Reporting Initiative (GRI) (2016). GRI G4 - Sustainability reporting guidelines. Obtenido de: https://www2.globalreporting.org/standards/g4/

Gong, V., X., Daamen, W., Bozzon,A., Hoogendoorn,S. (2019). Estimate Sentiment of Crowds from Social Media during City Events. Transportation Research Record Journal of the Transportation Research Board, June 2019.

Gray, R. (2006). Does sustainability reporting improve corporate behaviour?: Wrong question? Right time? Accounting and Business Research, 36(1), 65–88Hu, M., & Liu, B. (2004). Mining opinion features in customer reviews. National Conference on Artificial Intelligence.

Grover, P., Kar, A.K., Ilavarasan, P. V. (2019). Impact of corporate social responsibility on reputation-Insights from tweets on sustainable development goals by CEOs. International Journal of Information Management, 48 pp. 39-52

Hu, M. & Liu, B. (2004). Mining and summarizing customer reviews, in Proceedings of the tenth ACM SIGKDD international conference on Knowledge discovery and data mining, pp. 168–177, ACM.

Huang, Y.F., Do, M.H., Kumar, V. (2019). Consumers' perception on corporate social responsibility: Evidence from Vietnam. Corporate Social Responsibility and Environmental Management, 26(6) pp.1272-1284

Inditex. Memoria Anual de Inditex (2017). https://static.Inditex.com/annual_report_2017/

Inditex. Memoria Anual de Inditex (2018). http://static.Inditex.com/annual_report_2018/es/?v1=967754

Kemna, T. (2013) Dialogue, Twitter and New Technology-based firms: the communication practice on a social Medium. Retrieved from http://su-beta.diva-portal.org/smash/get/diva2:629600/FULLTEXT01.pdf;jsessionid=cf4e559c4a22712b67 093e831943

Knebel, S., & Seele, P. (2015). Quo vadis GRI? A (critical) assessment of GRI 3.1 A+ non-financial reports and implications for credibility and standardization. Corporate Communications: An International Journal, 20(2), 196–212. https://doi.org/10.1108/CCIJ-11-2013-0101

Kotler, P., Lee, N. (2004). Corporate social responsibility: doing the most good for your company and your cause. Hoboken, N.J.: Wiley.

KPMG (2015). Currents of Change. The KPMG survey of corporate responsibility reporting 2015. Amsterdam: KPMG.

Kumar Ravindran,S., Garg,V. (2015). Mastering Social Media Mining with R. Packt Publishing Ltd., pp.21-54.

KPMG, (2017). Executive Summary: The KPMG Survey of Corporate Responsibility Reporting 2017. Obtained from: https://home.kpmg/xx/en/home/insights/2017/10/executive-summary-the-kpmg-survey-of-corporate-responsibility-reporting-2017.html , 12 October pp16.

Liu, B. (2015). Sentiment analysis: Mining opinions, sentiments, and emotions. Ed Cambriedge University Press. New York (USA).

Lovejoy, K. and G.D. Saxton (2012). 'Information, Community, and Action: How Nonprofit Organizations Use Social Media'. Journal of Computed-Mediated Communication 17 (2012) 337-353.

Luoma-aho, V. (2015). Understanding Stakeholder Engagement: Faith-holders, Hateholders & Fakeholders 2(1)

Maignan, I., (2001). 'Consumers' perceptions of Corporate Social Responsibilities: A Cross-Cultural Comparison'. Journal of Business Ethics, 30(1), 57-72.

Merco Talent (2020). https://www.merco.info/es/que-es-merco#!

Michelon, G., Pilonato, S., & Ricceri, F. (2015). CSR reporting practices and the quality of disclosure: An empirical analysis. Critical Perspectives on Accounting, 33, 59–78. https://doi.org/10.1016/j.cpa.2014.10.003

Michelon, G., & Rodrigue, M. (2015). Demand for CSR: Insights from shareholder proposals. Social and Environmental Accountability Journal, 35(3), 157–175. https://doi.org/10.1080/0969160X.2015.1094396

Molina-González, Martínez-Cámara, Martín-Valdivia y Perea-Ortega (2013). iSOL Semantic orientation for polarity classification in Spanish reviews. Expert Systems with Applications, 40(18): 7250-7257. Obtained from: http://sinai.ujaen.es/wp-content/uploads/2013/05/isol.tar.gz

Morsing, M. and Schultz, M. (2006). "Corporate social responsibility communication: Stakeholder information, response and involvement strategies". Business Ethics 15 (4):323–33

Pomering, A., Johnson, L. and Nobl, G. (2013). "Advertising corporate social responsibility. Results from an experimental manipulation of key message variables", Corporate Communications: An International Journal, 18 (2), pp. 249-263

Ravindran SK, Garg V.(2015). Mastering social media mining with R. Packt Publishing Ltd. pp. 21-54

Rinker, T. (2019). sentimentr: Calculate Text Polarity Sentiment.R package version 2.7.1 https://CRAN.R-project.org/package=sentimentr

Ruf, B.M., K. Muralidhar, R.M. Brown, J.J. Janney and Paul K., (2001), 'An Empirical Investigation of the Relationship between Change in Corporate Social Performance and Financial Performance: A Stakeholder Theory Perspective', Journal of Business Ethics, 32(2), 143-156.

Rundong Du. et al. (2015). Toward Social Media Opinion Mining for Sustainability Research. Georgia Institute of Technology. Atlanta, Georgia 30332. Computational Sustainability: Papers from the 2015 AAAI Workshop.

Slater, K. (2000). Environmental Impact of Textiles: Production, Processes and Protection. Cambridge, UK: Woodhead Publishing Limited/The Textile Institute.

Smith, N.C., (2003). Corporate Social Responsibility: Whether or How?. California Management Review, 45(4), 52-76.

Song, B., Jing, W. (2019). Online corporate social responsibility communication strategies and stakeholder engagements: A comparison of controversial versus noncontroversial industries. Corporate Social Responsibility and Environmental Management, 27(2) 881-896

Srivastava, S.K. Green Suplly Chain Management: A State-of-the-Art Literature Review. International Journal of Management Reviews, v.9 No 1, pp. 53-80, 2007.

Szumniak-Samolej, J. (2019). Social Media for Corporate Social Responsibility Strategy Creation and Communication in Poland: Strategies, Opportunities and Challenges. In book: Corporate Social Responsibility in Poland. Publisher: Springer, Cham

Wang, Z., Hsieh, T.S., Sarkis J. (2018). CSR Performance and the Readability of CSR Reports: Too Good to be True?. Corporate Social Responsibility and Environmental Management, 25(1) 66-79

Wickham, H. (2016). plyr: Tools por Splitting, Applying and Combining Data. R package version 1.8.4.https://CRAN.R-project.org/package=plyr

Zizka, L. (2017). The (mis) use of social media to communicate CSR in hospitality: Increasing stakeholders'(dis) engagement through social media. Journal of Hospitality and Tourism Technology, 8 (1), pp. 73-86