A Fuzzy Logic Model for the Analysis of Social Corporate Responsibility

A. C. Golzio[1], M. Puerta-Díaz[1,*] and D. Martínez-Ávila[2]

[1]São Paulo State University, Brazil
[2]University Carlos III of Madrid, Spain

Abstract

INTRODUCTION: Critical public opinion, based on information that is made available to the public through different systems, has led companies that operate in the environment to continually improve their social, environmental, and ethical performance.

OBJECTIVES: This paper aims to propose a fuzzy-logic-based model for the analysis of social corporate responsibility in cases of environmental accidents.

METHODS: Our study employs techniques derived from social network analysis. The data was collected from the online database of The New York Times for the timespan from March 24, 1989, to September 1, 2017.

RESULTS: The results show that the proposed model can be replicated, after some adjustments.

CONCLUSION: We conclude that, despite the complexity of an analysis of this kind in which the model is applied considering isolated words in the text and not the semantic aspects, the proposed model based on fuzzy logic is adequate for the analysis of social corporate responsibility.

* Corresponding author. Email: Mireblue18@gmail.com

1. Introduction

One of the most important aspects for companies dealing with environmental disasters is the way they are informed to the public. Environmental accidents harm society and often leave irreparable marks on the planet, so the public has the right to know and to participate in the process of demanding accountability. In this context, the agents who caused the accident (directly or indirectly) may be asked through different sources to be accountable for taking measures to mitigate the damage. These sources could be legal, that is, responsible agents could be legally penalized and forced to take corrective measures; or, due to some type of social pressure, agents could be led to repair the damage done and producing benefits for society, for instance, favouring the quality of life of the affected population and the preservation of nature. The latter case is characteristic of what is known as corporate social responsibility (CSR). The concept of CSR is linked to the notion of obligation to constituent groups in society [7].

In the last decade, companies worldwide have increasingly faced demands related to transparency problems and the revelation of accounting or legal scandals, reputation trembles, and corporate collapses, such as the famous cases of Enron, WorldCom and Tyco in the United States. Critical public opinion has led companies that operate in the environment to continually improve their social, environmental and ethical performance. An example is the Green Chemistry that emerge around 1990, in order to reduce the pollution in the chemistry industries [9]. In addition, increased media interest and public pressure have also reinforced the stakeholder's awareness [13]. In this scenario, corporate social responsibility went beyond being seen as a potential competitive advantage for companies and it has been started to be considered as a real strategic need [12]. Those companies that do not embrace CSR, especially operating in environment and delicate fields, are at risk of losses, law suits, and falling into disgrace. In the business world, CSR has moved from ideology to reality [8].

The growing interest in CSR in the business world has been accompanied by an increasing interest in the academic field and the publication of articles on CSR in journals on management and other areas, mainly focusing on ethical and environmental issues [14]. The second half of the 20th century witnessed a long debate on corporate social responsibility, so the renewed interest on the topic in recent decades has resulted in the proposal of new concepts and interdisciplinary relations to its study [11, 4, 3, and 10]. In practice, CSR theories have four dimensions: profits, political performance, social demands, and ethical values [4]. In relation to the latter dimension, theories are based on the ethical responsibilities of companies to society.

CSR aims to raise awareness on ethical and environmental issues within and outside the organizations, alleging that organizations adopt socially responsible and sustainable attitudes that aim at protecting the environment and the ethical regulation of the conduct of their members. In other words, companies are aware of the risks of their activities for the environment and are adopting a proactive attitude towards minimizing the dangers and compensating society beforehand. Awareness and attribution of the reasons for CSR are essential aspects that need to be evaluated, as they represent the main drivers of the type of assessment that society will make of the company and the crisis situation [6]. Ultimately, the key for the companies and the impression they leave on society is an information game that depends on the transmission and analysis of information.

Although there has been a proportional growth in the interest and number of scientific publications that analyze CSR, the critical approaches based on methods of data mining and fuzzy logic to identify the assignation and recognition of corporate social responsibility are insufficient. Fuzzy logic is the logic underlying approximate, rather than exact, reasoning modes [16] and, therefore, it can be considered adequate for the analyses that involve subjective properties or inaccurate attributes. This aspect is an innovation for the research that is traditionally conducted on the information studies that relate to CSR.

The present research aims to fill this gap in the traditional literature as it examines the capacity of corporate response when dealing with environmental accidents based on the processing of massive data on public opinion. For the case of environmental disasters, we aim to study the existence of initiatives in those companies that caused major environmental disasters to adopt a corporate social responsibility over time.

Finally, we also aim to propose a model that draws on the analyses of an ethical vocabulary and a Fuzzy Rules-Based System to predict possible ethical and social responsibility responses offered by the companies that have caused environmental disasters reported in the media. Our model is applied to the environmental accident that happened on March 24, 1989, in which the oil tanker Exxon Valdez, one of the ships belonging to the Exxon Mobile Company, ran aground in the north of Prince William Sound, Alaska, spilling more than 42 million liters of crude oil (roughly 10.8 million gallons) and contaminating 1,990 kilometers of coastline. The relevance of this case led environmental organizations to develop what they called "Valdez Principles" for corporations to sign [1].

We believe that the conceptual model proposed in the present paper can be used as a basis for understanding the dynamics of other environmental accidents, as well as for identifying ethical actions for assigning or recognizing

corporate social responsibility initiatives in companies that are based on the pressure by public opinion.

2. Literature Review

The history of fuzzy logic dates back to the past century. In the 1960s, Lotfi Zadeh, suggested an alternative set theory that was less rigid than usual and called it fuzzy set theory [14]. Over the years, there has been an exponential growth of scientific interest in applying its fuzzy set theory in science too, as represented in Figure 1.
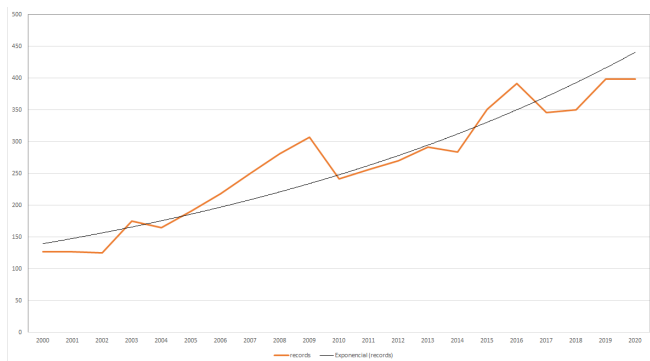

Figure 1. Trend of publications on fuzzy set theory and Fuzzy Rules Based System in WoS (2000-2020)

This figure presents the behavior over time of the 6,539 publications retrieved from the WoS database on May 30, 2021, searching for the terms "fuzzy set theory" OR "Fuzzy Rules Based System". The figure shows a continuous increase in the scientific production indexed by the WoS database that can be extrapolated to further dates. The figure was limited to the publications in the present millennium, for the sake of clarity. The oldest record retrieved, that is not shown in the figure, was published by Kickert and Koppelaar in 1976. In that study, the authors applied the fuzzy set-theory to the recognition of "syntactic patterns" of capital letters. Since that study, its application has been adopted in several scientific areas, gaining a marked interest in the field of artificial intelligence and computer science. The years 2019 and 2020 appear as the most productive ones on the topic, so it would be reasonable to predict an increase trend in the number of publications for the forthcoming years.

In the fuzzy set theory, proposed by Zadeh, the change from pertinence to non-pertinence is gradual and not abrupt. Thus, in the fuzzy sets, for each element of the discourse universe, we have a corresponding degree of relevance in the fuzzy set that is given by a real number between 0 and 1. With fuzzy sets, the possibility of interpreting non-quantitative phenomena is vague, also increasing the need to seek mechanisms for inferences from these data. One of these mechanisms was the Fuzzy Rules Based Systems (FRBSs) [2]. It should be noted that, due to its

multidisciplinary nature, the Fuzzy Rules Based System is also known by several other names, such as "fuzzy rules-based inference system," "expert fuzzy system," "fuzzy model," and "controller logical fuzzy."

A FRBS is a system that uses fuzzy logic to produce outputs from fuzzy inputs. In these systems, the linguistic variables play a fundamental role, since the linguistic terms, translated by fuzzy sets, are subject to a base of rules to obtain a fuzzy inference relation in which an output is produced for each input of the system. In this context, linguistic variables are variables whose values are fuzzy set names. For example, the temperature of a given process can be a linguistic variable assuming low, medium, high, etc. These values can be described by fuzzy sets.

In a nutshell, a FRBS consists of four components connected as shown in Figure 2 below:

i. Fuzzification module
ii. Rules base module (rules)
iii. Inference module
iv. Defuzzification module
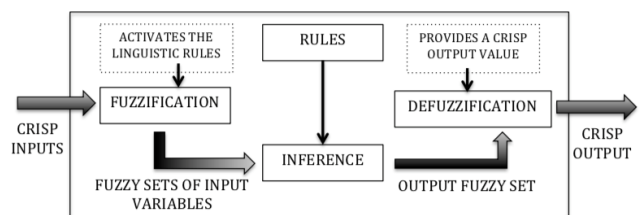
Figure 2. Fuzzy Rules Based System

In the fuzzification module, the system inputs must be modeled by fuzzy sets. In this module, the role of the proponent of the model is fundamental, as they are the ones who formulate the pertinent functions for each fuzzy set involved in the process.

In this step, even if the input has a precise value (crisp input), that is, it is not a fuzzy type value, it is fuzzified through its characteristic function. In the rules base module (rules) the fuzzy parameters are sent to an inference machine (inference) that processes a series of rules of the IF-THEN type, that is, propositions involving terms of linguistic variables. In the final stage of rules processing, the fuzzy value, which is obtained as an answer, goes through a process of defuzzification of the fuzzy values obtained as an output of the process, and, thus, a precise output is obtained (crisp output).

In a FRBS we can use several methods of inference to build the rules that make up the basis of inference. The most well-known are Mamdani and Takagi-Sugeno-Kang (TSK). Here we use the Mamdani method which is widely used due to its simplicity and also it is the one that best applies to our

problem. This method was introduced by Mamdani and Assilian in 1975. The rules of a Mamdani model are as follows:

If x is A and y is B then z is C

If x is A or y is B then z is C

If x is A then z is C

There are many methods of defuzzification that can be adopted, some of the most common are: centroid or (center of mass), mean of maximum, smallest of maximum, and highest of maximum. The method that proved to be the most suitable to be adopted in the FRBS built here is the largest of the maxima (LOM), which basically calculates the output of the FRBS by choosing the value that corresponds to the highest value of relevance among the maximums.

2. Materials and methods

The data to test our model was collected from the news database of The New York Times (NYT), a prestigious and popular medium for the information and creation of opinion in society. This source is also widely used and accepted in the scientific literature. For instance, in [5], the authors explain their preference of this data source over other newspapers based on, among other things, the free access to its abstracts, something that is also important for the purposes of our research. Additionally, we believe in the relevance of using a database of news collected from online newspapers to test our hypothesis because, although they do not directly reflect public opinion, they indicate the way the accident affected the social environment.

The search and retrieval of data used the Article Search API to obtain news that is available on the developer platform (https://developer.nytimes.com/docs/articlesearch-product/1/routes/articlesearch.json/get). The use of the API's function get /articlesearch.json allowed us to filter the search by date of publication of the news and keywords. With the application of filters, it was possible to reduce the existing computational cost for handling large amounts of data and thus to make subsequent processes affordable using current hardware and software specifications. We entered the following parameters:

- begin_date: 19890324
- end_date: 20170901
- facet_fields: null
- Query: exxon
- Sort: null

It should be noted that the parameter "sort" allows sorting by relevance criteria or by the most recent publication data or by the oldest one that is published on the NYT platform. We chose the value null for this parameter as we were more

interested in the analysis of the data than in their presentation at this stage. The retrieved news included the word "exxon" in the title field, keywords or lead paragraph. The choice of this term allowed us to conduct a more comprehensive search of the news about the company responsible for the accident and thus guaranteed the retrieval of the entire universe of data for the time span under analysis. The temporary window was established to study the dynamics of the public opinion about the event, since the collection allowed us to retrieve a representative data set from the date of the accident until the year 2017 (including a total of 28 years of media coverage).

Using these search criteria, a total of 2,001 news items were retrieved and incorporated into the file named 'DB Exxon 1989 – 2017.txt' (data set available at: https://doi.org/10.6084/m9.figshare.13555661). Other criteria for the characteristics of the data records of the news that were retrieved are specified in the 'Data collection, filtering and pre-processing' section.

We applied fuzzy system techniques to the design of a fuzzy rules-based system. Fuzzy logic can be applied to make rational decisions in an environment of inaccurate information [17] that is characteristic of the object of our study.

Our research begins with the presentation of a conceptual model that includes in one of its phases the application of a fuzzy model in massive databases. On the other hand, the conceptual model based on fuzzy logic uses inference rules that allow the adaptation to various contexts, especially those that involve the processing of some degree of uncertainty present in the data analysis.

An important phase of our work is based on the identification of an ethical vocabulary for the assignation and recognition of corporate social responsibility. We categorized the vocabulary in three main groups:

a. Concepts of assignation of responsibility
b. Concepts of recognition of responsibility
c. Neutral concepts (whose role is defined by their proximity to the terms of assignation or recognition, that is, they can perform both functions depending on the context).

3. Development and application of the model

Figure 3 represents the workflow of the model that is proposed here. The workflow is divided into four stages. The viability of each stage was meticulously evaluated considering that the ability to abstract and understand the dynamics of the public opinion in relation to companies that caused major environmental disasters was an important objective in our study.
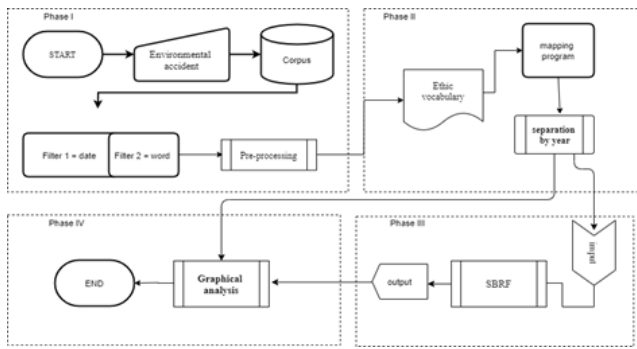
Figure 3. Workflow of the model for the analysis of environmental accidents

Each stage of the model is explained in more detail and discussed in the next section for the case of the Exxon Valdez environmental accident.

3.1 Data collection, filtering and pre-processing

Collection

In this stage, the news database in which the environmental accident is represented is selected and the data is retrieved.

Environmental accident to be analyzed: Oil spill by the tanker Exxon Valdez.

News database: we selected the New York Times Article Archive and used the *Article Search API* to retrieve items by keyword in a specific period. The search uses filters to refine result*s* and to specify the fields and values for our query. We limited the search to articles whose source is The New York Times with the following query: fq=source:("The New York Times").

Filtering: The model uses two types of filters, one by time period and another one by keyword. The purpose of applying the filters is to limit the volume of data that is retrieved and to locate the specific news that are related to the environmental accident. In our case:

- By period: we retrieved news from March 24, 1989, to September 1, 2017.
- Keyword: items that include the term ("exxon").

In our case, we retrieved 2,001 items from the NYT online database. Then, we assigned a consecutive identification number (ID) that went from "0" to "2000" to each record. These collected data were very important for the mapping to the ethical vocabulary and the preparation of the fuzzy system (section 3.2).

Pre-processing

After the metadata filtering, we applied some Natural Language Processing (NLP) methods. For the scripts to be applied in the processing, we considered, adapted, and optimized the methods published on the website https://www.machinelearningplus.com/nlp/topic-modeling-gensim-python/ to the context of our research data and interests of our study.

In the metadata filtering stage, we removed some irrelevant values of the metadata article schema for our specific research purposes. The following information of the description of the data was not necessary for the application of our model: 'web_url', 'print_page', 'source', 'multimedia', 'headline', 'keywords', 'document_type', 'news_desk', 'type_of_materia', '_id', 'word_count', and 'uri'. The fields that we considered are the following ones: "headline", "lead_paragraph", "pub_date", "snippet", and "abstract". It is important to note that the removed data could be included in other cases depending of the objectives of the study (as well as the fields we included removed for similar reasons).

NLP methods: this sub-stage included techniques of tokenization, removal of stopwords, lemmatization and n_grams and bigram of the news retrieved and stored in the file 'DB Exxon 1989 - 2017.txt' from the dataset.

The description of the NLP at the morphological level is as follows:

i. Tokenization: to transform a set of records into a matrix in which each document is transformed into a list of words and each generated word list occupies a line in the matrix, using the Gensim library. The following symbols were removed: punctuation marks, isolated letters (a, e, o), numbers and special symbols present in the analyzed corpus.
ii. Removal of stopwords: we removed the stopwords or common words in the language that do not seem to have a relevant meaning, such as "I", "that", "for", etc. using the NLTK library for the English.
iii. Lemmatization: this method was applied as part of the NLP at the morphological level to transform the word into its "root", the meaningful base form.
iv. n_grams: we apply a bigram to identify those terms that frequently appear together, such as 'social network', etc.

3.2 Mapping the ethical vocabulary

In the second stage, we conducted the mapping of the ethical vocabulary using the filtered and pre-processed database. Figure 4 shows the deployment of the terms of the ethical vocabulary following the categorization described in the methodology:

i. Group 1 - concepts of assignation of responsibility (23 concepts),

ii. Group 2 - concepts of recognition of responsibility (9 concepts)

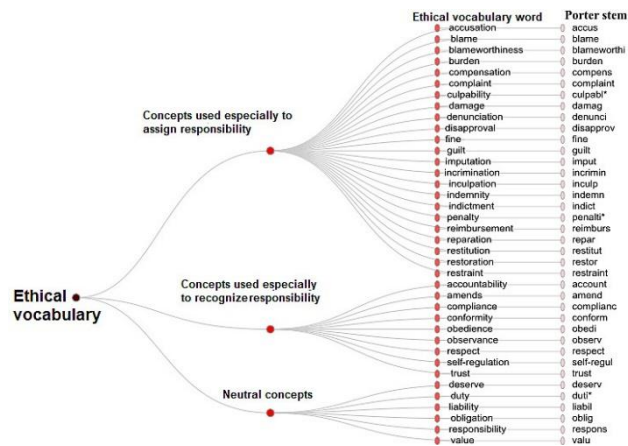iii. Group 3 - neutral concepts (6 concepts).

Figure 4. Ethical vocabulary terms (also available in full resolution at https://doi.org/10.6084/m9.figshare.13582898)

This stage was developed in four steps:

i. First step: Identification of the items (news) in the filtered and pre-processed database that include the terms from the ethical vocabulary.

ii. Second step: Selection and ordering of just the items (news) IDs that include the terms from the ethical vocabulary.

iii. Third step: Association of the selected items (news) that include the terms from the ethical vocabulary to their respective terms from the ethical vocabulary and their years of publication, resulting in a table that includes the IDs of the items (news), the year in which they were published and the terms from the ethical vocabulary that are present in the news.

iv. Fourth step: Division of the previous table by year, indicating the frequency of terms from the ethical vocabulary for each of the three groups of concepts (assignation, recognition, and neutral).

Application of the ethical vocabulary mapping

In the first and second steps we mapped and located those items-news (from the clean and pre-processed database) that include terms from the ethical vocabulary. This task was divided into three parts:

- We identified the *Porter stem* of each term in the ethical vocabulary using the application available at: http://morphadorner.northwestern.edu/morphadorner/lemmatizer/example/.
- We searched the *Porter stem* of the terms in the clean and pre-processed database using a searcher developed

in Python, using the following script to search for the term root

```
import re
x = input('Type the root word by leaving a space before the word and using .+ at the end: ')
with open('/Users/autor/Desktop/Exxon.txt') as f:
for l_num, l in enumerate(f, 1):
while re.search(x, l):
print(l_num - 1)
break
else: print('The search is over')
```

In those cases in which the *Porter stem* changed some letters from the original term, we searched both versions. For instance, for "duty" the *Porter stem* returned "duti", so we searched for "duty" and "duti". Our search program used a *wildcard* to search for variations of the term in the ethical vocabulary. A *wildcard* is a symbol used to replace or represent one or more characters in the term. For instance, the asterisk symbol "*" is a *wildcard* that is commonly used at the end of a root word, as a suffix, to search for variable endings of a root word. In our case we used the *Porter stem* of the word as the root.

- Our search program, developed in Python, returned a list with the IDs of the items (news) that included the term. We organized these data relating the terms of the ethical vocabulary searched to the ID of the news in which they appear. It is important to note that our search does not take into account the frequency of the term in the news, we only signal if the term appears in the news or not.

In the third step we related the data from the previous step to the year in which the news were published. We created the relationship between the ID of the item (news), the year in which it was published, and the terms from the ethical vocabulary that are present in the news (where "1" indicates the presence of the term in the news and "0" its absence).

In the fourth step, we divided this information by year (from 1989 to 2017) and calculated the frequency of terms from the ethical vocabulary for each group of concepts (assignation, recognition, and neutral). The table available at https://doi.org/10.6084/m9.figshare.13567838 illustrates how the calculation was made for all years. For instance, for the year 2000 (Figure 5), we identified 12 news items that include terms from the ethical vocabulary. In these 12 news items, we have 276 possible occurrences of terms from the ethical vocabulary for the group of responsibility assignation, 108 possible occurrences of terms for the group of responsibility recognition, and 72 possible occurrences of terms from the group of neutral concepts. For each group of concepts, we counted the actual occurrences of the terms, resulting in 4 actual occurrences out of 276 possible

occurrences of assignation concepts, that is 1%; 7 actual occurrences out of 108 possible occurrences of recognition concepts, that is 6%; and 4 actual occurrences out of 72 possible occurrences of neutral concepts, that is also 6%. This was calculated for all the years analyzed with our model.

Figure 5. Frequency of terms from the ethical vocabulary present in the news published in the year 2000 (available in full resolution at https://doi.org/10.6084/m9.figshare.13567838)

The data from the four steps of the mapping to the ethical vocabulary were organized in tables (Figures 6 and 7) and used in the next stages of application of the model. The data from the mapping of the ethical vocabulary were used to construct the variables of the fuzzy system, described in the next section.

Figure 6. Data mapping of the ethical vocabulary (available in full resolution at https://doi.org/10.6084/m9.figshare.13567838)

Year	v – News with ethical vocabulary concepts (Mean)	v – News with ethical vocabulary concepts (%)	a – Assignation concepts (Mean)	a – Assignation concepts (%)	r – Recognition concepts (Mean)	r – Recognition concepts (%)	n – Neutral concepts (Mean)	n – Neutral concepts (%)
1989	0.183036	18%	0.032874	3%	0.008130	1%	0.060976	6%
1990	0.262500	26%	0.034161	3%	0.029101	3%	0.059524	6%
1991	0.410256	41%	0.048007	5%	0.032407	3%	0.048611	5%
1992	0.309735	31%	0.058385	6%	0.015873	2%	0.052381	5%
1993	0.243902	24%	0.030435	3%	0.011111	1%	0.116667	12%
1994	0.418182	42%	0.051040	5%	0.014493	1%	0.043478	4%
1995	0.342105	34%	0.036789	4%	0.042235	4%	0.076923	8%
1996	0.296296	30%	0.043478	4%	0.069444	7%	0.000000	0%
1997	0.333333	33%	0.027668	3%	0.060606	6%	0.045455	5%
1998	0.236842	24%	0.021739	2%	0.055556	6%	0.064815	6%
1999	0.277777	28%	0.014493	1%	0.044444	4%	0.088889	9%
2000	0.285714	29%	0.014493	1%	0.064815	6%	0.055556	6%
2001	0.245902	25%	0.034783	3%	0.007407	1%	0.055556	6%
2002	0.292683	29%	0.028986	3%	0.027778	3%	0.083333	8%
2003	0.243902	24%	0.021739	2%	0.033333	3%	0.066667	7%
2004	0.333333	33%	0.014493	1%	0.064815	6%	0.041667	4%
2005	0.316667	32%	0.038902	4%	0.040936	4%	0.096491	10%
2006	0.354545	35%	0.040134	4%	0.028490	3%	0.064103	6%
2007	0.236842	24%	0.033816	3%	0.043210	4%	0.064815	6%
2008	0.121622	12%	0.043478	4%	0.012346	1%	0.037037	4%
2009	0.172414	17%	0.034783	3%	0.011111	1%	0.066667	7%
2010	0.212766	21%	0.013043	1%	0.016667	2%	0.116667	12%
2011	0.179104	18%	0.014493	1%	0.018519	2%	0.111111	11%
2012	0.191489	19%	0.048309	5%	0.000000	0%	0.055556	6%
2013	0.179487	18%	0.037267	4%	0.015873	2%	0.023810	2%
2014	0.128205	13%	0.008696	1%	0.111111	11%	0.066667	7%
2015	0.373194	37%	0.036522	4%	0.026667	3%	0.026667	3%
2016	0.166667	17%	0.006689	1%	0.034188	3%	0.089744	9%
2017	0.212121	21%	0.018634	2%	0.047619	5%	0.023810	2%
Maximum		42%		6%		11%		12%

Figure 7. Annual summary of data on mapping ethical vocabulary (available in full resolution at https://doi.org/10.6084/m9.figshare.13567871)

3.3 Application of the FRBS

Our FRBS, named "Mars" (after *Modelo de análise de responsabilidade social*, Model for social responsibility analysis in Portuguese), was developed using Mamdani's inference method and implemented using the MATLAB software and the "Fuzzy Logic Designer" toolbox (see Figure 8).

Figure 8. FRBS Mars created using MATLAB

Fuzzification and defuzzification

The FRBS has three input variables, namely "assignation", denoted by "a" (which refers to the occurrences of concepts of assignation of responsibility), "recognition", denoted by "r" (which refers to occurrences of concepts of recognition of responsibility), and "neutral", denoted by "n" (which refers to occurrences of neutral concepts). These data were obtained in the previous stage by mapping the ethical vocabulary and were used here in the construction of the system variables. Table 1 shows the definition of membership functions and characteristics of input variables for a universe of discourse: delimiters [0; 12].

Table 1. Definition of the membership functions of the input variables: neutral (n), recognition (r) and assignation (a)

Notation	Type	Delimiters	Terms
x1	trapezoidal	[0 0 3 5]	low
x2	trapezoidal	[3 5 7 9]	medium
x3	trapezoidal	[7 9 12 12]	high

Here we present the membership function (trapezoidal type) for each input variable. As they are identical in the three cases - neutral (n), recognition (r) and assignation (a) – they all are represented by Figure 8. The values of the input variables are mapped by their respective membership functions created based on the occurrences in the corpus (in percentage) of the terms from the ethical vocabulary. The input variables were classified as "low", "medium", and "high", according to an equidistant partition of the discourse universe, taking into account the results of the mapping of the ethical vocabulary. The universe of discourse was defined considering the maximum value obtained in the table in Figure 9, in relation to the three variables (recognition, assignation, and neutral).
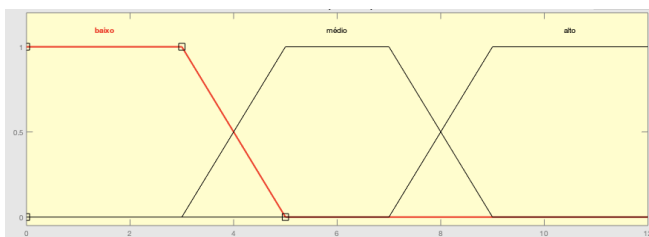
Figure 9. Pertinence functions of the input variables in MATLAB

In the FRBS Mars we also have an output variable named "indicator_SR", denoted by "i". Figure 10 presents the membership functions of the output variable following a triangular type. We chose the triangular-type functions as this type was the one that presented the best result, according to the expectations for the model. Table 2 presents the definition of the membership functions and the characteristics of the output variable: indicator_SR (i).
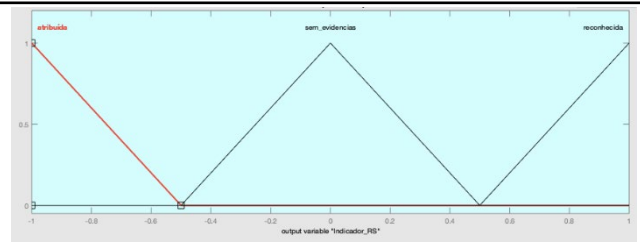

Figure 10. Pertinence functions of the output variables in MATLAB

Table 2. Definition of the membership functions of the output variable

Notation	Type	Delimiters	Terms
z1	triangular	[-1 -1 -0,5]	assigned
z2	triangular	[-0,5 0 0,5]	no_evidence
z3	triangular	[0,5 1 1]	recognized

The universe of discourse of the output variable was defined considering 3 values: -1, 1, 0. These values refer, respectively, to the ideas that "assigned" is a negative value for the company responsible for the environmental disaster, "recognized" is a positive value, and 0 means that a value could not be determined. The output variable was classified as "assigned", "no_evidence" or "recognized", according to the results of the mapping to the ethical vocabulary and from the analysis in the database. The intervals of the triangular functions were constructed considering the idea of "around", that is, we classified as "assigned" the values around -1, as "recognized" the values around 1, and as "no_evidence" the values around 0.

Rules and inference

As mentioned above, the rules were created after a careful analysis of the database and the results of the mapping of the ethical vocabulary, taking into account the following premises:

1. If $(r > a)$ and $(n \leq r)$, then (no_evidence)
2. If $(r > a)$ and $(n > r)$, then (recognized)
3. If $(r < a)$ and $(n \leq r)$, then (no_evidence)
4. If $(r < a)$ and $(n > r)$, then (assigned)
5. If $(r = a)$, then (no_evidence)

We need at least 27 rules to cover all possible combinations of terms for the input variables, that is, the number of rules we created was 3 ^ 3 = 27 rules. The fuzzy rules used in Mars are shown in the table in Figure 11.

Rules								
Rule number		n		r		a		i
1	IF	x1	AND	x1	AND	x1	THEN	z2
2	IF	x1	AND	x1	AND	x2	THEN	z2
3	IF	x1	AND	x1	AND	x3	THEN	z2
4	IF	x1	AND	x2	AND	x1	THEN	z2
5	IF	x1	AND	x2	AND	x2	THEN	z2
6	IF	x1	AND	x2	AND	x3	THEN	z2
7	IF	x1	AND	x3	AND	x1	THEN	z2
8	IF	x1	AND	x3	AND	x2	THEN	z2
9	IF	x1	AND	x3	AND	x3	THEN	z2
10	IF	x2	AND	x1	AND	x1	THEN	z2
11	IF	x2	AND	x1	AND	x2	THEN	z1
12	IF	x2	AND	x1	AND	x3	THEN	z1
13	IF	x2	AND	x2	AND	x1	THEN	z2
14	IF	x2	AND	x2	AND	x2	THEN	z2
15	IF	x2	AND	x2	AND	x3	THEN	z2
16	IF	x2	AND	x3	AND	x1	THEN	z2
17	IF	x2	AND	x3	AND	x2	THEN	z2
18	IF	x2	AND	x3	AND	x3	THEN	z2
19	IF	x3	AND	x1	AND	x1	THEN	z2
20	IF	x3	AND	x1	AND	x2	THEN	z1
21	IF	x3	AND	x1	AND	x3	THEN	z1
22	IF	x3	AND	x2	AND	x1	THEN	z3
23	IF	x3	AND	x2	AND	x2	THEN	z2
24	IF	x3	AND	x2	AND	x3	THEN	z1
25	IF	x3	AND	x3	AND	x1	THEN	z2
26	IF	x3	AND	x3	AND	x2	THEN	z2
27	IF	x3	AND	x3	AND	x3	THEN	z2

Figure 11. FRBS Mars Rules (available in full resolution at https://doi.org/10.6084/m9.figshare.13582988)

The inference was based on the Mamdani method [2]. The Mars code developed in MATLAB is as follows:

[System]
Name='Mars'
Type='mamdani'
Version=2.0
NumInputs=3
NumOutputs=1
NumRules=27
AndMethod='min'
OrMethod='max'
ImpMethod='min'
AggMethod='max'
DefuzzMethod='lom'

[Input1]
Name='Neutral'
Range=[0 12]
NumMFs=3
MF1='low':'trapmf',[0 0 3 5]
MF2='medium':'trapmf',[3 5 7 9]
MF3='high':'trapmf',[7 9 12 12]

[Input2]

Name='Recognition'
Range=[0 12]
NumMFs=3
MF1='low':'trapmf',[0 0 3 5]
MF2='medium':'trapmf',[3 5 7 9]
MF3='high':'trapmf',[7 9 12 12]

[Input3]
Name='Assignation'
Range=[0 12]
NumMFs=3
MF1='low':'trapmf',[0 0 3 5]
MF2='medium':'trapmf',[3 5 7 9]
MF3='high':'trapmf',[7 9 12 12]

[Output1]
Name='Indicator_SR'
Range=[-1 1]
NumMFs=3
MF1='assigned':'trimf',[-1 -1 -0.5]
MF2='no_evidence':'trimf',[-0.5 0 0.5]
MF3='recognized':'trimf',[0.5 1 1]

[Rules]
1 1 1, 2 (1) : 1
1 1 2, 2 (1) : 1
1 1 3, 2 (1) : 1
1 2 1, 2 (1) : 1
1 2 2, 2 (1) : 1
1 2 3, 2 (1) : 1
1 3 1, 2 (1) : 1
1 3 2, 2 (1) : 1
1 3 3, 2 (1) : 1
2 1 1, 2 (1) : 1
2 1 2, 1 (1) : 1
2 1 3, 1 (1) : 1
2 2 1, 2 (1) : 1
2 2 2, 2 (1) : 1
2 2 3, 2 (1) : 1
2 3 1, 2 (1) : 1
2 3 2, 2 (1) : 1
2 3 3, 2 (1) : 1
3 1 1, 2 (1) : 1
3 1 2, 1 (1) : 1
3 1 3, 1 (1) : 1
3 2 1, 3 (1) : 1
3 2 2, 2 (1) : 1
3 2 3, 1 (1) : 1
3 3 1, 2 (1) : 1
3 3 2, 2 (1) : 1
3 3 3, 2 (1) : 1

4. Model validation

In order to be acceptable, a scientific model must be reliable, that is, it must be susceptible to be replicated and validated. The validity of the model must be external and internal.

Internal validity concerns the analysis of the causality relationship in the model independently from subjective factors, that is, the following question must be answered: Can we obtain the same type of relationship between the elements by other means in addition to those used in a specific experiment? External validity is related to the generalizability of the model.

In our proposed model, reliability and external validity are guaranteed, as the model is described in detail in this paper and it can be easily replicated for the analysis of other environmental disasters (after some adjustments of the data, mainly related to the input variables). The internal validity was evaluated by comparing the results obtained in Mars to a manual classification by a specialist (expert) list (in: "no_evidence", "assigned", and "recognized") of 30% (10 random years) of the 29 years that were analyzed in the corpus. The full classification of each individual article for the 10-year period is available at https://doi.org/10.6084/m9.figshare.13583555. Table 3 below presents a summary of the results.

Table 3. Most frequent values in the expert evaluation for the nature of the news.

Year	Number of news with ethical	Expert assessment	Results in the Mars	Comparison
1989	41	no_evidence	no_evidence	equal
1990	42	no_evidence	no_evidence	equal
1991	48	assigned	assigned	equal
1994	23	assigned	assigned	equal
1996	8	no_evidence	no_evidence	equal
1999	15	no_evidence	recognized	different
2000	12	no_evidence	no_evidence	equal
2008	9	assigned	assigned	equal
2014	5	no_evidence	no_evidence	equal
2017	7	no_evidence	no_evidence	equal

Although the percentage of disagreement/error seems high (10%) when presented year by year, we observe that only 30% of the analyzed years were tested this way. There is a total of 518 news items with terms from the ethical vocabulary (available at https://figshare.com/articles/dataset/Anex_3_Ethical_vocabulary/13567838) while in total we manually classified 210 news items (referring to the 10 years tested), which represents 40% of the result of mapping the ethical vocabulary. The year 1999 presents only news items and the application of the model presents disagreement/error in the classification of 4 of them. From these data, we can infer that the model had an error in 4 out of 210 news classified manually, reducing the error rate to less than 2%. Considering the large number of tested news and the complexity involved in the analysis, we consider that the model performs well in relation to the classification of news.

5. Discussion and final remarks

The application of our model reveals that in the years 1991, 1992, 1994, 1995, 2005, 2006, 2008 and 2012 the news published about Exxon brought up information on topics related to the assignation of social responsibility to Exxon for the environmental disaster in Alaska.

Regarding the events that occurred in those years, we highlight that in 1991 a federal judge accepted a package of $1 billion USD in criminal and civil agreements to close the state and federal lawsuits against Exxon Corporation. Subsequently, in 1992, transcripts of telephone conversations were released among employees of the oil industry during the response to the environmental disaster. A federal jury in 1994 imposed a penalty of $5 billion USD in punitive damages to 34,000 fishermen and other Alaskans, but the American public opinion found the sums imposed on Exxon to be totally inadequate and insufficient to prevent it from being penalized in the future. In 1995, Exxon announced a donation of $5 million USD to the protection of the tiger habitat in the wild, which, in our opinion, was a strategy to manipulate the public opinion as the company was accused of not being fast enough to act upon the oil spill. CEO Lawrence G. Rawl, who led the review of Exxon's operations in 1980 and shaped the company's response to the Exxon Valdez oil spill, died in 2005 at the age of 76. A year later, under the George W. Bush Administration, the federal appeals panel reduced the punitive damages from $4.5 billion USD against Exxon Mobil to $2.5 billion USD, arguing that the company's negligent conduct was unintentional. In 2008, the United States Supreme Court further reduced the claim to $500 million USD. In 2012, the book "Private Empire" was published, revealing the true extent of the Exxon Mobil corporation's power in American and foreign politics. In this sense, it should be noted that the "assigned" values obtained using our model are related to the relevant events in the same period.

Regarding the manual classification of the specialist, we should also note that out of the 210 analyzed items (news), only 21 of them were classified manually as "recognized", which means that only 10% of the news that present terms from the ethical vocabulary indicate the recognition of the corporate social responsibility. The number of items classified manually as "no_evidence" was 118, that is, more than half of the news that were classified did not show enough evidence to be classified as assignation or recognition of social responsibility, while the news items classified as "assigned" totalled 71 out of the 210 that were tested, representing 34% of them.

The results of the annual analysis show that 3% of the analyzed years indicate topics related to the recognition of corporate social responsibility by Exxon; 28% refer to topics surrounding the assignation of responsibility to Exxon for causing the accident, and 69% do not show enough evidence to indicate whether the most popular texts about the accident of the year suggest the assignation or recognition of corporate social responsibility.

Despite comparing different magnitudes, we noticed that the results are similar in both types of analysis (by year and by news). This aspect shows that, despite the complexity of an analysis of this kind in which the model is applied considering isolated words in the text and not the semantic aspect, the proposed model proved to be adequate to conduct this type of analysis. In this sense, we also highlight the advantage of the high potential for adapting this model to the study of assigned and recognized social responsibility in relation to other environmental accidents, with the only requirement of some adjustments in the inputs as aforenoted.

Acknowledgements.

References

[1]   Barnard, J.W. Exxon collides with the "Valdez Principles". Faculty Publications (1990). Available at: 1039.https://scholarship.law.wm.edu/facpubs/1039

[2]   Barros, L. C.; Bassanezi, R. C. Tópicos de Lógica Fuzzy e Biomatemática. Unicamp/IMECC (2015).

[3]   Campbell, J. L. Why would corporations behave in socially responsible ways? An institutional theory of corporate social responsibility. Academy of management Review, 32(3), 946-967 (2007).

[4]   Garriga, E.; Melé, D. Corporate social responsibility theories: Mapping the territory. Journal of business ethics 53(1-2), 51-71 (2004).

[5]   Hicks, D.; Wang, J. The New York times as a resource for mode 2. Science, technology, & human values 38(6), 851-877 (2013). DOI https://doi.org/10.1177/0162243913497806

[6]   Janssen, C.; Sen, S.; Bhattacharya, C. B. Corporate crises in the age of corporate social responsibility. Business Horizons 58(2), 183-192 (2015). DOI: http://dx.doi.org/10.1016/j.bushor.2014.11.002

[7]   Jones, T.M. Corporate Social Responsibility revisited, redefined. California Management Review 22(3), 59-67 (1980). DOI: https://doi.org/10.2307/41164877

[8]   Lindgren, A.; Swaen, V. Corporate Social Responsibility. International Journal of Management Reviews 12(1), 1-7 (2010). DOI: https://doi.org/10.1111/j.1468-2370.2009.00277.x

[9]   Marcelino, L.V.; Pinto, A.L.; & Marques, C.A. Scientific specialties in Green Chemistry. Iberoamerican Journal of Science Measurement and Communication; 1(1), (2020). DOI: https://doi.org/10.47909/ijsmc.06

[10]   Matten, D.; Moon, J. "Implicit" and "explicit" CSR: A conceptual framework for a comparative understanding of corporate social responsibility. Academy of management Review 33(2), 404-424 (2008).

[11]   McWilliams, A.; Siegel, D. Corporate social responsibility: A theory of the firm perspective. Academy of management review 26(1), 117-127 (2001).

[12]   Story, J; Neves P. When corporate social responsibility (CSR) increases performance: exploring the role of intrinsic and extrinsic CSR attribution. Business Ethics: A European Review 24(2), 111-124 (2015). DOI: https://doi.org/10.1111/beer.12084

[13]   Våland, T.; Heide, M. Corporate Social Responsiveness: Exploring the Dynamics of "Bad Episodes". European Management Journal 23(5), 495-506 (2005). DOI: https://doi.org/10.1016/j.emj.2005.09.005

[14]   Visser, W. Corporate social responsibility in developing countries. In: The Oxford handbook of corporate social responsibility. (2008). DOI: 10.1093/oxfordhb/9780199211593.003.0021

[15]   Zadeh, L. A. Fuzzy Sets. Information and Control 8, 338-353 (1965).

[16]   Zadeh, L. A. Fuzzy logic. Computer 21(4), 83–93 (1988). DOI:10.1109/2.53

[17]   Zadeh, L A. Is there a need for fuzzy logic? Information Sciences 178 (13), 2751-2779 (2008). DOI: 10.1016/j.ins.2008.02.012.