

### How to cite this article in bibliographies / References

S Arce García, N Orviz Martínez, T Cuervo Carabel, I Fernández Suárez (2017): “Occupational accidents and their prevention in the Spanish digital press”. *Revista Latina de Comunicación Social*, 72, pp. 1.608 to 1.625.

<http://www.revistalatinacs.org/072paper/1237/87en.html>

DOI: [10.4185/RLCS-2017-1237en](https://doi.org/10.4185/RLCS-2017-1237en)

# Occupational accidents and their prevention in Spanish digital press

**Sergio Arce García** [CV] [ORCID] [G] Universidad Internacional de La Rioja, UNIR / International University of La Rioja, Spain / [sergio.arce@unir.net](mailto:sergio.arce@unir.net)

**Natalia Orviz Martínez** [CV] [ORCID] [G] Universidad Internacional de La Rioja, UNIR / International University of La Rioja, Spain / [natalia.orviz@unir.net](mailto:natalia.orviz@unir.net)

**Tatiana Cuervo Carabel** [CV] [ORCID] [G] Universidad Internacional de La Rioja, UNIR/ International University of La Rioja, Spain / [tatiana.cuervo@unir.net](mailto:tatiana.cuervo@unir.net)

**Iván Fernández Suárez** [CV] [ORCID] [G] Universidad Internacional de La Rioja, UNIR/ International University of La Rioja, Spain / [ivan.fernandez@unir.net](mailto:ivan.fernandez@unir.net)

## Abstract

**Introduction.** Occupational accidents and their prevention have become a serious day-to-day human, social and economic problem for businesses, while the news media have traditionally tended to give central coverage to fatal work accidents. The objective of this article is to study how born-digital newspapers address this issue. **Methods.** The study is based on the content analysis of media framing by text mining techniques of a sample of news pieces related to the work accidents and their prevention, published over a period of two years and seven months, by five of the main digital newspapers in Spain. **Results and conclusions.** Evidence shows that digital news media respond reactively to workplace accidents, just like their print counterparts, although there are important differences in the treatment of this subject according to the editorial line of each online newspaper.

## Keywords

Occupational accidents; prevention; text mining; online press; framing.

## Contents

1. Introduction. 1.1. Occupational accidents in Spain. 2. Objectives and methods. 2.1. Population and sample. 2.2. Information processing procedure. 3. Results. 3.1. Hierarchical cluster analysis. 3.2. Correspondence and co-occurrence analysis. 4. Conclusions. 5. References.

Translation by **CA Martínez-Arcos**  
(PhD in Communication, University of London)

## 1. Introduction

The media, and the press in particular, are a reflection of society and the public opinion, but at the same time are a channel to transmit opinions and demands from leaders to citizens and vice versa. They are able to set and manage their own agenda as well as the agenda of the parties that make up society (Crespo, Moreno and Delgado, 2003:13-14). Informative texts need narrative structures that articulate the discourse, and the media organises it according to a series of perceptions that limit the perception of reality. In this way “some aspects of the reality perceived through the news will be more prominent than others” (Ardèvol-Abreu, 2015: 424) and the media will set, following their editorial line, the salience of the information as well as the frames to interpret it. Individuals, after consulting the media in a regular basis, will adopt this agenda “due to an inherently psychological need to be in harmony with their environment” (Álvarez-López, 2004:58).

Occupational or work accidents in Spain and their prevention are an issue that has been present in the media’s agenda throughout time. Already in 1933, during the Second Republic, María Palancar and her husband Eugenio Pérez Botija won the Marvá award with the work titled “Prevention of work accidents” (Galán-García, 2008), which criticised the media for focusing on the scabrous and social aspects of work accidents, instead of educating workers. Publications on this subject have been scarce, limited in general to magazines and professional conferences. In recent years, however, academic studies on the local and national press have begun to emerge. News pieces focus on the fatal nature of occupational accidents, lack technical explanations and do not address accident prevention, while the new preventive legislation has not change much this media treatment. It should be noted that after the deep economic crisis broke in 2007 in Spain, this issue almost disappeared from the agenda of the print press (Arce-García, 2017: 18-21).

Of all the studies carried out on this subject only one focuses on the digital press, and was published in 2012 by Spain’s National Institute for Safety and Hygiene at Work (INSHT), currently known as the National Institute for Safety, Health and Well-being at Work (INSSBT). Berjón, Urquijo and Fraile (2012:20-21) analysed 338 print newspapers and 373 digital newspapers, but in a global and joint manner. These authors compared the print press against its digital counterpart, developing an eminently quantitative study that involved some qualitative aspects. They found out that there was a greater presence of denunciations of working conditions and the non-existent preventive measures, although the fatal nature of work accidents continued to be the central element of the news. The presence of reports, interviews and opinion articles is almost non-existent, especially in online newspapers. Moreover, public institutions and trade unions were scarcely used as sources of information in the news about work accidents.

### **1.1. Occupational accidents in Spain**

Occupational accidents are one of the most transcendental social and economic problems for a country. The economic costs are important for companies, ranging from sanctions, losses in working-days, materials and insurance coverage, but also represent social costs for the whole country, with losses reaching 2% of the gross domestic product (GDP) of Spain in the 2000-2007 period (Remo-Díez, 2011:221-225). For companies, there are studies that indicate that the real costs of workplace injuries range from four times the direct costs, i.e., the easily identifiable and quantifiable costs, to even twenty times that amount for certain sectors, such as construction (Narocki, 1999: 108).

In the personal field, occupational accidents are an important human loss, with its consequent social pain. Spain is one of the European countries with the highest work accident rates, way higher than the average of the countries in the Euro zone. With the coming into effect of the basic legislation on accident prevention at work in 1995 there was an increase in the rates up to the year 2000, when the rates begin to decrease, getting closer to the average of the rest of countries. With the start of the economic crisis in Spain in the year 2007, there was a sharp decline in these rates (Arce-García, 2017:94). Once the worst years of this crisis passed, the accident rates increased again. In 2015 there were 529,248 accidents causing sick leave and 515 causing death; in 2016, there were 566,235 accidents causing sick leave and 496 causing death; and from January to July 2017, 333,993 accidents and 323 deaths were recorded (Ministry of Employment and Social Security of Spain, 2017).

The correspondence between work accidents in a country and its economic cycles has been verified in various studies, like the one carried out in the United States by Asfaw, Pana-Cryan and Rosa (2010:6), who found out that 1% increase in the GDP increases the incidence of workplace injuries by 1.6% (calculation of accidents per thousand or million workers according to the size of the group analysed), while a 1% decline in the GDP implies a decrease of 0.2% in these rates. In Spain, the studies carried out by Fernández-Muñiz, Montes-Peón and Vázquez-Ordás (2016) corroborate the dependence between the incidence of workplace injuries and economic cycle.

## **2. Objectives and methods**

The initial research hypothesis is that born-digital newspapers include occupational accidents and their prevention in their agenda, but that there are differences according to the editorial lines of each medium.

The objective is to analyse digital newspapers' treatment of the incidence of workplace injuries and their prevention to determine whether this topic is included in their agenda and identify the treatment given to it. Previous studies on the print press will be complement with this work on the increasingly important born-digital press, which so far has not been qualitatively examined nor has been subjected to a comparative study.

The study of how born-digital newspapers treat work accidents and their prevention in their agenda starts with a quantitative approach that seeks to determine the number of journalistic pieces published by a sample of five newspapers. This way, the analysis takes into account all the journalistic pieces that contain terms related to the subject in question in their digital body of text.

These pieces were subsequently subjected to content analysis to determine their main frames by means of framing theory, which indicates that the treatment of the event and other journalistic pieces will

produce a change in opinions and attitudes “due to subtle alterations in the definition of the problem” (Igartua & Humanes, 2004:55). This analysis uses text mining techniques to determine the main media frames with minimal human intervention or interference (Fu-ren, De Hao and Dachi, 2016). Text mining techniques have many applications in the information society and can be used to summarise contents and information, extract useful ideas, and analyse public opinion, social networks, marketing, e-commerce and the internet-user relationship (Justicia-de-la-Torre, 2017:31-37).

The application of text mining techniques to the media is recent and has very few practical applications in the scientific literature, and it is practically non-existent in the Spanish language. However, there are some research applied to case studies in the field of communication, which show that “text mining techniques allow the purely automatic organisation and classification of documents, as well as the recovery and extraction of specific information, evaluation and prediction. In these techniques, human intervention has been reduced to a minimum, thus reducing the problems generated by the human codification of textual information (Álvarez-Gálvez, Plaza, Muñiz and Lozano-Delmar, 2014:923).

The degree of accuracy and reliability of this technique and its use in Spanish-language texts has been corroborated in various studies, such as those carried out by Contreras-Barrera (2016:41), Álvarez-Gálvez, Plaza, Muñiz and Lozano-Delmar (2014:929) and Aureli (2017:20-21). In this way, its reproduction is favoured, and the researcher’s personal biases are eliminated. Moreover, all the authors emphasise that the results offered by text mining should be later analysed and verified by researchers to eliminate data that make no sense, to avoid misunderstandings or unrealistic situations. The study carried out by Yan and Minnhagen (2018:561) shows that the use of this technique of analysis does not have differences across different languages, but that the problems come from the use of polysemic words, which requires a subsequent analysis of the results.

The main use of text mining is based on hierarchical cluster analysis, where the initial steps and the pre-processing strategy determine their results, and have shown that the analyses based on K-means type algorithms present the best results (Baojun, Hua, Ye, 2017). For the treatment of text there are several methods on how algorithms should collect and treat the text. For European languages there are two types: *Stemming*, which is based on the reduction of words to a common stem by clipping off the unnecessary morphemes, and *Lemmatisation*, which is based on the clustering of words by morphemes guided by the knowledge by the computer program of the dictionary and morphology of the language for this association (Singh, Gupta, 2017:159). Although in principle one might think that *lemmatisation* could offer more reliable results, various studies such as those carried out by Kettunen, Kunttu and Järvelin (2005) and Balakrishnan, Humaidi and Lloyd-Yemoh (2016) point out that the results in different languages between both methods present insignificant differences. Anyways, *stemming*, the most used method in computer programs for Latin languages, can present two types of errors in certain cases: false positives and false negatives, respectively, with words that have an almost equal morphology and different meaning or polysemic words (Hajeer, Ismail, Badr, Tolba, 2017).

The choice of this method of analysis, therefore, comes from the advancement of these mathematical algorithms that, in addition to eliminating the possible human intervention in the selection of the frames, can process much larger amounts of text than any researcher could process manually. Its use, accompanied by a subsequent review to detect frames that make no sense or have no relation, allows us to interpret correctly the results obtained.

## 2.1. Population and sample

To carry out the analysis of the journalistic pieces we have chosen five born-digital newspapers that do not have a print version/edition. For their selection we identified the ten most-visited born-digital newspapers: *Elconfidencial.com* (with 9.9 million unique users per month in 2016), *Elespañol.com* (5.7 million), *Okdiario.com* (5.4 million), *Eldiario.es* (4.5 million) and *Publico.es* (3.5 million), which are the 1<sup>st</sup>, 2<sup>nd</sup>, 3<sup>rd</sup>, 5<sup>th</sup> and 6<sup>th</sup> most visited online newspapers, respectively (Prnoticias, 2016). The fourth most visited newspaper, the *Huffingtonpost.es*, was ruled out because its search engine did not offer the same ease and reliable results as the rest of digital newspapers. The period of time established for the collection of journalistic pieces was 2 years and seven months: from 1 January 2015 to 31 July 2017.

For the collection of the sample we introduced the following terms in the search engines of each of the digital newspapers: *accidente laboral* (“Occupational accident”), *accidente de trabajo* (“work accident”) and *prevención de riesgos laborales* (“prevention of occupational hazards”). Following the guidelines set by Kayser (1979:151), to be able to compare different multiple editions of the same newspapers we eliminated those pieces that appeared in the search engines that belong to the local editions of the newspapers, which is an aspect that happens mainly in *Eldiario.es*, which has published numerous pieces related to the subject under study in its local sections.

## 2.2. Information processing procedure

The identification information of the pieces, such as date, title, author, sources, presence of photograph and author, if any, journalistic genre, subject and sector, was collected in the free computer statistical program PSPP version 0.10.2. The same program was used to perform the quantitative study of frequencies of documents per newspaper, year, genre, presence of photographs, author and sources.

After the collection of the pieces, the text of each newspaper was stored in flat text files, separated each one by publication date with a tag that allowed the analysis programs to interpret a jump. For the analysis of text mining we used KH Coder, a free software based on R modules for text mining. The pre-process used for this treatment was stemming with the snowball algorithm adapted for Latin languages, where we introduced in Spanish the articles, prepositions and conjunctions to be eliminated from the analysis.

Each one of the newspapers included in the study were subjected to the following process:

- A hierarchical cluster analysis, through Ward’s minimum variance criterion with cosine distance, to graphically represent the main frames described by the newspapers in a dendrogram that shows the relationship between words and their frequency.
- A correspondence analysis among the most frequently used words, to determine whether they follow a pattern of behaviour and appearance in each journalistic piece. This study took into account the sixty most used terms in each newspaper. They have been expressed on two dimensions that cross the number of terms per document and the relationships of greater dependence between words per chi-square test between them. This allowed us to see the clustering or scattering of terms in the diagram, which shows whether the story is similar or different between the journalistic pieces.



- The establishment of co-occurrence networks between the words that reach a Jaccard coefficient greater than 0.2, so that the relationships between the different frames can be verified. This level makes the studies of co-occurrences to be analysed between words with a minimum term frequency (tf) that will be different for each newspaper. Thus *Eldiario.es* has a tf of 35, *Elconfidencial.com* a tf of 5, *Elespañol.com* a tf of 25, *Okdiario.com* and *Publico.es* a tf of 10. This analysis aims to show the relationship between pattern words with similar behaviour. The thickness of the lines that unite the terms are represented according to the degree of co-occurrence between the two, while the size of the circles representing the words indicates their frequency. The main frame is represented in different colours.

### 3. Results

The number of pieces about occupational accidents and their prevention that were published by the five main digital newspapers in Spain has been uneven across newspapers and over time. As Table 1 shows, *Eldiario.es* published the largest number of pieces, followed by *Okdiario.com* and *Elespañol.com*. They are followed by *Publico.es* and, with only four pieces, *Elconfidencial.com*. In the first year analysed, 2015, almost all the newspapers published similarly low numbers of pieces, between two and four, but there was increase in the number of pieces published over time (taking into account that data collection covered until July 2017), mainly in *Eldiario.es*, and with the exception of *Publico.es*, which maintained the same amount of pieces, and *Elconfidencial.com*, which even decreased the number of pieces

**Table 1: Number of Pieces**

	<i>Elconfidencial.com</i>	<i>Eldiario.es</i>	<i>Elespañol.com</i>	<i>Okdiario.com</i>	<i>Publico.es</i>
2015	3	4	4	2	4
2016	0	10	9	14	4
2017 (31 July)	1	18	10	10	4
Total	4	32	23	26	12

Source: Authors' own creation.

The analysis of the journalistic genre of the pieces shows notable differences between the different newspapers as shown in table 2. Only two newspapers, *Eldiario.es* and *Elespañol.com*, published pieces that belong to the opinion genre, mainly articles written by columnists. Of the informative genres, the news report is the most used, except for *Elconfidencial.com* and *Publico.es*, where the feature is predominant.

**Table 2: Journalistic genres**

	<i>Elconfidencial.com</i>	<i>Eldiario.es</i>	<i>Elespañol.com</i>	<i>Okdiario.com</i>	<i>Publico.es</i>
Feature- Informative	4	2	3	6	11
News- Informative		22	17	20	1
Article- Opinion		8	3		

Source: Authors' own creation.

With regards to the topic that motivates the journalistic piece (table 3), the information of news about the incidence of workplace accidents is not usually the main reason for the pieces. Instead, these news stories are used as lateral explanation for other subjects where there is a work accident that conditions the main subject, which is qualified under “others”. In this way, the incidence of workplace accidents or their prevention is not usually the main element of the piece, except for *Publico.es*, which privileges statistical information that shows the evolution of occupational accidents and its comparison with surrounding countries. Likewise, the consequences of sentences derived from work accidents are scarce and only appear in *Eldiario.es* and *Publico.es*.

The productive sector in which the work accidents occur is not usually mentioned in the pieces and when it is mentioned it does not coincide with the sectors with the highest rates of accidents. Thus, *Eldiario.es* focuses on construction (3) and transport (1), *Elespañol.com* on construction (2) and transport (1), *Okdiario.com* on construction (3), metal (1), transport (1) and others (4) and *Publico.es* on transport (1) and others (1). *Elconfidencial.com* does not specify the sectors.

**Table 3: Themes**

	<i>Elconfidencial.com</i>	<i>Eldiario.es</i>	<i>Elespañol.com</i>	<i>Okdiario.com</i>	<i>Publico.es</i>
Accident	2	8	5	9	1
Diseases			1		
Sentences		2			2
Statistics				1	5
Other	2	22	17	16	4

Source: Authors' own creation.

With the exception of three pieces (one in *Eldiario.es*, one in *Elconfidencial.com* and one in *Elespañol.com*), they all contain photographs accompanying the information. These photographs are mainly generic photographs (100% in *Eldiario.es*, 75% in *Elconfidencial.com*, 95.65% in *Elespañol.com*, 85.71% in *Okdiario.com* and 83.33% in *Publico.es*) and were obtained from agencies (76.19% in *Eldiario.es*, 100% in *Elconfidencial.com*, 40% In *Elespañol.com*, 50% in *Okdiario.com* and 66.67% in *Publico.es*). When a photograph does not come from agencies it is signed by a journalist.

Many pieces do not specify the information sources, but there are differences in this regard across newspapers. As Table 4 shows, trade unions are the main source in *Eldiario.es* and *Publico.es*, while staff and public authorities are the main source in the remaining newspapers. The workers from the company where the accident occurred are used as sources on only one occasion, in *Elespañol.com*.

**Table 4: Sources**

	<i>Elconfidencial.com</i>	<i>Eldiario.es</i>	<i>Elespañol.com</i>	<i>Okdiario.com</i>	<i>Publico.es</i>
Unions		3	1	3	3
Government	2		2	5	2
Workers			1		
Other	1	1		1	1

Source: Authors' own creation.

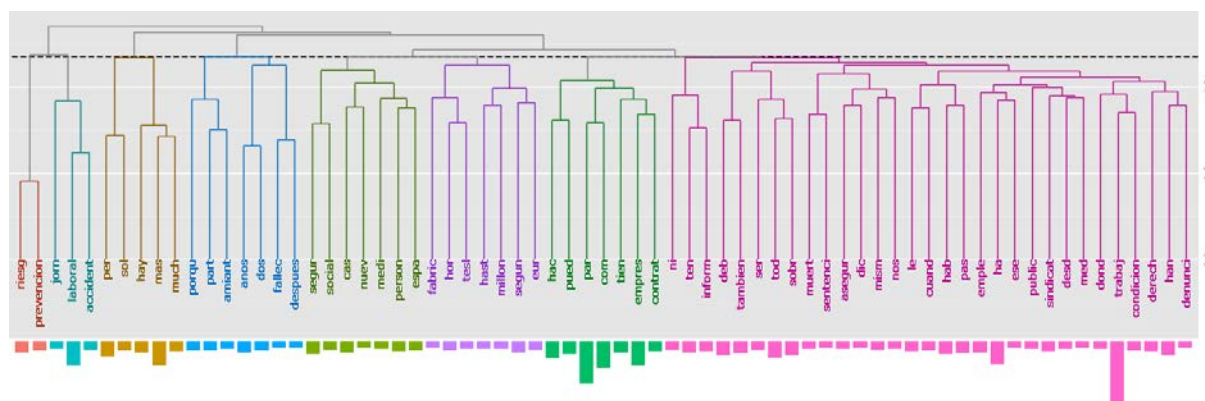
### 3.1. Hierarchical cluster analysis

Text mining analysis was performed in the different journalistic pieces to identify their most frequently used and, therefore, most representative conceptual frames. Hierarchical cluster analysis was performed in each of the newspapers to identify the main frames.

Seven clusters were identified in *Eldiario.es*, as shown in Figure 1. As we can see, the main cluster, which is formed by the words with the most frequency of appearance, revolves around workers, the denunciation of work conditions and the fight for workers' rights. This is related to trade unions, sentences, death, denunciation, insurance, information. The second cluster of frequently used words is related to companies and types of contracts. The words risk and prevention appear in another cluster, just like accident and work. In the rest of the cluster there are relations between asbestos and deceased years later, social security in Spain and manufacturing. In this way, we can see that the main frame of *Eldiario.es* is the denunciation of the conditions of workers, including unions in their relations and exposing the judicial sentences for workers' death as proof, while entrepreneurs and contracts are the second frame. The newspaper speaks about occupational accidents and risks and prevention, but in a lateral way and not mainly related to the main frames.



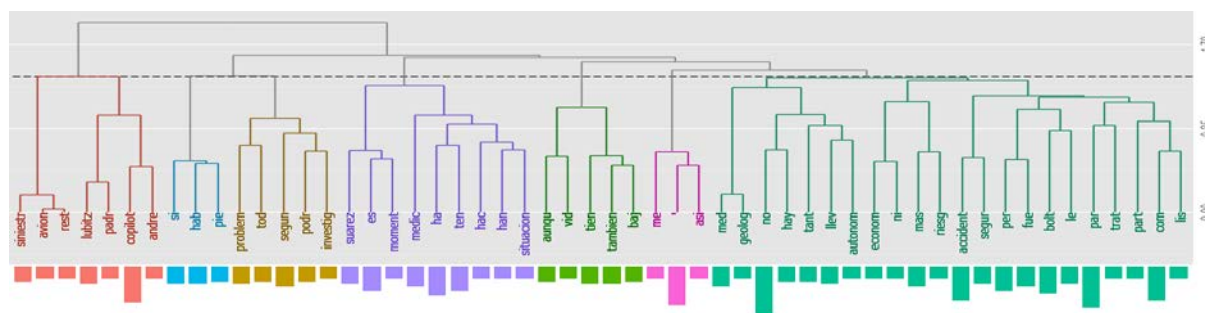
**Figure 1: Hierarchical cluster analysis of *Eldiario.es***



Source: Authors' own creation.

Seven clusters were identified in *Elconfidencial.com*, as we can see in Figure 2, which shows that the most frequently used terms were accidents and safety, but related mainly to the economy, and accompanied by a lack of words of denunciation. In the rest of the cluster one can see some cases, like the place crash caused by a German pilot who committed suicide and the investigation around one accident. Therefore, there is no discourse nor denunciation of tragic work accidents or their relation to prevention.

**Figure 2: Hierarchical cluster analysis of *Elconfidencial.com***



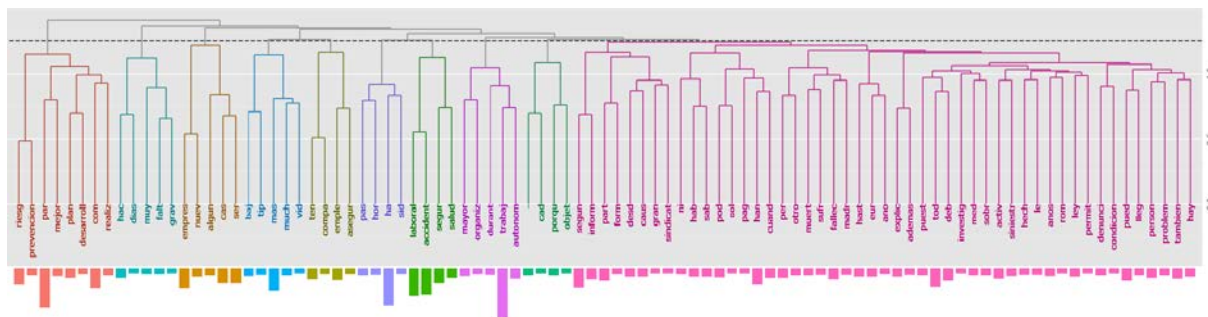
Source: Authors' own creation.

Ten clusters were identified in *Okdiario.com* (see Figure 3). The cluster that uses the term worker most frequently is related to self-employment, but has no more associations. The second cluster with words that appear the most uses the words prevention and risks, which are related to the need for more plans and development. In the cluster with the largest number of associated words includes the terms trade unions, investigation, death, complaints and conditions, but it is a very extensive frame with very little frequency of appearance and with many words associated at a large distance with worker, accident and prevention.

The cluster analysis in *Okdiario.com* indicates that the newspaper addresses prevention and risks, but not from a point of view that denounces the conditions of workers, and instead from the point of view of development and planning. Several actors are introduced, but they are rarely employed and are not

related to the incidence of work accidents and their prevention, but to death and unions' complaints about the working conditions.

**Figure 3: Hierarchical cluster analysis of Okdiario.com**

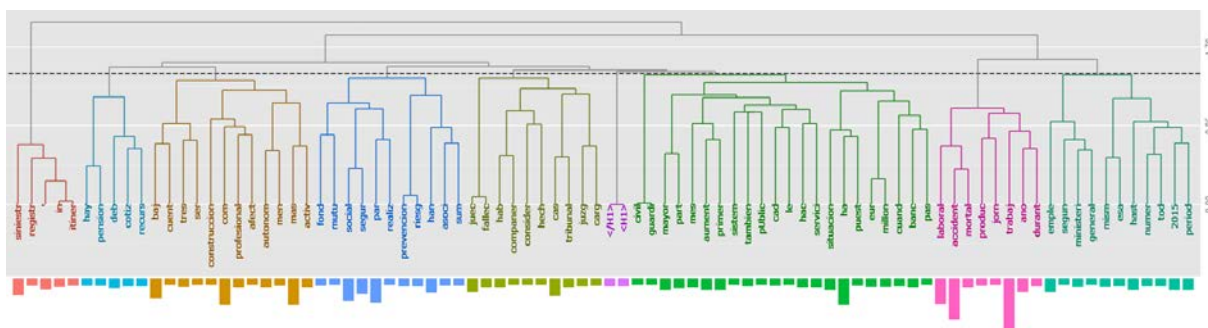


Source: Authors' own creation.

Eight clusters were identified in *Publico.es*, as shown in Figure 4. The main cluster is formed with the most frequently used terms accident and worker. The word accident is clearly related to the words labour and deadly, as well as to working-day and production. The second largest cluster of frequently used words refers to self-employment and construction, one of the sectors with the highest accident rate. The third cluster is formed with the words risk prevention, benefit societies and social security. In the rest of the clusters the accident is related to *in itinere* accidents (those that happen on workers' way to or from work), which must be endowed with resources, and the appearance of judges and tribunals for cases of fatal accidents.

*Publico.es* therefore associates fatal accidents with working conditions and resources, without putting this assertion in the mouth of unions, highlighting sectors with high accident rates, such as construction and self-employment. Risk prevention is limited to benefit societies and social security as a relief aid that is external to the main frames of the pieces.

**Figure 4: Hierarchical cluster analysis of Publico.es**

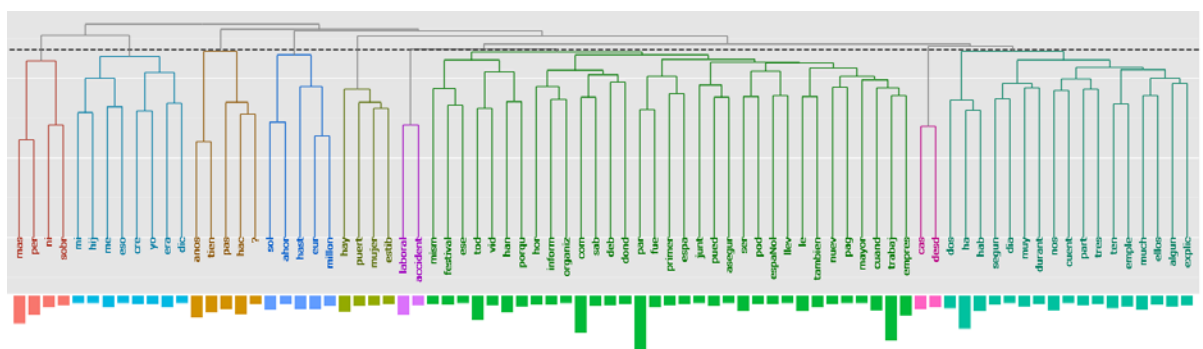


Source: Authors' own creation.

Nine clusters appeared in *Elespañol.com*, as we can see in Figure 5, where the frame with the most-used terms is quite extensive, but workers are related with companies and the main terms are associated with information and organisation. The rest of the clusters do not show clear lines and are limited to very specific events mentioned in the pieces under analysis. There is a cluster formed with the words

accident and work, but it does not correspond with other words. The terms prevention and risk do not appear in the analysis as they do not reach the minimum frequency required for analysis. Therefore, it can be concluded that for *Elespañol.com*, workplace accidents do not constitute a main frame and are only a side issue to the main frames of the journalistic pieces under analysis.

**Figure 5: Hierarchical cluster analysis of *Elespañol.com***

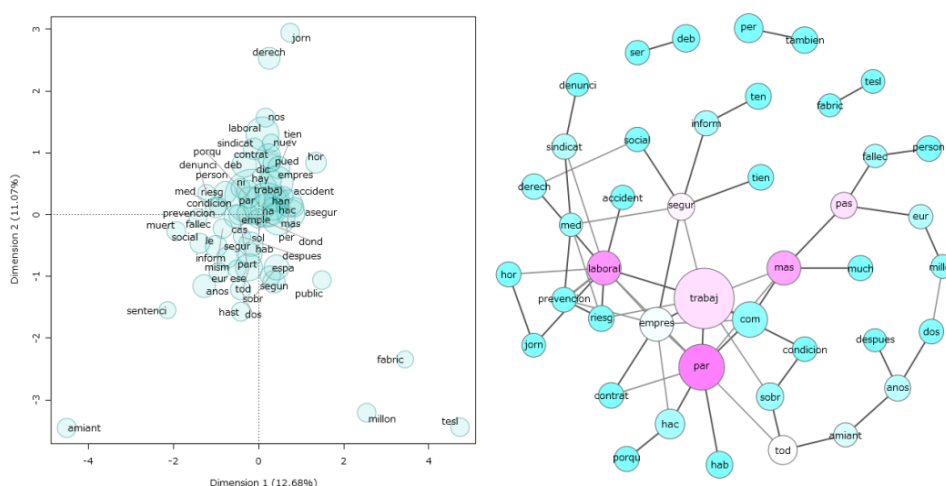


Source: Authors' own creation.

### 3.2. Correspondence and co-occurrence analysis

As Figure 6 shows, the correspondence analysis of *Eldiario.es* indicates that the terms are quite grouped together, so there is a very similar pattern between the terms, which indicates that the different journalistic pieces are quite similar in their discourse. The co-occurrence network revolves around the main axis, marked by worker and work, which have important and close relations with accident, prevention and risk, as well as right and union with complaint. As we can see, other relations of co-occurrences are about working conditions and asbestos. Therefore, the main discourse of complaint and relation with the incidence of work accidents and their prevention is made around the term work.

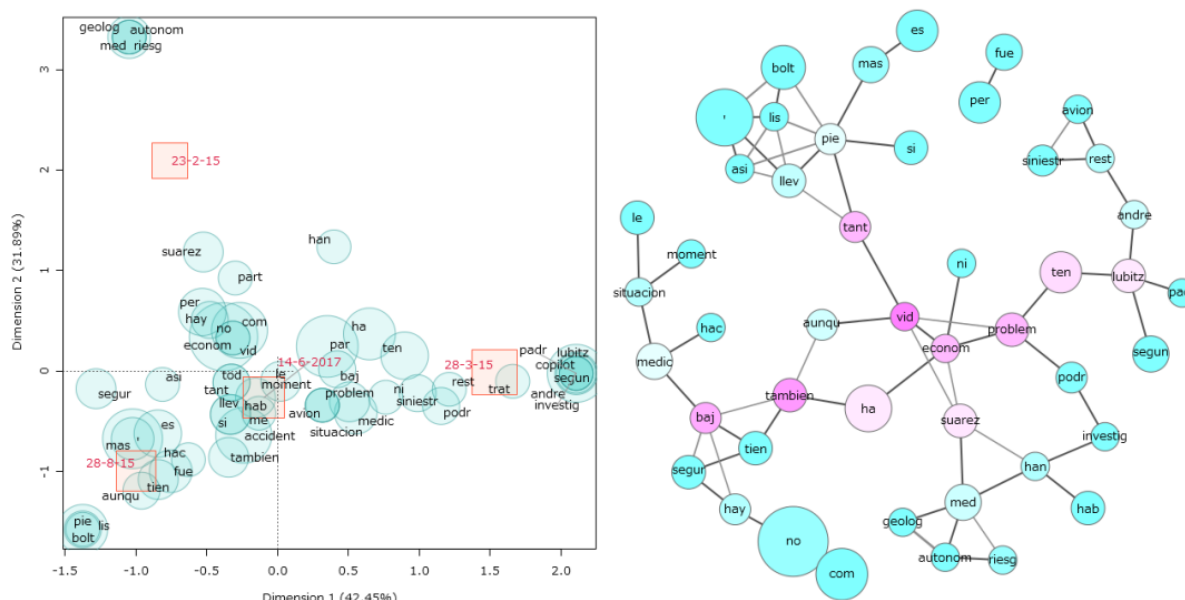
**Figure 6: Analysis of correspondence and co-occurrence of *Eldiario.es***



Source: Authors' own creation.

In *Elconfidencial.com* there is a great dispersion of terms in the correspondence analysis, so it can be induced that there are enough differences between what is written in the journalistic pieces. There is no common pattern among their writings. In the analysis of co-occurrence networks, the main pattern has small sizes, so they have few frequencies. This main frame includes the economic problem while the incidence of accidents is left in a distant and minor relationship, as shown in Figure 7.

**Figure 7: Analysis of correspondence and co-occurrence of *Elconfidencial.com***

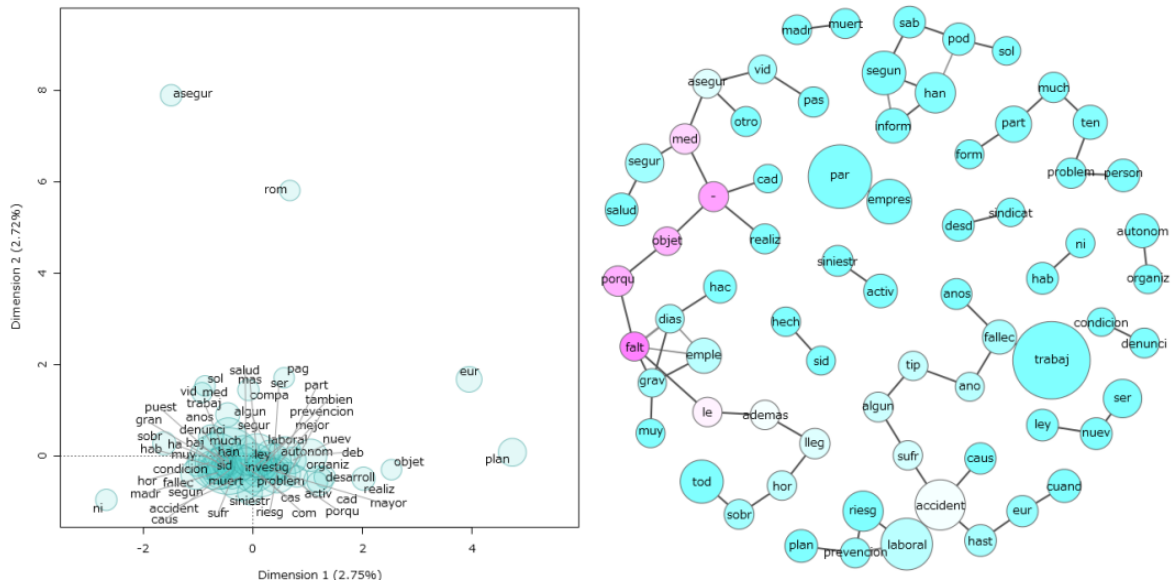


Source: Authors' own creation.

In *Okdiario.com* there is a great clustering of terms in the correspondence analysis, so there is a pattern of behaviour between the different pieces published by this newspaper. There are few words scattered off the centre of dimensions so that the different pieces follow a similar line. In the co-occurrence network, the worker is related to death and years of work and occupational accidents, which is also linked with prevention, risk and plan. All this line of co-occurrences is not linked to others, presenting a wide variety of clusters, many of them small, related to each other. Therefore *Okdiario.com*, as shown in Figure 8, presents a coherent line between its pieces, but there is also a multitude of aspects, being one of them the relation between workers' accidents and the prevention of occupational risks and their planning.

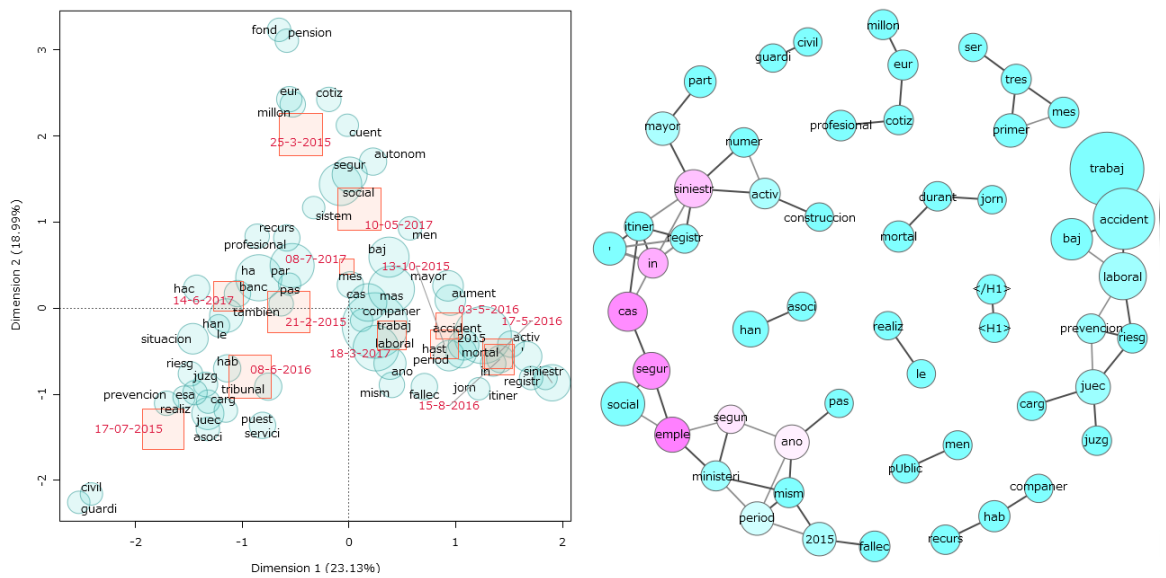
In *Publico.es*, the correspondence analysis shows great scattering of terms, which indicates that there are no clear patterns of behaviour across the different pieces, and there is great variety across texts. The analysis of the co-occurrence network shows that the words with the highest frequency of appearance, worker, accident and work, are directly related with risk, prevention, as well as judges, charges and courts. Likewise, there is a main frame line from the statistical point of view of relationship, and it is equally distributed around accident, security and employment, which is related to a lesser extent with ministry, construction and deceased. subsequently, there are other clusters with relationships, but these are minor and have few interconnections, as shown in Figure 9.

**Figure 8: Analysis of correspondence and co-occurrence of *Okdiario.com***



Source: Authors' own creation.

**Figure 9: Analysis of correspondance and co-occurrence of *Publico.es***



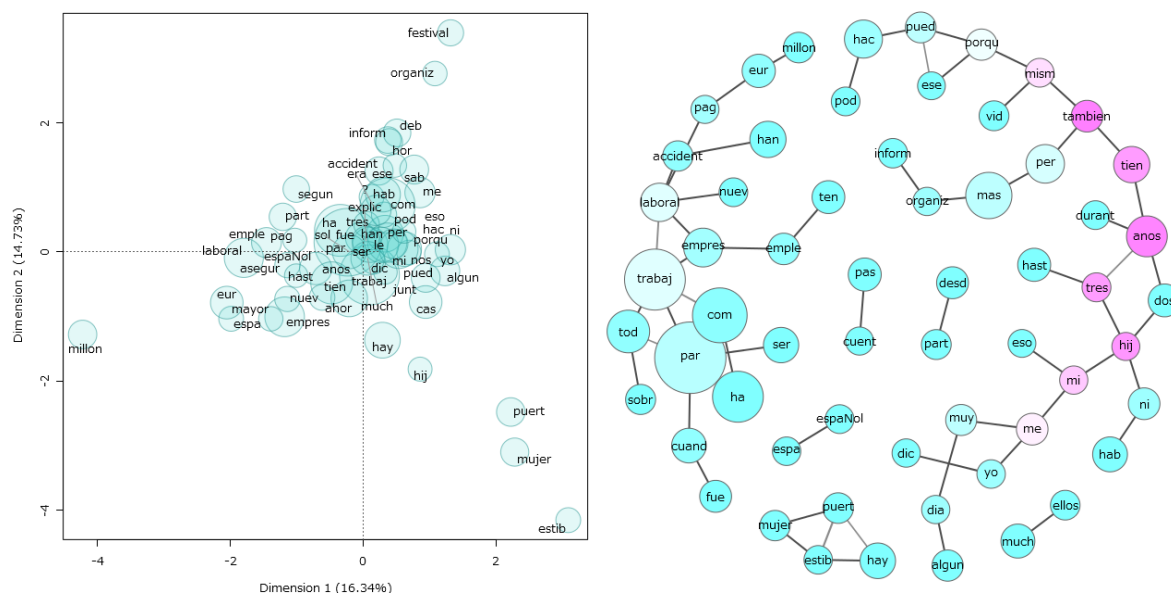
Source: Authors' own creation.

In *Elespañol.com* there is no great scattering of terms in its correspondence analysis, as shown in Figure 10. Most words are quite grouped together in the two-dimensional figure, which indicates that there is a certain pattern that links the texts of the pieces of this newspaper. The analysis of the co-occurrence network shows a main cluster that has no relation with the topic under analysis. On the



other hand, the group of most frequently used words includes worker and work, linked to accident, company-entrepreneur, million and euros. There is no greater relationship to accidents than the economic relationship.

**Figure 10: Analysis of correspondence and co-occurrence of *Elespañol.com*.**



Source: Authors' own creation

#### 4. Conclusions

The study of the treatment of occupational or work accidents and their prevention in the main born-digital newspapers in Spain has revealed a rather similar treatment in terms of what the print press has manifested over time, focusing on the fatal nature of the accident and the accident as a workplace event. The pieces are motivated by accidents, which makes the newspaper to treat these events mainly from an informative point of view, but also to sometimes address it in opinion articles. This is a topic which, being an important human, social and economic aspect for society, appears differently in the newspapers under analysis, in their media agenda and their treatment.

Of the five newspapers analysed over a period of two years and seven months, *Eldiario.es* is the digital newspaper that published the largest number of pieces about occupational accidents as well as opinion articles about the subject. *Eldiario.es* and *Publico.es* were the newspapers that used unions as their main source of information (in fact, *Eldiario.es* only quotes unions), whereas the rest of the newspapers used public authorities as their main source. Moreover, these two newspapers are the only ones that published pieces about court rulings relating to occupational accidents as examples for social knowledge.

A news report about a work accident is the preferred piece in all the newspapers except for *Publico.es*, which publishes more statistical facts about occupational accidents to show their evolution and



compare it with the situation in surrounding countries. *Okdiario.com* is the online newspaper that publishes the largest number of news pieces about work accidents, but its level is very similar to that of *Eldiario.es*. With regards to journalistic genres, the news or information report is the predominant, except in *Publico.es* which usually use feature articles. Only *Eldiario.es* and, to a lesser extent, *Elespañol.com* address work accidents in the opinion genre.

The framing method was based on text mining techniques and the combination of three statistical analyses: hierarchical cluster analysis, correspondences analysis, and co-occurrence networks. Each of these techniques reveals a series of aspects and complement each other to give an overall view of the treatment performed by each of the newspapers. It has been verified that *Publico.es* and *Elconfidencial.com* have not exhibited a line that follows a pattern of exposure throughout their pieces, while the rest exhibit a clearer line throughout the pieces.

*Eldiario.es* and *Publico.es* add to the news about work accident the denunciation of the working conditions of workers, either by giving unions a channel to voice their opinion, or by sharing their own opinion. These papers refer to cases and sectors such as construction, asbestos and *in itinere* accidents as problems. Risk prevention appears as something lateral to the main event, as a necessity, but without becoming a relevant aspect.

In *Elespañol.com*, *Elconfidencial.com* and *Okdiario.com* work accident is not related to the denunciation of working conditions, but is related to the economic aspect. Only *Okdiario.com* makes a clear reference to the prevention of risks and the need for planning but, as the other newspapers, as something lateral to the main frame of the occupational accident.

Therefore, and continuing with the initial hypothesis, it has been possible to confirm that the incidence of work accidents appears scarcely in the agenda of digital media, and mainly as part of a reactive process. The prevention of occupational hazards is not part of the main news frame and is only a derivation of the main event of the accident or the alarming statistical data. There are differences between the editorial lines of each newspaper. In fact, two of them have focused on denouncing the events and on giving voice to trade unions, while the other three have focused more on the economic aspects. As predicted in the study on web and print media carried out by Spain's National Institute for Safety and Hygiene at Work (INSHT), digital media focus on the fatal nature of the accident, while the denunciation and treatment as well as the use of sources and genres is very diverse across the different newspaper, which is an aspect that was analysed in a global way by the aforementioned study, which compared web and print newspapers without any specific qualitative techniques.

The use of the statistical methods based on text mining has shown that they are good options to identify the main frames in the texts, although they need further analysis to filter the results and only retain the coherent ones. It has been possible to verify that the method works better the greater the amount of text analysed and the greater the number of pieces analysed, because there are less results of clusters that make no sense, which is an issue that emerged in this study in the newspapers with fewer pieces. However, the global analysis of the use of this method allows us to recommend it to be used more commonly in the future for large volumes of information. It may present certain limitations with the use of certain polysemic words and non-coherent relationships, but its degree of analysis is becoming more and more advanced.

\* This article has been partially financed by UNIR Research (<http://research.unir.net>), the International University of La Rioja (UNIR, <http://www.unir.net>), within the framework of its Institutional Research Plan 2016-2018, and the research group “Liquid work and emerging risks in the information society”.

Dates:

-Start of research: August 2017

-End of research: November 2017

## 5. References

- I Álvarez-López (2004). “Los usos estratégicos del *framing*: (la campaña electoral municipal de Madrid 2003)” (PhD thesis). Facultad de Ciencias de la Información, Universidad Complutense de Madrid, Spain.
- J Álvarez-Gálvez, JF Plaza, J.A. Muñiz, J Lozano-Delmar (2014). “Aplicación de técnicas de minería de textos al frame analysis: identificando el encuadre textual de la inmigración en la prensa”. *Estudios sobre el Mensaje Periodístico*, 2, pp. 919-932.
- S Arce-García (2017). “La Prevención de Riesgos Laborales y la accidentalidad laboral en la prensa española: representación y cobertura a partir de la Ley 31/1995 de Prevención de Riesgos Laborales (1994-2014)” (PhD thesis). Facultad de Humanidades y Comunicación, Universidad de Burgos, Spain.
- A Ardèvol-Abreu (2015). “Framing o teoría del encuadre en comunicación. Orígenes, desarrollo y panorama actual en España”. *Revista Latina de Comunicación Social*, 70, pp. 423-450.
- A Asfaw, R Pana-Cryan, R Rosa (2011). “The business cycle and the incidence of workplace injuries: Evidence from the U.S.A.”. *Journal of Safety Research*, 42, pp. 1-8.
- S Aureli (2017). “A comparison of content analysis usage and text mining in CSR corporate disclosure”. *The International Journal of Digital Accounting Research*, 17, pp. 1-32.
- V Balakrishnan, N Humaidi, E Lloyd-Yemoh (2016). “Improving document relevancy using integrated language modeling techniques”. *Malaysian Journal of Computer Science*, 1, pp. 45-55.
- M Baojun, Y Hua, W Ye (2017). “Exploring performance of clustering methods on document sentiment analysis”. *Journal of Information Science*, 43, 1, pp. 54-74.
- MA Berjón, M Urquijo, A Fraile (2012). “La salud de los trabajadores en la prensa: del papel a internet”. Centro Nacional de Verificación de Maquinaria. Instituto Nacional de Seguridad e Higiene en el Trabajo. Available at <http://www.oect.es/Observatorio/Prensa/Informe%20sobre%20Salud%20Laboral%20en%20la%20prensa%20diaria%202010-2011.pdf>

I Crespo, C Moreno, I Delgado (2003). “El estudio de las campañas electorales”. In Crespo, Ismael (Ed.) *Partidos, medios de comunicación y electores. Los efectos de la campaña electoral de 2000 en España*, pp. 13-24. Buenos Aires: Planeta.

M Contreras-Barrera (2016). “Minería de texto en la clasificación de documentos digitales”. *Biblios: Journal of Librarianship and Information Science*, 64, pp. 33-43.

B Fernández-Muñiz, JM Montes-Peón, CJ Vázquez-Ordás (2016). “Occupational accidents and the economic cycle in Spain 1994–2014”. *Safety Science*. Available at <http://dx.doi.org/10.1016/j.ssci.2016.02.029>

L Fu-ren, D Hao, L Dachi (2016) “Automatic Content Analysis of Media Framing by Text Mining Techniques”. *49th International Conference on System Sciences (HICSS)*, Hawaii.

A Galán-García (2008). “*La prevención de los accidentes del trabajo. (Palancar y Pérez Botija, 1933): facsimil*” (Edición facsimil de la colección Ptha, introduction by Agustín Galán García, Antonio Fernández García and Carlos Álvarez Jiménez). Sevilla: Consejería de Empleo de la Junta de Andalucía and Fundación Mapfre.

SI Hajeer, RM Ismail, NL Badr, MF Tolba (2017). “A New Stemming Algorithm for Efficient Information Retrieval Systems and Web Search Engines”. In: A Hassanien, MM Mostafa, A Manaf, M Zamani, R Ahmad, J Kacprzyk (eds) *Multimedia Forensics and Security*. Intelligent Systems Reference Library, 115. Cham: Springer.

JJ Igartua, ML Humanes (2004). “Imágenes de Latinoamérica en la prensa española. Una aproximación empírica desde la Teoría del Encuadre”. *Comunicación y Sociedad*, 17, 1, pp. 47-75.

MC Justicia de la Torre (2017), *Nuevas técnicas de minería de textos: Aplicaciones* (PhD thesis). E.T.S. de Ingeniería Informática y de Telecomunicación, Universidad de Granada, Spain.

J Kayser (1979). *El diario francés*. Barcelona: ATE.

K Kettunen, T Kunttu and K Järvelin (2005). “To stem or lemmatize a highly inflectional language in a probabilistic IR environment?”. *Journal of Documentation*, 61, 4, pp. 476-496.

Ministerio de Empleo y Seguridad Social (2017). “Anuario de Estadísticas del Ministerio”. Available at <http://www.empleo.gob.es/es/estadisticas/index.htm>

C Narocki (1999). “Si ‘La prevención es rentable’ ¿Por qué no lo han descubierto los empresarios? Una revisión de propuestas para políticas en salud laboral”. *Cuadernos de Relaciones Laborales*, 14, pp. 101-133.

Redacción PRnoticias (24 November, 2016). “Comscore: ¿Cuáles fueron los diarios nativos más leídos de octubre?”. *PRnoticias.com*. Available at <http://prnoticias.com/internet-y-redes-sociales/diarios-digitales/20158237-comscore-diarios-nativos-mas-leidos-de-octubre>

N Remo-Díez (2011). “Costes sociales de siniestralidad laboral (2000-2007)”. *Pecunia. Revista de la Facultad de Ciencias Económicas y Empresariales*, Monográfico, pp. 213-231.

J Singh, V Gupta (2017). “A systematic review of text stemming techniques”. *Artificial Intelligence Review*, 48, 2, pp. 157-217.

X Yan, P Minnhagen (2018). “The dependence of frequency distributions on multiple meanings of words, codes and signs”. *Physica A: Statistical Mechanics and its Applications*, 490, pp. 554-564.

JH Ward (1963). “Hierarchical grouping to optimize an objective function”. *Journal of the American Statistical Association*, 58, pp. 236-244.

---

### How to cite this article in bibliographies / References

S Arce García, N Orviz Martínez, T Cuervo Carabel, I Fernández Suárez (2017): “Occupational accidents and their prevention in Spanish digital press”. *Revista Latina de Comunicación Social*, 72, pp. 1.608 to 1.625.

<http://www.revistalatinacs.org/072paper/1237/87en.html>

DOI: [10.4185/RLCS-2017-1237-87en](https://doi.org/10.4185/RLCS-2017-1237-87en)

Article received on 28 July 2017. Accepted on 11 December.  
Published on 17 December 2017.