Annotated Corpus for Citation Context Analysis

M. Hernández-Álvarez, José Gómez Soriano and Patricio Martínez-Barco

***Abstract*—In this paper, we present a corpus composed of 85 scientific articles annotated with 2092 citations analyzed using context analysis. We obtained a high Inter-annotator agreement; therefore, we assure reliability and reproducibility of the annotation performed by three coders in an independent way. We applied this corpus to classify citations according to qualitative criteria using a medium granularity categorization scheme enriched by annotated keywords and labels to obtain high granularity. The annotation schema handle three dimensions: PURPOSE: POLARITY: ASPECTS. Citation purpose define functions classification: use, critique, comparison and background with more specific classes stablished using keywords: Based on, Supply; Useful; Contrast; Acknowledge, Corroboration, Debate; Weakness and Hedges. Citation aspects complement the citation characterization: concept, method, data, tool, task, among others. Polarity has three levels: Positive, Negative and Neutral. We developed the schema and annotated the corpus focusing in applications for citation influence assessment, but we suggest that applications as summary generation and information retrieval also could use this annotated corpus because of the organization of the scheme in clearly defined general dimensions.**

***Index Terms*— Corpus, annotation, methodology, machine-learning, function, polarity, aspects, schema, keywords, labels, classification.**

I. INTRODUCTION

IT is necessary to overcome the absence of a common framework to facilitate research progress in collaborative conditions for citation context analysis. This framework should include a standard annotation scheme, and an annotated corpus according to such scheme. In fact, [1] suggested that the biggest problem facing researchers in this field is that there is not a public available annotated corpus that responds to a medium or high granularity scheme that can be used on a shared basis for scholars. The few annotated corpus available present some of the following problems: different ad hoc classification schemes are developed for each application; corpus are not publicly available for shared work; or, they are not presented in a standard format that other researchers could understand and use. Moreover, most of the previous citation work do not take into account citation context but only the sentence that contains the citation, method that results in loss of information that difficult achieving better classification results [2].

Different annotation approaches present diverse levels of granularity in citation function definitions. These schemes define from three to 35 different classes. Less granularity often refers to polarity (positive, negative, or neutral/objective). Schemes that are more complete correspond to diverse approaches and applications.

In [3], they categorize annotation schemes in two classes according if they have acceptable results using manual or automatic methods. In that study, we observe that manual classification schemes have medium granularity, while automatic processed schemes have low granularity. Annotated corpora with medium or high granularity provide valuable information indispensable to citation context analysis, but its annotation is a complex task, even for human coders, because even people have problems to achieve a good Inter-annotator agreement. Of course, challenges for automatic annotation are even greater [4]. The schemes with medium or high granularity need to be manually labeled by their authors; because attempting automatic labeling of this kind of corpora until now generates poor and not reliable results [1]. Even manual labeling without an adequate methodology results in a poor Inter-annotator agreement [5].

We could not find a classification scheme for citation function that combines a sufficient granularity with a simple structure, in a way that allows it to be useful in Citation Context Analysis, also having the necessary clarity to yield good Inter-annotator agreement; index that is indispensable to assure reproducibility and reliability.

We had three objectives to fulfill in the present study. First, to define a simple but complete structure scheme with enough information about purpose which is defined as aim and intention of the reference; and, citation polarity defined as author's disposition towards a reference that could be favorable or positive, unfavorable or negative and neutral [6]. Second, to annotate a corpus using this scheme obtaining a good Inter-annotator agreement, and make it available for collaborative work in the University of Alicante digital repository[1] and in the LRE map. Third, to apply in the previously mentioned citation corpus a machine-learning algorithm to classify automatically function and polarity with an acceptable outcome. In further work, we intent to identify influence levels of the citations applying in the developed corpus a machine-learning algorithm taking as inputs: function, polarity, and features related to position of the references.

M. Hernández-Álvarez, *Escuela Politécnica Nacional, Facultad de Ingeniería de Sistemas, Quito, Ecuador, myriam.hernandez@epn.edu.ec*

José Gómez Soriano, *Universidad de Alicante, Departamento de Lenguajes Informáticos, Alicante, España, jmgomez@dlsi.ua.es*

Patricio Martínez-Barco, *Universidad de Alicante, Departamento de Lenguajes Informáticos, Alicante, España, patricio@dlsi.ua.es*

[1] http://hdl.handle.net/10045/47416

II. Description of the Classification Scheme

As mentioned, we designed a classification scheme in order to maintain a simple structure with two dimensions: function that is associated to purpose, and polarity related to the disposition of the citing author towards the cited paper (Figure 1).

Function defines purpose; therefore, they have to do with categories such as use, comparison, critique and background. In some of these categories, we have classes that are more specific. USE: The functions *Based on, Supply* correspond to citation content that the citing author use in the same paper. Detecting related aspects separate this grouped class. *Based on* have to do with concept, methods and similar aspects. *Supply* has aspects such as tools, data, task and so on. The function *Useful* corresponds to a citation mentioned as used in other work, but that the citing paper does not apply.

COMPARISON: The function *Contrast* performs a comparison between aspects of different studies with positive, negative and neutral outcome. Frequently positive outcome results from a comparison with citing author's work.

CRITIQUE: This type of purpose corresponds to the functions *Weakness* and *Hedges*. *Weakness* is a direct criticism, *Hedges* is a concealed critique as defined by Hyland (1998).

BACKGROUND: This type of purpose relates with work that the citing paper nor other mentioned studies applied. It corresponds to functions *Acknowledge*, *Corroboration* and *Debate*. These grouped functions are separated using aspects. *Acknowledge* is a simply recognition of previous work. In *Corroboration*, there are aspects that determine agreement with the cited paper. *Debate* involves aspects that express difference of opinion with some of the content of a citation.

Polarity could be Positive, Negative and Neutral according to a favorable, unfavorable or neutral disposition from the author of the citing paper. Polarity definition relates to sentiment analysis.

We combined the two-dimension structure PURPOSE: POLARITY, with keywords and labels that indicate citation aspects: concept, method, data, tool, task, etc.; and positive, negative or neutral features. This more complete combination PURPOSE: POLARITY: ASPECTS yields high granularity, comparable with exhaustive ontologies as CiTO[2]. In [4], it is noticed that ontologies like CiTO present difficulty for annotation and obtain a low Inter-annotator agreement due to their complexity. In contrast, our proposed scheme facilitates understanding and application in the annotation process. The keywords and labels work both ways: to clarify function and polarity for the annotators, and later, they will serve as inputs for the automatic classification of function and polarity of the corpus.

Function	Description
Based on, Supply	Citing paper uses work from the citation. *Based on* refers to aspects such as concept, method and similar. Aspects of Supply function are data, tool, task, etc.
Useful	Citing paper does not use work from the citation, but it mentions citation as used in other studies. Aspects of this function are concept and method, but also data, tool, task, etc.
Acknowledge, Corroboration, Debate	Citation is mentioned as background to recognize prior work. Aspects separate the grouped functions. Paper could be mentioned just in passing (*Acknowledge*); to agree with cited paper (*Corroboration*); or to discuss cited paper (*Debate*). Citing paper does not use cited work. Other paper mentioned in citing paper does not use cited work.
Contrast	Citation is compared to citing paper or other work. Result can be a criterion positive, negative or neutral.
Weakness	Citing paper notes an error or weakness from cited paper.
Hedges	Citing paper uses careful language to disguise a criticism directed to the reference.

Table 1: Function classification scheme

Figure 1 shows classification dimensions, while Table 1 presents the function classification scheme.
Results for Inter-annotator agreement will demonstrate that our scheme is easy to apply. Annotators are able to take advantage of all possibilities of classification, because they need to understand and remember only six functions clearly defined and three levels for polarity; as opposed to what happen with complex ontologies as CiTO, where coders have to apply 92 object properties.

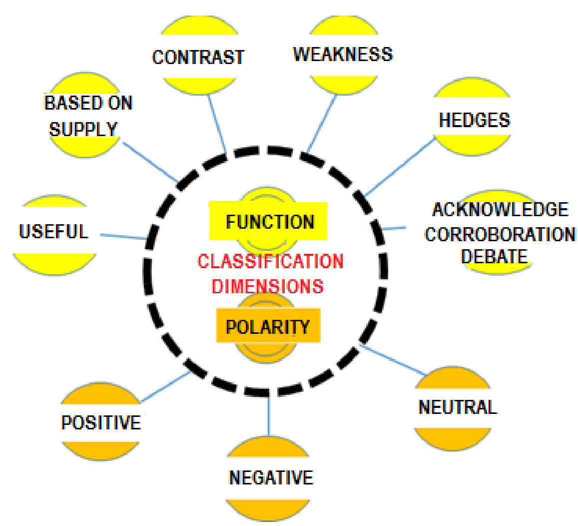

Figure 1: Function and polarity classification levels.

[2] http://purl.org/spar/cito

III. Dataset and Methodology

We applied the proposed scheme in a citation corpus composed by 85 articles taken randomly from ACL Anthology[3] with 2092 citations. We developed a program for converting text to XML, labeling paper titles, authors, sections, paragraphs and citations. After this initial pre-processing, we annotate citation function and polarity according to the suggested scheme using a methodology that includes a step of pre-annotation in which keywords and semantic tags are marked to clarify and standardize an internal representation that a coder or annotator creates about citation context. Using this method, the mental model is more likely to coincide with the ones produced for other coders, and consequently we obtain a good rate of Inter-annotator agreement in function and polarity classification. Experimentally we observed that with this pre-annotation step, we dramatically improve the agreement among annotators, which is indispensable to validate reliability and reproducibility of the annotation scheme.

Reliability and reproducibility of a classification scheme show whether it is possible to generalize results obtained in the annotation test to the complete process, in which probably are going to participate new annotators and not only the ones that codify the sample [8].

According to [9], annotation reliability and reproducibility is achieved if annotation process comply three conditions: a clear scheme with detailed instructions, specific criteria to choose annotators; and, the process must have at least three annotators working in an independent way. In our experiment, we fulfilled with these three requirements. We proposed a guide with a clear scheme, very detailed and with enough application examples; annotators are familiar with computational linguistics and with our guide, they revised the scheme carefully; and, we had three annotators working separately.

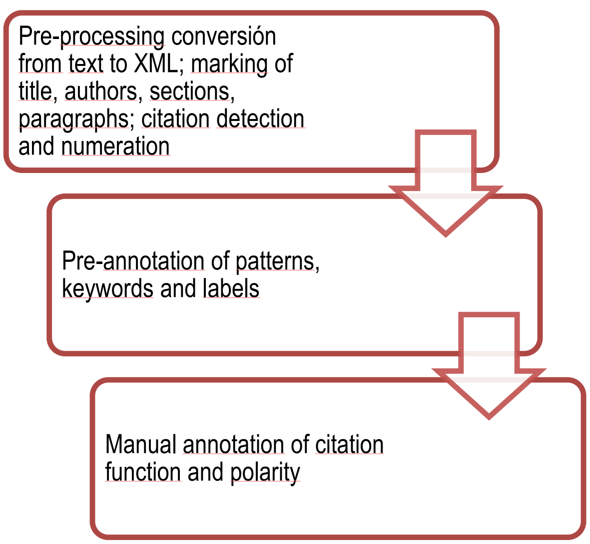

Figure 2: Corpus annotation process

Annotators chose keywords and labels from a list that corresponds to the most used words and phrases for each function and polarity classification; during the annotation process, we created new entries to this list as necessary. Annotators recognized relevant citation context inside a paragraph in which a citation is located. The keywords and labels list was refined while annotating the corpus and was part of the annotation guide. Figure 2 shows the corpus annotation process.

Illustrative keywords associated to polarity are “robustly” for *Positive*; “however” mostly for *Negative*; “previous work” for *Neutral*. Examples for keywords related to function: “build on” for *Based on, Supply*; “available” for *Useful*; “approach is not very satisfactory” for *Weakness*; “similarly” for *Contrast*; “another possibility for” for *Acknowledge, Corroboration, Debate.* Examples for labels associated with aspects of the citation function are “cited work”, “author”, “method”, “theory”, “task”, “tool”, “result”, “feature”, “positive feature”, and “negative feature”. Annotators can take these words or sequences of words from a specialized lexicon, but for our experiments, we defined these keywords and labels during the design of the coded corpus and through the course of the annotation process. In later experiments, we plan to annotate automatically keywords and labels, detecting those using bag_of_words and n-gram techniques from the lexicon we developed in the manual annotation.

For instance, if we have an original citation sentence: “Our classifier is built on the detailed previous work by Dong and Schäfer, 2011”. Resulting XML with annotation will be “<author>Our</author> <tool>classifier</tool> <kw>is built on </kw> the <posfeature>detailed</posfeature> previous work by <cite id=’citation_number_ identification’ function=’based on, supply’ polarity=’pos’>Dong and Schäfer, 2011</cited>”. The pattern is “AUTHOR TOOL *is built on* POSFEATURE CITE”, the different features of this pattern will be the input for the classification both manual and automatic. In this example, the classification is *Supply, Positive*. We improve Inter-annotator agreement marking first keywords and labels, but we also used these patterns to improve the granularity of the corpus in combination with function and polarity to disaggregate grouped functions and to define citation aspects. In this example, we classified the citation as *Supply* because it refers to a tool used by the author, and it is *Positive* for the kind of feature associated to it. Keywords were important to clarify the classification. To illustrate the role of the keyword, if the aspect were a method and not a tool, the classification for function would be *Based on.*

A special treatment is required for the recognition of the *Hedges* function. For instance, the classification should recognize the combination of a positive feature followed by a negative one.

For example if we have the quote: “The only recent work on citation sentiment detection using a relatively *large corpus* is by Athar (2011). However, this work *does not handle* citation context”.

In this example, the author intention is to make a disguised criticism softened with a prior recognition of a positive characteristic. The result is a *Hedges* function because the real intention is criticism (Hyland, 1998). Here, the positive feature is “large corpus”; the negative feature is “doesn´t handle”.

[3] Released Dec. 2013 http://clair.eecs.umich.edu/aan/index.php

Another case for the detection of the *Hedges* function involves not expressing categorically a negative expression (Hyland, 1998).

For example in the citation: "The first experiments in Argumentative Zoning used Naïve Bayes (NB) classifiers Kupiec et al., 1995; Teufel, (1999), which assume conditional independence of the features. However, this assumption ***is rarely true*** for the types of rich feature representations we want to use for most NLP tasks".

The negative opinion is softening by the words "rarely true" to avoid making a more categorical affirmation but the intention is again criticism, and therefore the function is *Hedges.*

Our scheme is simple but powerful because of the three dimensions used for classification: function, polarity and annotated patterns formed by keywords and labels: FUNCTION: POLARITY: ASPECTS. The combination of the three criteria produce high granularity without a complex structure.

In Figure 3, we present an example of the high granularity achieved using these three dimensions. A citation function classified as useful can refer to different aspects as tool, data, task, method; also, it can be mentioned with positive, negative or neutral features that facilitate definition of polarity, also it can be defined with its name. With all these elements, we obtained a complete citation description.

For instance, a citation could be referred as a specific tool, which is reported as useful because it is applied in other study and not in citing paper, and have positive reports that are detected by a positive feature annotated as a label. In this case, the function is *Useful*; polarity is *Positive*; and its aspect is that it is a tool. In general terms, the aspect is a third very important dimension that will specify if the citation refers to a tool, data, task or method or other; besides it will tell if it has positive, negative or neutral features which will define polarity.

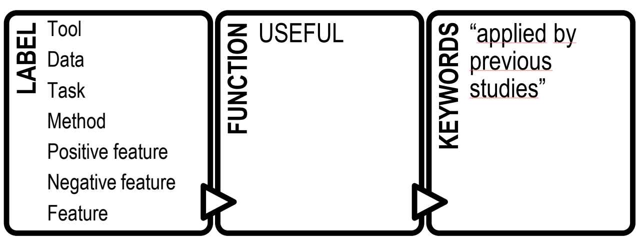

Figure 3: Improved granularity using labels and keywords.

IV. Annotation Results: Inter-Annotator Agreement

We validate Inter-annotator agreement and show results for function in Table 2, and for polarity in Table 3. We can see that the values of Fleiss' Kappa are as high as 0.862 for function and 0.912 for polarity. These values correspond to an almost perfect agreement in accordance to the scale of [10].

Using keywords and labels, we obtain a considerable improvement, because without this step, with the same annotators, there were low results for this index: 0.386 and 0.259 for function and polarity respectively, because of the difficulty to form coincident mental models among different coders.

The pre-annotation step allows forming these matching mental models and in addition, it provides information to feed as input to classifiers. Therefore keywords and labels added in the pre-annotation step, help both manual and automatic classification. Other studies [5] showed that it is very difficult to obtain a Kappa value for Inter-annotator agreement higher than 0.75 for a scheme with more than three classes.

Fleiss	**Krippendorff**	**Pairwise avg.**
A_obs=0.911	D_obs = 0.089	% agr = 91.1
A_esp=0.354	D_esp = 0.648	Kappa=0.862
Kappa=0.862	Alpha = 0.862	

Table 2: Inter-annotator agreement for function annotation

Fleiss	**Krippendorff**	**Pairwise avg.**
A_obs = 0.98	D_obs = 0.02	% agr = 98
A_exp=0.776	D_exp = 0.225	Kappa=0.913
Kappa=0.912	Alpha = 0.912	

Table 3: Inter-annotator agreement for polarity annotation

Regarding the context length for classification, in the annotation results, we noticed that the context length chosen by coders largely corresponds to just one statement: the one with the citation. With less frequency appears a length context of two or three sentences. It is probable that the context should not include more than three sentences to cover all the necessary information about the reference.

Context length	**Number of occurrences**
One sentence	1502
Two sentences	377
Three sentences	127
Four sentences	56
Five or more sentences	30

Table 4: Citation context length chose for annotators

Table 4 shows the number of sentences chose for annotators for citation context. In 95.6% of cases, the context length refers to one, two and three sentences including the one that contains the citation.

We evaluate performance indexes for function and polarity classification that uses the annotated keywords and labels as inputs. Results rated high for F-Measure, which demonstrate suitability of the chosen features for those classifications. We chose algorithms after the recommendations of our initial study [1]. In our results, SVM with SMO training has the best values; we show our experiment outcomes for function classification in Table 5; and, for polarity classification in Table 6. Used relation between tests vs. training datasets was 10% - 90%.

In Table 7, we present the relationship between function and its polarity.

Class	**F-Measure**
Useful	0.89
Weakness	0.94
Acknowledge, Corroboration, Debate	0.92
Based on, Supply	0.86
Contrast	0.89
Hedges	0.67
Weighted Avg.	**0.896**

Table 5: Function classification performance with SVM - SMO algorithm.

Previous studies presented results not as good for similar o less granularity. In [11], they used the model of [12], with four facets and their F1 scores varied from 0.68 for discriminating idea from a tool, to 0.51 for conformational / negational facets (similar to polarity), with scores between this minimum and this maximum for the other two classes. In [13], they classified fundamental idea /technical basis /comparison with F1 values of 0.66. In [5], they achieved F1 of 0.71 but just for polarity classification. In [14], they classified two function: corroborate and contrast with a recall of 0.83 for the first and 0.67 for the other. In [15], it was implemented a citation-classification algorithm through pattern matching, with a highest Recall of 0.49. In [16], they classified 10 citation functions to discover only 6 of them and a very variable F1 scores that go from 0.05 to 0.802 with an average of 0,49. In [17], they used a six-function scheme to obtain an average F macro of 0.58.

Class	**F-Measure**
Positive	0.94
Negative	1
Neutral	0.96
Weighted Avg.	**0.957**

Table 6: Polarity classification performance with SVM – SMO algorithm.

	Positive	Neutral	Negative
Useful	226	479	0
Weakness	0	0	123
Acknowledge, Corroboration, Debate	62	708	12
Based on, Supply	280	57	0
Contrast	14	69	25
Hedges	0	0	37

Table 7: Relationship between function and polarity classification

We noted that certain functions do not have results for some polarities. *Useful* do not appear as Negative; *Weakness* and *Hedges* are always with Negative polarity; and, Based on, Supply do not have occurrences with Negative polarity. All of that make sense from function and polarity definitions.

V. CONCLUSIONS

The developed scheme are consistent to citation purpose and citing author's disposition towards references. In further work, we intent to use this scheme and corpus for citation analysis to obtain influence levels of a citation in a paper. With this scheme, we annotated 85 ACL articles obtained randomly with 2092 citations. We suggest that this scheme and developed corpus could also be applied for summary generation and information retrieval, because of the clear organization of the scheme in general dimensions: PURPOSE: POLARITY: ASPECTS.

Annotation results are high with an Inter-Annotator agreement of 0.862 and 0.912 for citation function and polarity classification respectively. This kind of results we could not have obtained without our annotation methodology that has a pre-process of labeling patterns formed by keywords and labels that clarify the scheme dimensions. Later we also use these patterns as input features for the machine-learning algorithm for function and polarity classification.

We use the annotated corpus to perform automatic classification of citation function and polarity and we obtained an F1 weighted average of 0.896, which are higher than results in other studies. However, it is important to notice that annotated data in our corpus is relevant and delivers a sufficient amount of information to feed classifiers to yield optimal results; marked keywords and labels define what we called Aspects. For some other corpus, automatic annotation generally is performed just in a lexical and / or syntactic level and have lots of not pertinent information (noise). When other studies use these noisy annotations, they achieve low algorithm performance.

In contrast, we manually annotated our corpus, using an annotation scheme with relevant features organized according the scheme, that take into account citation context (inside a paragraph). These criteria form the basis for building a good model for automatic citation classification. Aspects annotated in a variable context length, give a great amount of information and allow achieving satisfactory results for function and polarity citation classification. According our results optimal context length could be from one to three sentences around a citation.

Classification results in our experiments confirm the validity of our classification scheme. If an application requires a trusty classification, it is important to define relevant features that should be included in any annotation effort, manual or automatic; they give information that is indispensable for good results.

In summary, in the present work, we intent to contribute with the following:

- A proposed annotation scheme simple in its structure, but with high granularity thanks to the combination of information from function, polarity, keywords and semantic labels, organized in three dimensions: PURPOSE: POLARITY: ASPECTS.

- The annotation methodology, particularly regarding to the pre-annotation process to detect keywords and labels that are useful to create mental models in the annotators. These characteristics also serve as input features in classification algorithms. Therefore, we used keywords and labels to improve Inter-annotator agreement, but also we applied those to increase the granularity of the corpus.
- An annotated corpus with a sufficient size that contains those relevant features and is accessible for collaborative work. The XML files for our annotated corpus is available in the University of Alicante digital repository [18].
- The experimental finding that the significant context around a citation usually takes no more than three sentences including the one with the mention.

As future work, we will continue populating the corpus with new annotated documents and new collaborative tools for manual annotation.

There are controversies regarding counting approaches to measure citation impact, because they consider all citations as equal regardless of the purpose or the polarity with which they were mentioned. In [19], it was showed that incomplete, erroneous, or controversial papers have higher citation counts. Therefore, we plan to use the corpus to obtain citation influence in a paper using a machine-learning algorithm using as features the same dimensions: PURPOSE: POLARITY: ASPECTS, with additional information: citation position in an IMRAD paper structure, and frequency of the citation in the different sections of the paper. For this new challenge, we are labeling the training dataset with answers of authors of citing papers that will state influence of the works they cited. We are sending a survey with this request to the authors of the articles in our corpus and we are in the process of receiving and tabulating answers. We will use this information to measure precision in our influence classification.

Due to the reliability that is obtained in our manual corpus annotation, we suggest that, in the near future, the data continue to be annotated manually using our methodology. We state that it is necessary to improve current automatic annotation techniques marking relevant information for obtaining reliable results when applied to an annotation scheme with medium or high granularity.

Regarding to automatic annotation, as future work, we consider that our scheme and detected features determine a clearer path for the development of automatic annotation techniques, because we divide a complex task in ones that are more manageable. It would be easier for an automatic classifier to recognize characteristic patterns for each of our defined dimensions. From the lexicon created for this study, we intent to develop an automatic annotation process for marking keywords and labels using simple techniques as bag_of_words and n-gram detection.

REFERENCES

[1] Hernández Álvarez, M., & Gómez Soriano, J. (2015b). Survey about citation context analysis: Tasks, techniques, and resources. *Natural Language Engineering*. Cambridge University Press. Available on CJO 2015 doi: 10.1017/S1351324915000388

[2] Athar, A. (2014). Sentiment analysis of scientific citations. Technical Report, University of Cambridge. (UCAM-CL-TR-856).

[3] Mandya, A. A. (2012). Enhancing Citation Context based Information Services through Sentence Context Identification. Doctoral dissertation, University of Otago. Retrieved from: http://hdl.handle.net/10523/2520

[4] Ciancarini, P., Di Iorio, A., Nuzzolese, A. G., Peroni, S., & Vitali, F. (2014). Evaluating citation functions in CiTO: cognitive issues. In *The Semantic Web: Trends and Challenges* pp. 580-94. Springer International Publishing.

[5] Teufel, S., Siddharthan, A., & Tidhar, D. (2006, July). Automatic classification of citation function. In *Proceedings of the 2006 conference on empirical methods in natural language processing* (pp. 103-110). Association for Computational Linguistics.

[6] Hernández Álvarez, M., & Gómez Soriano, J. (2015a). Esquema de anotación para categorización de citas en bibliografía científica. *Procesamiento del Lenguaje Natural*, 54, 45-52.

[7] Hyland, K. 1998. *Hedging in Scientific Research Articles*, Vol. 54. Amsterdam: John Benjamins Publishing.

[8] Artstein, R., & Poesio, M. 2008. Inter-coder agreement for computational linguistics. *Computational Linguistics*, 34(4), 555-96.

[9] Krippendorff, Klaus. 2004. Reliability in content analysis: Some common misconceptions and recommendations. *Human Communication Research*, 30(3):411–33.

[10] Landis, J. R., & Koch, G. G. (1977). The measurement of observer agreement for categorical data. *Biometrics*, 159-74.

[11] Jochim, C., and Schütze, H. 2012. Towards a generic and flexible citation classifier based on a faceted classification scheme. In *Procedings of COLING'12* (pp. 1343–58).

[12] Moravcsik, M. J., & Murugesan, P. (1975). Some results on the function and quality of citations. *Social studies of science*, *5*(1), 86-92.

[13] Dong, C., and Schäfer, U. 2011. Ensemble-style self-training on citation classification. In *Proceedings of 5th International Joint Conference on Natural Language Processing*, pp. 623–31. Chiang Mai, Thailand: Asian Federation of Natural Language Processing.

[14] Meyers, A. 2013. Contrasting and corroborating citations in journal articles. In *Proceedings of the International Conference Recent Advances in Natural Language Processing RANLP 2013*, Hissar, Bulgaria, pp. 460–6.

[15] Iorio, A., Di, Nuzzolese, A. G., and Peroni, S. 2013. Towards the automatic identification of the nature of citations. In *SePublica*, Montpellier, France, pp. 63–74.

[16] Li, X., He, Y., Meyers, A., and Grishman, R. 2013. Towards fine-grained citation function classification. In *Proceedings of Recent Advances in Natural Language Processing*, Hissar, Bulgaria, pp. 402–7.

[17] Abu-Jbara, A., Ezra, J., and Radev, D. 2013. Purpose and polarity of citation: Towards NLP-based bibliometrics. In *Proceedings of the 2013 Conference of the North American Chapter of the Association for Computational Linguistics: Human Language Technologies*. ACL. Atlanta, GA. pp. 596–606.

[18] Concit – Corpus (2015), Universidad de Alicante digital repository. http://hdl.handle.net/10045/47416.

[19] Radicchi, F. 2012. In science "there is no bad publicity": Papers criticized in comments have high scientific impact. *Nature Scientific Reports* 2: 815.

Myriam Hernández-Álvarez received the Electronic and Telecomunications Engineering degree from Escuela Politécnica Nacional, Quito, Ecuador (1982); the Master of Science degree in Computer Science from Ohio University at Athens, Ohio, USA (1987); a Specialist degree in Business Management, Universidad Andina Simón Bolivar (1998); PhD Degree in Informatic Applications, Universidad de Alicante, Alicante, España (2015). Currently, she is doing research in the field of Natural Language Engineering and is Dean of the Systems Engineering Faculty of the Escuela Politécnica Nacional.

Patricio Martínez Barco, received his Master Degree in Computer Science Engineering by the Universidad de Alicante (1994) and his PhD in Computer Science (2001). He is member of the Language Processing and Information System Research Group (GPLSI) at the Universidad de Alicante; Member of the Natural Language Processing InterUniversity Group (Universidad Politécnica de Valencia and Universidad de Alicante) and Vicepresident of the Spanish Society for Natural Language Processing - SEPLN.

José Manuel Gómez Soriano, received his PhD degree by the Universidad Politécnica de Valencia in 2007. Currently he works as a researcher at the Universidad de Alicante, Alicante, Spain, as a member of the Natural Language Processing and Information System Group. Project: Intelligent, Interactive and Multilingual Text Mining based on Human Language Technologies.