IJCLA VOL. 1, NO. 1-2, JAN-DEC 2010, PP. 45-59
RECEIVED 28/11/09 ACCEPTED 16/01/10 FINAL 08/03/10

Exploiting Higher-level Semantic Information for the Opinion-oriented Summarization of Blogs

ALEXANDRA BALAHUR[1], MIJAIL KABADJOV[2], JOSEF STEINBERGER[2]

[1] *University of Alicante, Spain*
[2] *European Commission – Joint Research Centre, Italy*

ABSTRACT

Together with the growth of the Web 2.0, people have started more and more to communicate, share ideas and comment in blogs, social networks, forums and review sites. Within this context, new and suitable techniques must be developed for the automatic treatment of the large volume of subjective data, to appropriately summarize the arguments presented therein (e.g. as "in favor" and "against"). This article assesses the impact of exploiting higher-level semantic information such as named entities and IS-A relationships for the automatic summarization of positive and negative opinions in blog threads. We first run a sentiment analyzer (with and without topic detection) and subsequently a summarizer based on a framework drawing on Latent Semantic Analysis. Further on, we employ an annotated corpus and the standard ROUGE scorer to automatically evaluate our approach. We compare the results obtained using different system configurations and discuss the issues involved, proposing a suitable method for tackling this scenario.

Keywords: opinion mining, sentiment analysis, text summarization, social media.

1 Introduction

The recent growth in access to technology and the Internet, together with the development of the Web 2.0 (Social Web), has led to the birth of new and interesting social phenomena. On the one hand, the possibility to express opinion "by anyone, anywhere, on anything", in blogs, forums, review sites has made it possible for people all around the world to take better and more informed decisions at the time of buying products and contracting services. On the other hand, the companies and public persons are more informed on the impact they have on people, because the large amount of opinions expressed on them offers a direct and unbiased, global feedback. Moreover, people all over the world can express their opinion on the issues that affect their lives – events in politics, economics, the social sphere – or simply discuss on their hobbies and everyday lives. Thus, the past few years, due to the growing access to the Internet and the development of such Web 2.0 phenomena, have lead to the creation on the web of extensive quantities of subjective and opinionated data. Such information cannot be manually processed, although their analysis (discovery of opinions, their classification into positive and negative), could be useful to a high diversity of entities (potential customers, companies, public figures and institutions etc.), for a large variety of tasks (opinion analysis for marketing, sociological or political studies, decision support etc.). Therefore, automatic systems must be built, with the aim of processing the subjective data available and extracting the information that is relevant to the users.

For example, when a potential customer is interested in buying a new digital camera, they would like to know what others think about the features of the different models available on the market, within a price range, and whether others recommend the product or not. An automatic system assisting such a user would have to retrieve all the opinionated texts on the customer's products of interest, extract the product features and the opinions expressed on them, classify the opinions as positive or negative and present the user with percentages of positive and negative opinions on each of the product features. One step further could be that of summarizing the positive and negative opinions, so that the users can read for themselves the reasons for liking or disliking the product.

Another example involving the treatment of subjective data is that of a public person constantly monitoring his/her public image. Such a person would require the daily or weekly analysis of all the opinions expressed on them and their actions. An automatic system

implementing this task would have to gather all the opinions expressed on the person every day, analyze them to determine whether they are positive or negative and present the user with an overview of the general opinion (in percentages, organized depending on the opinion source, or in the form of an extractive summary).

Finally, an example of a system analyzing subjective data to respond to the needs of different users is one that is capable of extracting, from discussion threads, such as those present in blogs, the arguments "in favor" and "against" a topic, be it the economic crisis or cooking recipes. Such a system can extract the relevant opinions expressed on the topic and eliminating the redundant information, presenting the user with a clear list of arguments explaining the general view on the matter.

This article presents and compares different methods implemented with the aim of creating a system of the latter type. We show how the subjective content can be analyzed from the pure opinion and combined topic-opinion point of view and how the relevant parts can subsequently be summarized, based on the polarity of the opinions expressed. In what follows, Section 2 presents the related work and previous experiments in related tasks. Further on, Section 3 motivates the approaches proposed and indicates the contribution of this article to the task. In Section 4, we present the data we employ in our experiments and in Section 5, we depict the preliminary experiments conducted on it. Section 6 presents an in-depth description of the experiments performed and the results of the different evaluations. Finally, we conclude in Section 7, by discussing our findings and proposing the lines for future work.

2 Related Work

While the task of summarization has been tackled for a longer period of time within the field of Natural Language Processing (NLP), literature in sentiment analysis has only flourished in the past few years, due to the massive growth in the quantity of subjective data available on the web. Thus, whilst there is abundant literature on text summarization [1, 2, 3, 4, 5] and sentiment analysis [6, 7, 8, 9, 10], there is still limited work at the intersection of these two areas [11, 12, 13]. This is easily explainable by: a) the fact that both systems performing opinion mining, as well as those automatically summarizing must have a certain level of maturity, so that errors do not propagate along the processing

pipeline; b) the task of summarization within the opinion context may be different from the traditional view on text summarization [14].

The 2008 edition of the Text Analysis Conference (TAC 2008), organized by the US National Institute of Standards and Technology (NIST), contained a pilot task, within the summarization track – i.e. Summarization Opinion Pilot. Being a pilot task within the summarization track, most of the techniques employed by the participants were based on the already existing summarization systems. New characteristics were added to these systems to account for the assessment of opinions present in the text (sentiment, positive/negative sentiment, positive/negative opinion). Examples of such systems are: CLASSY [15]; CCNU [16]; LIPN [17]; IIITSum08 [18]. Other participants, outside de summarization track, focused more on the opinion mining part of the task, thus doing the retrieval and filtering based on polarity - DLSIUAES [19]- or on separating information rich clauses – italica [20]. The results of the competition showed that, on the one hand, systems concentrating on the summarization part lost on the opinion content, and, on the other hand, systems lacking proper summarization components lose as far as the linguistic quality of the results is concerned and introduce much noise due to not being able to filter out redundant or marginal information.

Zhou and Hovy [21] and [22] present approaches to summarizing threads in blogs and online discussions, but focusing on the factual content. They demonstrate why this type of summarization is more difficult than traditional summarization in newswire and model subtopics and topic drifts.

Recently, [12] propose an approach to summarize threads in blogs using a combination of an opinion mining and a summarization system. They analyze the output as far as linguistic quality is concerned, to assess the difficulty of the task in the context of blogs, demonstrating that the difficulty in performing opinion summarization of blog threads resides in the language used, the topic inconsistency and the high redundancy of information. [13] claim that topic detection is crucial to the summarization of blog threads, but no experiments are done in this sense.

[14] assess the difference between the traditional task of summarization and opinion summarization in blogs, showing that through the nature of blog texts and the high subjectivity they contain, opinion summarization differs to a large degree from the traditional task. They experiment with the hypothesis of whether, in this context, the intensity of polarity is a good summary indicator.

3 Motivation and Contribution

As demonstrated by the body of research that has tackled this issue, summarizing opinion is a difficult task, especially when pursued in the context of blogs.

Even if the behavior of bloggers has changed in the past few years, as shown by the Technorati "State of the blogosphere" reports[1] in 2008[2] and 2009[3], one of the main difficulties when addressing opinion expression in blogs is that it contains many references to outside sources, as well as "copy+paste"s from newspaper articles, photos, videos and other types of multimodal information that supports the argument that is made. While in 2006, Zhou and Hovy (2006) were writing that the predominance in blogs is given by the original blog message of the blog author, in 2009, we find that the vast majority of the thread body is given by comments written by other bloggers. This fact is supported by the Technorati report on the state of the blogosphere in 2009, where commenting in other blogs is found to be one of the strategies employed for attracting audience to one's blog, along with the tagging of content, regular updating of content and others.

Contrary to the general belief, blogs are mainly written by highly educated people and they can constitute a manner to consult expert opinion on different subjects. That is why, our first motivation in our experiments to search for and summarize opinions on different topics in blogs is given by the possibility blogs give to acquire useful and timely information.

Secondly, the research done so far in this area has not taken into consideration the use of methods to detect sentiment that is directly related to the topic. In the experiments we have performed, we detect sentences where the topic is mentioned, by using Latent Semantic Analysis.

Thirdly, most summarization systems do not take into consideration semantic information or include Named Entity variants and co-references. In our approach, also employed in the TAC 2009

[1] The Technorati reports on the state of the blogosphere have been published online since 2004, and are available at http://technorati.com/. They present statistics and overviews on the number of blogs, their topics, the social background and motivation of bloggers, as well as results of questionnaires enquiring on the behavior of bloggers.

[2] http://technorati.com/blogging/feature/state-of-the-blogosphere-2008/

[3] http://technorati.com/blogging/feature/state-of-the-blogosphere-2009/

summarization, we employ these methods and show how we can obtain better results through their use.

4 DATA

The data used in our experiments is described in [12] and it has also been used in [13] and [14]. It consists of 51 blog entries with their corresponding comments (threads) in English, summing up to a total of 1829 posts with 299.568 words. This corpus was selected, on the one hand, because it gives us the possibility to compare the results obtained with the ones reported in the related studies and, on the other hand, because it contains the annotations of the topics discussed in the posts and labeling of the topic-relevant sentences as far as source (the author of the text snippet), target (the topic it addresses), polarity (positive and negative) and intensity of the polarity (low, medium, high) are concerned. Although the threads are centered mostly on economy, science and technology, cooking, society and sport, their annotation contains a finer-grained identification of subtopics – e.g. the economic crisis, idols, VIPs and so on.

The gold standard for the summarization process is marked by the annotations on this corpus. We consider that the correct sentences that should appear in the final summaries (separately considering the positive and negative arguments on a topic) are the ones that are relevant for the topic, have the required polarity and score high on intensity.

5 PRELIMINARY EXPERIMENTS

Before reaching the present configuration of the system, we have performed several experiments on the presented blog data, as well as on quotations (reported speech)- shorter pieces of text representing a direct statement of opinion, from a source to a target. From these preliminary experiments, we could extract several useful conclusions, which influenced the final setting of the experiments presented.

5.1 Preliminary sentiment analysis approach

The first and easiest approach that we carried out was based on two processing phases: the first one identified the subjective sentences -

using the Subjectivity Indicators in [23] - and, in the second phase, the polarity of the sentences classified as subjective was computed as sum of the opinion words found in them - using different combinations of affect and opinion lexicons: MicroWordNet Opinion [24], SentiWordNet [25], WordNet Affect [26] and a list of in-house terms denominated the JRC List. In order to perform these two steps on the data, the blog threads were split into files containing the initial post and, individually, the comments given by other bloggers on this post and subsequently the posts were split into sentences using LingPipe[4]. The best results on the blog data presently used were obtained when a combination of all resources was employed, leading to a precision of classification for positive opinion of 0.67, with a recall of 0.22 and a precision of classification for negative opinion of 0.53, with a recall of 0.89. The low results were attributed mostly to the lack of topic determination; the analysis of the accuracy for sentence classification revealed that many of the sentences had been correctly classified from the opinion polarity point of view, but they were not on the topics identified in the blogs. The summarization process, based on Latent Semantic Analysis [27] had a performance, given by the ROUGE scores, of 0.21 and 0.22 (R_1 for positive and negative, respectively) and 0.05 and 0.09 (for R_2 and R_{SU4} for positive and negative, respectively).

5.2 Opinion classification around Named Entities

Filtering sentences according to their topic, when the latter is a wide concept, such as economics or politics, is not a trivial task. However, when the topic is a Named Entity – its mentions, under its name or title (e.g. Gordon Brown, mentioned as such, or as Gordon, or "the British prime-minister") – the task becomes easier. Thus, in a parallel experiment, we tested, under the same conditions, the possibility to classify opinion on different public persons, by assessing the context surrounding their mentions in newspaper quotations. The results of these experiments showed significant improvements over the previous results, with an accuracy of 83% in classifying opinion among positive, negative and neutral (objective), using a combination of MicroWNOp, the JRC List and the General Inquirer (Stone et al. 1966). We employ this same strategy in order to compute the opinion on the topic of interest, using the topic words discovered with LSA as anchors around which opinion words are sought.

[4] http://alias-i.com/lingpipe/

6 Experiments and evaluation

As seen in the preliminary experiments, the performance of the opinion summarization, as it was tackled so far (without taking into consideration the topic) was rather low. From the human evaluation of the obtained summaries, we could see that the sentiment analysis system classified the sentences correctly as far as opinion, polarity and intensity are concerned. However, many topic irrelevant sentences were introduced in the summaries, leaving aside the relevant ones. On the other hand, we could notice that in the experiments taking into consideration the presence of the opinion target and its co-references and computing the opinion polarity around the mentions of the target reaches a higher level of performance. Therefore, it became clear that a system performing opinion summarization in blogs must include a topic component.

6.1 Sentiment analysis system

In the first stage, we employ the same technique as in the preliminary approach, but using only the resources that best scored together (MicroWordNet Opion, JRC Lists and General Inquirer). We map each of these resources into four classes (of positive, negative, high positive and high negative, and assign each of the words in the classes a value, of 1, -1, 4 and -4, respectively. We score each of the blog sentences as sum of the values of the opinion words identified in it (Fig.1).

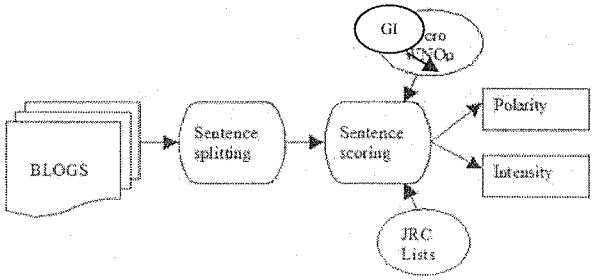

Fig. 1. Sentiment analysis system

In the second stage, we first filter out the sentences that are associated to the topic discussed, using LSA. Further on, we score the sentences identified as relating to the topic of the blog post, in the same manner as in the previous approach (Fig. 2).

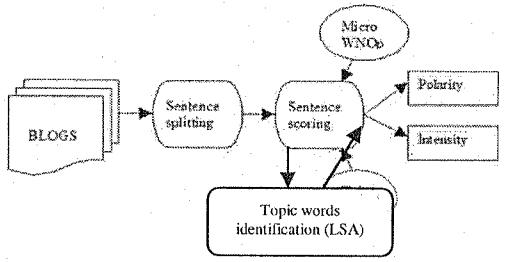

Fig. 2. Sentiment analysis system with topic words identification through LSA

Topic words identification using LSA. In order to filter for processing only the sentences containing opinions on the post topic, we first create a small corpus of blog posts on each of the topics included in our collection. These small corpora (30 posts for each of the five topics) are gathered using the search on topic words on http://www.blogniscient.com/. For each of these 5 corpora, we apply LSA, using the Infomap NLP Software[5]. Subsequently, we compute the 100 most associated words with 2 of the terms that are most associated with each of the 5 topics and the 100 most associated words with the topic word. For example, for the term "bank", which is associated to "economy", we obtain (the first 20 terms):

```
bank:1.000000;money:0.799950;pump:0.683452;
switched:0.682389;interest:0.674177;easing:0.661366
;
authorised:0.660222;coaster:0.656544;roller:0.65654
4;
maintained:0.656216;projected:0.656026;apf:0.655364
;
requirements:0.650757;tbills:0.650515;ordering:0.64
8081;
eligible:0.645723;ferguson's:0.644950;proportionall
y:0.63358;
integrate:0.625096;rates:0.624235
```

[5] http://infomap-nlp.sourceforge.net/

6.2 *Summarization system*

The summarization process is based on LSA, which is enriched with semantic information coming from two sources: the Medical Subject Headings (MeSH) taxonomy[6] and a Named Entity recognizer and disambiguator [28].

The LSA approach to summarization entails a two-fold process: firstly, a term-by-sentence matrix from the source is built and secondly, Singular Value Decomposition (SVD) is applied to the initial matrix. The decomposition is then used to select the most informative sentences. The enrichment of semantic information takes place during the step of building the term-by-sentence matrix. Full details of the approach can be found in [29].

6.3 *Evaluation*

We include the usual ROUGE metrics: R_1 is the maximum number of co-occurring unigrams, R_2 is the maximum number of co-occurring bigrams, R_{su4} is the skip bigram measure with the addition of unigrams as counting unit, and finally, R_L is the longest common subsequence measure (Lin, 2004). In the cases of the baseline systems we present the average *F1* score for the given metric and within parenthesis the 95% confidence intervals.

Table 1. Summarization performance.

System	R_1	R_2	R_{SU4}	R_L
Sent+BLSumm$_{neg}$	0.22 (0.18-0.26)	0.09 (0.06-0.11)	0.09 (0.06-0.11)	0.21 (0.17-0.24)
Sent+Summ$_{neg}$	**0.268**	0.087	0.087	**0.253**
Sent+BLSumm$_{neg}$	0.21 (0.17-0.26)	0.05 (0.02-0.09)	0.05 (0.02-0.09)	0.19 (0.16-0.23)
Sent+Summ$_{neg}$	**0.275**	0.076	0.076	**0.249**

[6] The MeSH thesaurus is prepared by the US National Library of Medicine for indexing, cataloguing, and searching for biomedical and health-related information and documents. Although, it was initially meant for biomedical and health-related documents, since it represents a large IS-A taxonomy it can be used in more general tasks (http://www.nlm.nih.gov/mesh/meshhome.html). Additionally, thanks to NGO Health-on-the-Net (HON, http://www.hon.ch/), a tool for recognizing terms in free text and grounding them to the MeSH taxonomy was available to us.

There are four rows in table 1: the first one, *Sent+BLSumm$_{neg}$*, is the performance of the baseline LSA summarizer on the negative posts (i.e., using only words), the second one, *Sent+Summ$_{neg}$*, is the enhanced LSA summarizer exploiting entities and IS-A relationships as given by the MeSH taxonomy, the third one, *Sent+BLSumm$_{pos}$*, presents the performance of the baseline LSA summarizer on the positive posts and the fourth one, *Sent+Summ$_{pos}$*, is the enhanced LSA summarizer for the positive posts.

Based on table 1 we can say that the results obtained with the enhanced LSA summarizer are overall better than the baseline summarizer. The numbers in bold show statistically significant improvement over the baseline system (note they are outside of the confidence intervals of the baseline system). The one exception where there is a slight drop in performance of the enhanced summarizer with respect to the baseline system is in the case of the negative posts for the metrics R_2 and R_{su4}, however, the *F1* is still within the confidence intervals of the baseline system, meaning the difference is not statistically significant.

We note that the main improvement in the performance of the enhanced summarizer comes from better precision and either no loss or minimal loss in recall with respect to the baseline system. The improved precision can be attributed, on one hand, to the incorporation of entities and IS-A relationships, but also, on the other hand, to the use of a better sentiment analyzer than the one used to produce the results of the baseline system.

We conclude that exploiting higher-level semantic information such as entities and IS-A relationships does bring a tangible improvement for the opinion-oriented summarization of blogs.

7 Conclusions

In this paper we measured the impact of exploiting higher-level semantic information such as named entities and IS-A relationships for the automatic summarization of positive and negative opinions in blog threads. We ran in tandem a sentiment analyzer and an LSA-based summarizer in two configurations: one using only words which we set as our baseline system, and another one making use in addition of entities and IS-A relations which we called the enhanced LSA summarizer. We used an annotated corpus and the standard ROUGE scorer to automatically evaluate the performance of our system. We

conclude that making use of higher-level semantic information as given by named entities and IS-A relationships does bring a tangible improvement for the opinion-oriented summarization of blogs.

In future work, we intend to analyze in more detail the cases where our system fails as well as the cases where a standard framework for evaluating summarization system falls short in providing adequate results for the task of producing opinion-oriented summaries.

REFERENCES

1. Kabadjov, M., Steinberger, J., Pouliquen, B., Steinberger, R., and Poesio, M. Multilingual statistical news summarisation: Preliminary experiments with English. In Proceedings of theWorkshop on Intelligent Analysis and Processing of Web News Content at the IEEE / WIC / ACM International Conferences on Web Intelligence and Intelligent Agent Technology (WI-IAT), (2009)
2. Steinberger, J., Poesio, M., Kabadjov, M., and Jezek, K. Two uses of anaphora resolution in summarization. Information Processing and Management, 43(6):1663–1680. Special Issue on Text Summarization (Donna Harman, ed.), (2007)
3. Hovy, E. H. Automated text summarization. In Mitkov, R., editor, The Oxford Handbook of Computational Linguistics, pages 583–598. Oxford University Press, Oxford, UK, (2005)
4. Erkan, G. and Radev, D. R. LexRank: Graph-based centrality as salience in text summarization. Journal of Artificial Intelligence Research (JAIR), (2004)
5. Gong, Y. and Liu, X. Generic text summarization using relevance measure and latent semantic analysis. In Proceedings of ACM SIGIR, New Orleans, US, (2002)
6. Pang, B. and Lee, L. A sentimental education: Sentiment analysis using subjectivity summarization based on minimum cuts. In Proceedings of the Annual Meeting of the Association for Computational Linguistics, pages 271–278, Barcelona, Spain, (2004)
7. Riloff, E. and Wiebe, J. Learning extraction patterns for subjective expressions. In Proceeding of the Conference on Empirical Methods in Natural Language Processing, (2003)
8. Kim, S. and Hovy, E. (2004). Determining the sentiment of opinions. In Proceedings of the International Conference on Computational Linguistics (COLING), (2004)
9. Turney, P. and Littman, M. Measuring praise and criticism: Inference of semantic orientation from association. ACM Transactions on Information Systems, 21, (2003)

10. Balahur, A., Steinberger, R., van der Goot, E., Pouliquen, B. and Kabadjov, M. Opinion mining from newspaper quotations. In Proceedings of the Workshop on Intelligent Analysis and Processing of Web News Content at the IEEE / WIC / ACM International Conferences on Web Intelligence and Intelligent Agent Technology (WI-IAT), (2009)
11. Stoyanov, V. and Cardie, C. Toward opinion summarization: Linking the sources. In Proceedings of the COLING-ACL Workshop on Sentiment and Subjectivity in Text, Sydney, Australia. Association for Computational Linguistics, (2006)
12. Balahur, A., Lloret, E., Boldrini, E., Montoyo, A., Palomar, M., and Martínez-Barco, P. Summarizing threads in blogs using opinion polarity. In Proceedings of the Workshop on Emerging Text Types (eETTs 2009), satellite workshop to RANLP, (2009)
13. Balahur, A., Kabadjov, M., Steinberger, J., Steinberger, R. and Montoyo, A. Summarizing opinion in blog threads. In Proceedings of the 23rd Pacific Asia Conference on Language, Information and Computation (PACLIC 2009).
14. Kabadjov, M., Balahur, A. and Boldrini, E. (2009b). Sentiment Intensity: Is It a Good Summary Indicator? In Proceedings of the Language and Technology Conference (LTC 2009)
15. Conroy, J. and Schlesinger, S. Classy at TAC 2008 metrics. In Proceedings of the Text Analysis Conference, organized by the National Institute of Standards and Technology, (2008)
16. He, T., Chen, J., Gui, Z., and Li, F. CCNU at TAC 2008: Proceeding on using semantic method for automated summarization yield. . In Proceedings of the Text Analysis Conference, organized by the National Institute of Standards and Technology, (2008)
17. Bossard, A., Généreux, M., and Poibeau, T. Description of the LIPN systems at TAC 2008: Summarizing information and opinions. In Proceedings of the Text Analysis Conference, organized by the National Institute of Standards and Technology, (2008)
18. Varma, V., Pingali, P., Katragadda, R., Krisha, S., Ganesh, S., Sarvabhotla, K., Garapati, H., Gopisetty, H., Reddy, V., Bysani, P., and Bharadwaj, R. In Proceedings of the Text Analysis Conference, organized by the National Institute of Standards and Technology, (2008)
19. Balahur, A., Lloret, E., Ferrández, O., Montoyo, A., Palomar, M., and Muñoz, R. The DLSIUAES team's participation in the TAC 2008 tracks. In Proceedings of the Text Analysis Conference, organized by the National Institute of Standards and Technology, (2008)
20. Cruz, F., Troyano, J., Ortega, J., and Enríquez, F. (2008). The Italica system at TAC 2008 opinion summarization task. In Proceedings of the Text Analysis Conference, organized by the National Institute of Standards and Technology, (2008)
21. Zhou, L. and Hovy, E. Digesting Virtual "Geek" Culture: The Summarization of Technical Internet Relay Chats. In Proceedings of ACL 2005, (2005)

22. Zhou, L. and Hovy, E. On the summarization of dynamically introduced information: online discussions and blogs. In Proceedings of the AAAI Symposium on Computational Approaches to Analysing Weblogs (AAAI-CAAW), (2006)
23. Wilson, T., Wiebe, J., Hoffman, P. Recognizing contextual polarity in phrase-level sentiment analysis. In Proceedings of Proceeding of the Conference on Empirical Methods in Natural Language Processing, (2005)
24. Cerini, S., Compagnoni, V., Demontis, A., Formentelli, M., and Gandini, G. Language resources and linguistic theory: Typology, second language acquisition, English linguistics, chapter Micro-WNOp: A gold standard for the evaluation of automatically compiled lexical resources for opinion mining. Franco Angeli Editore, Italy, (2007)
25. Esuli, A., Sebastiani, F.. SentiWordNet: A Publicly Available Resource for Opinion Mining. In Proceedings of 4th International Conference on Language Resources and Evaluation, LREC (2006)
26. Strapparava, C. Valitutti, A. WordNet-Affect: an affective extension of WordNet. In Proceedings of the 4th International Conference on Language Resources and Evaluation, LREC (2004)
27. Deerwester, S., Dumais, S., Furnas, G. W., Landauer,T. K., and Harshman, R. Indexing by latent semantic analysis. Journal of the American Society for Information Science, 3(41), (1990)
28. Pouliquen, B. and Steinberger, R. Automatic construction of multilingual name dictionaries. In Cyril Goutte, Nicola Cancedda, Marc Dymetman, and George Foster, editors, Learning Machine Translation. MIT Press, NIPS series (2009)
29. Steinberger, J., Kabadjov, M., Pouliquen, B., Steinberger, R., Poesio, M. (2009). WB-JRC-UT's participation in TAC 2009: Update Summarization and AESOP tasks. In Proceedings of the Text Analysis Conference, organized by the National Institute of Standards and Technology (2009)

ALEXANDRA BALAHUR
UNIVERSITY OF ALICANTE,
DEPARTMENT OF SOFTWARE AND COMPUTING SYSTEMS,
AP. DE CORREOS 99, E-03080 ALICANTE, SPAIN
E-MAIL: <ABALAHUR@DLSI.UA.ES>

MIJAIL KABADJOV
EUROPEAN COMMISSION – JOINT RESEARCH CENTRE
IPSC - GLOBESEC - OPTIMA
(OPENSOURCE TEXT INFORMATION MINING AND ANALYSIS)
T.P. 267, VIA FERMI 2749. 21027 ISPRA (VA), ITALY
E-MAIL: <MIJAIL.KABADJOV@JRC.EC.EUROPA.EU>

JOSEF STEINBERGER
EUROPEAN COMMISSION – JOINT RESEARCH CENTRE
IPSC - GLOBESEC - OPTIMA
(OPENSOURCE TEXT INFORMATION MINING AND ANALYSIS)
T.P. 267, VIA FERMI 2749. 21027 ISPRA (VA), ITALY
E-MAIL: <JOSEF.STEINBERGER@JRC.EC.EUROPA.EU>