Creating Domain-Specific Semantic Lexicons for Aspect-Based Sentiment Analysis

Panos Alexopoulos* and Manolis Wallace†
*Expert System Iberia, Av. del Partenon 10, 28042, Madrid, Spain
Email: palexopoulos@expertsystem.com

† Knowledge and Uncertainty Research Laboratory
Department of Informatics and Telecommunications
University of Peloponnese, Tripolis, Greece 22 100
Email: wallace@uop.gr - Web: http://gav.uop.gr

***Abstract*—Aspect based sentiment analysis (ABSA) is an opinion mining process where texts are analyzed to extract the sentiments that their authors express towards certain features and characteristics of particular entities, such as products or persons. Key role in the effectiveness of this process plays the accurate and complete identification of the entities' discussed aspects within the text, as well as of the evaluation expressions that accompany these aspects. Nevertheless, what entities may be considered as aspects and what evaluation expressions may characterize them, depends largely on the domain at hand. With that in mind, in this paper we propose an approach for representing and populating semantic lexicons that contain domain-specific aspect-evaluation-polarity relations and, as such, can be (re-)used towards more effective ABSA in concrete domains and scenarios.**

I. Introduction

Given a set of texts discussing a particular entity (e.g., reviews about restaurants), aspect based sentiment analysis (ABSA) attempts to identify the most prominent aspects of the entity (e.g., prices, food, ambience, etc.) and the expressed sentiment for each aspect (e.g., positive or negative) [1]. This is a challenging problem that has received a great deal of interest in the research community, with several approaches trying to tackle it, mainly by means of natural language processing and statistical learning techniques [2] [3] [4] [5]. In fact, in a relevant challenge that took place last year [6], more than 70 ABSA systems were submitted and evaluated against a common golden dataset.

Many of these systems incorporate within their methods sentiment lexicons, i.e., set of terms associated with one or more numerical values that quantify their sentiment polarity. SentiWordNet[1], for example, is an extension of WordNet that assigns to each of the latter's synsets three sentiment scores, one for positivity, one for negativity and one for objectivity. Other known sentiment lexicons are the Bing Liu opinion lexicon[2], the General Inquirer lexicon[3], the MPQA subjectivity lexicon[4] and the NRC Hashtag Sentiment Lexicon[5]. Moreover, there are also ontologies and semantic resources that contain knowledge related to sentiment analysis like, for example, SenticNet [7] or the Visual Sentiment Ontology[6] [8]. The main characteristic of these latter resources is that they associate polarity (and other) scores with concepts rather than mere words or synsets (e.g., *"high price"*).

Our position in this paper is that, while these resources are very useful, the ABSA task could benefit from the existence of further, more domain-specific, background knowledge that would include the following:

- **The domain's entities and characteristics that are potential aspects.** For a given domain (e.g. films) not all entities's characteristics are necessarily subject to opinions (e.g., it is rather rare for one to express a positive or negative opinion about a film's genre). For that, knowing a priori what entities are potential aspects can increase ABSA's precision as non-aspect entities would be filtered out.
- **The evaluation expressions that may express opinion for the domain's aspects.** For example, the aspect *"Food"* for a restaurant can be characterized as *"tasteless"* while the aspect *"Price"* as *"expensive"*. Knowing what expressions go with what aspects can lead to better aspect identification by simply using these expressions as evidence towards the implied aspect.
- **The typical polarity that evaluation expressions carry for a given aspect.** For example, saying that *"prices are high"* for a restaurant is typically negative while saying that *"standards are high"* is positive. This means that i) the evaluation expression's polarity cannot be isolated from the aspect it characterizes and ii) knowing this domain-specific polarity can help identify more accurately the aspect's sentiment within the text.

To the best of our knowledge, existing lexicons or ontological semantic resources do not currently fully provide

[1] http://sentiwordnet.isti.cnr.it/
[2] http://www.cs.uic.edu/liub/FBS/sentimentanalysis. html
[3] http://www.wjh.harvard.edu/inquirer/homecat.htm
[4] http://mpqa.cs.pitt.edu/ 8http://sentiwordnet.isti.cnr.it/
[5] http://www.umiacs.umd.edu/saif/WebDocs/NRCHashtag-Sentiment-Lexicon-v0.1.zip
[6] http://visual-sentiment-ontology.appspot.com/

the above type of information, mainly due to lack of domain-specific aspect-evaluation relations. For example, SentiWordNet does not take into consideration the aspect dependence of the synset's polarity, as the example with *"high prices"* and *"high standards"* illustrated above. Similarly, SenticNet, suggests that *"high price"*[7] has a positive polarity, yet in the restaurant domain the opposite is true.

Given this, in this paper we describe an approach for representing and populating semantic lexicons that contain domain-specific aspect-evaluation-polarity relations and, as such, may be (re-)used towards more effective ABSA in concrete domains and scenarios. In particular, we define a SKOS-based ontological schema for representing such relations (section II) and we develop a semi-automatic method for populating these relations from text (section III). As a use case we consider the restaurant domain in which we evaluate the effectiveness of the lexicon population process (section IV).

II. Aspect-Evaluation-Polarity Ontology

The Aspect-Evaluation-Polarity ontology aims to enable the answering of the following competency questions:

1) Given a domain, what are the entities (classes, instances, etc.) that may have the role of an aspect, i.e., entities for which a sentiment may be expressed?
2) What are the evaluation expressions that may be used to express a sentiment for a given aspect?
3) What is the typical polarity an evaluation expression has for a given aspect?

To achieve that, we define the ontological schema of figure 1, comprising the following elements:

- Class *Aspect*: Consists of the characteristics for which an opinion or sentiment can be expressed in a given domain. If the domain is already represented as an ontology, then these characteristics may be classes (e.g. Restaurant, Location,Food, Author, Actor, etc.), individuals (e.g. Soup, Steak, Windows 8, etc.) object properties (e.g. hasSubject, hasLocation, servesFood, hasOperatingSystem, etc.) or even datatype properties (e.g. price, capacity, size, etc.). To capture this generality, we model the *Aspect* class as a subclass of *skos:Concept*. Another reason for this modeling choice is that aspects may form taxonomies in which a child node is a more specific aspect of its parent (e.g., "Linux" is more specific than "Operating System"). SKOS's broader and narrower relations effectively enable the representation of such taxonomies.
- Class *AspectEvaluation*: Describes an evaluation that a given aspect may assume and consists of an evaluation expression (e.g., "tasty") and a polarity score (e.g., "positive").
- Object Property *hasAspectEvaluation*: Links an aspect to one or more evaluations it may assume.
- Datatype Property *hasEvaluationExpression*: The evaluation expression of a given aspect evaluation.
- Datatype Property *hasPolarity*: The polarity of the evaluation expression in a given aspect evaluation. This can be a categorical value, such as "positive" or "negative", or a number in some sentiment scale.

As an example, consider the restaurant domain and the entity "Food". In a restaurant review, one may characterize food, among others, as "tasty" or "decent", the former expressing a positive sentiment while the latter a neutral one. Using the above ontology, we can model this example as follows:

```
@prefix : <http://vocab.isoco.net/absa/> .
@prefix ex:<http://example.org/> .
ex:DecentFood a :AspectEvaluation ;
    :evaluationExpression "decent":
    :polarity "neutral".
ex:TastyFood a :AspectEvaluation
    :evaluationExpression "tasty":
    :polarity "positive".
ex:Food a :AspectEntity ;
    :hasAspectEvaluation ex:DecentFood ;
    :hasAspectEvaluation ex:TastyFood ;
```

Moreover, in an aspect taxonomy, the children aspects typically inherit the evaluations their parents can assume. For example, the term *"tasty"* can apply to the entity *"Food"* but also to specific foods such as *"soup"* or *"chicken"*. Therefore, it makes sense to assign these common evaluations only to the most generic aspect they are applicable to and have its children aspect inherit them by means of reasoning. To facilitate that in our ontology we define an SWRL rule (see figure 1) suggesting that if an aspect A has an evaluation E then all the narrower aspects of A have also this evaluation.

III. Ontology Population

In this section we propose a semi-automatic process by which one may populate the above ontology in specific domains. The process is depicted in figure 2 and it is designed to serve two goals:

1) The discovery of evaluation expressions (and their polarity) for already known/identified aspects in a given domain.
2) The discovery of aspect-evaluation-polarity triples involving aspects that were previously unknown.

For both goals, we require as input a corpus of opinionated sentences about our target domain (e.g., restaurant reviews). A second input, required for the first goal, is a set of known domain aspect entities. Given these, the population process works as follows. First, the corpus sentences are processed by a subjectivity detection system in order to filter out sentences that do not express an opinion. Then, a Named Entity Resolution system is applied in order to

[7] http://sentic.net/api/en/concept/price_high

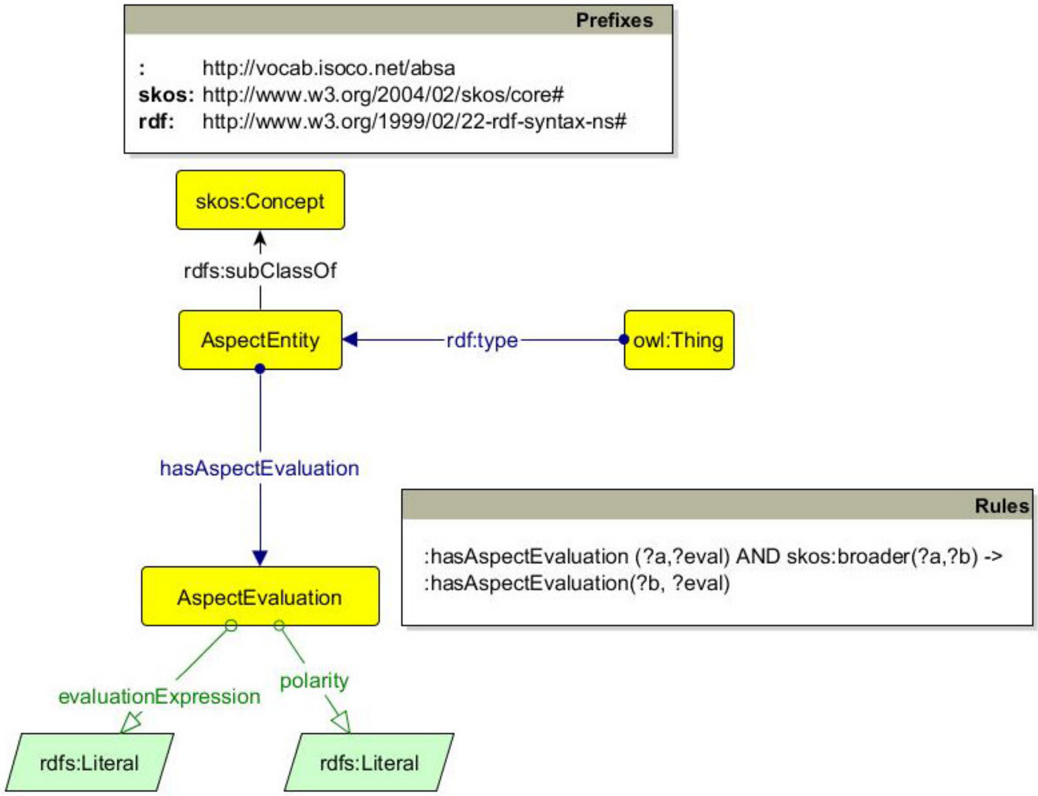

Figure 1. Aspect-Evaluation-Polarity Ontology.

identify which of these opinionated sentences mention a known aspect entity.

Subsequently, a relation extraction system is applied to all sentences in order to extract pairs of aspects and evaluation expressions. The frequency of the extracted pairs within the corpus is used as a parameter for ranking the extracted pairs and fine-tuning the precision/recall of the process; higher frequency provides more confidence for the extracted pair.

The final step of the process involves determining the polarity of each extracted pair. To do that we make the hypothesis that positive aspect evaluations appear mostly in positive contexts and vice versa. Given this, pair polarity is calculated as follows:

1) For each unique pair we gather (from the sentence collection) the textual contexts it was detected in (one context per sentence).
2) To each context we perform sentiment analysis (via some existing tool) in order to derive a polarity score for it.
3) The average of these scores is assigned to the pair.

The polarized aspect-evaluation pairs are formally represented by means of the previous section's ontology. For a given extracted pair related to an already known aspect we create an instance of the AspectEvaluation class and assign to it the extracted evaluation expression. If an evaluation expression is shared by two or more aspects, then it is assigned only to the most generic ones. Moreover, expressions that are shared between aspects at the same level of the taxonomy are manually checked for generality, i.e. whether they may refer to more generic aspects. For example the expression *"tasty"* should be linked to the aspect "Food" even if in the corpus only specific foods are characterized as such. If that's the case, then the expression is moved to the more generic aspect. For extracted pairs that contain previously unknown aspects we work in a similar fashion though first we need to manually validate the newly discovered aspects.

IV. Implementation and Evaluation

In order to assess the feasibility and effectiveness of our lexicon generation process, we applied it in the restaurant domain, using:

- A restaurant domain ontology that we developed and whose aspect entities we manually identified.
- A corpus of 2000 restaurant review sentences, derived from the dataset used in the Semeval-2014 ABSA challenge[8].
- A concrete implementation of figure's 2 pipeline.

[8]http://goo.gl/nOXbZb

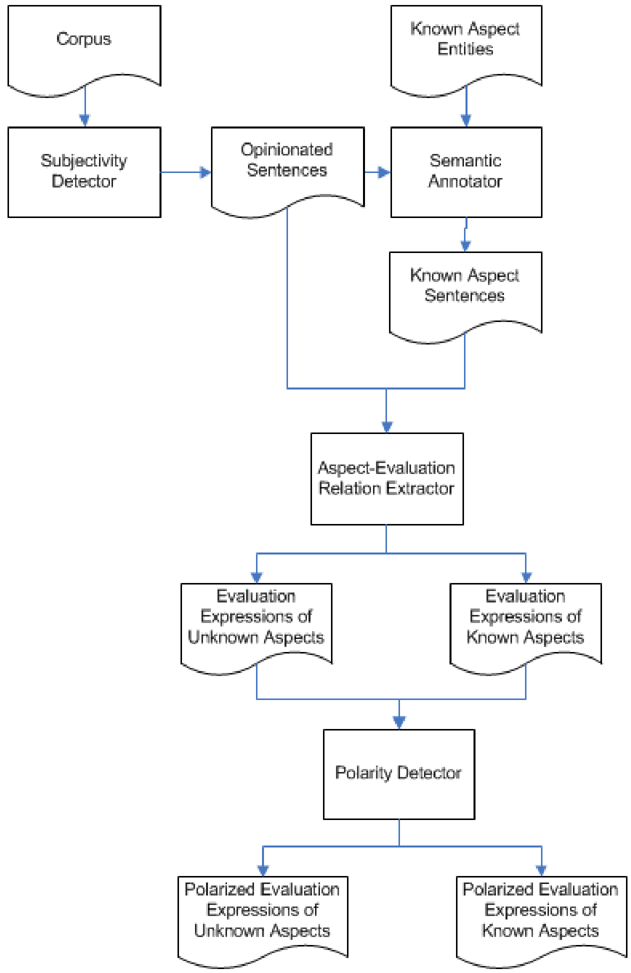

Figure 2. Aspect-Evaluation-Polarity Ontology Population Pipeline.

In the implemented pipeline, subjectivity detection was facilitated by OpinionFinder[9] while semantic annotation by Knowledge Tagger [9], an in-house developed named entity resolution tool, particularly applicable to domain-specific application scenarios. For the aspect-evaluation extraction part, we utilized a pattern-based framework [10] that uses dependency grammars to learn relation occurrence patterns which then applies to extract new relations from text. The framework requires as training input sentences that are already annotated with pairs of the target relations; for that we (manually) annotated the sentences of the restaurant reviews corpus with pairs of aspects and evaluation expressions like the following:

```
<sentence>But the staff was horrible to us.
</sentence>
<aspectTerms>
    <aspectTerm term="staff"
    eval_expression="horrible"/>
</aspectTerms>
```

```
<sentence>The food is exceptional, with a
very capable kitchen. </sentence>
<aspectTerms>
    <aspectTerm term="food"
    eval_expression="exceptional"/>
    <aspectTerm term="kitchen"
    eval_expression="capable"/>
</aspectTerms>
```

Finally, for the polarity detection part, we used the Stanford Sentiment Analysis Tool[10].

A. *Evaluation of Aspect-Evaluation Pair Extraction*

Our first task was to measure the precision and recall of the aspect-evaluation pair extraction process, both for previously known aspects (i.e., aspects that we had identified in our domain ontology) and for unknown ones. Precision was measured as the ratio of the correctly extracted pairs to the total extracted pairs, while recall as the ratio of the correctly extracted pairs to the total actual pairs in the

[9] http://mpqa.cs.pitt.edu/opinionfinder/

[10] http://nlp.stanford.edu/sentiment/index.html

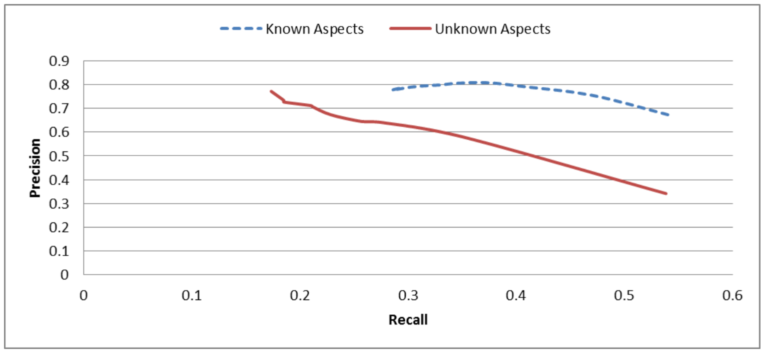

Figure 3. Precision-Recall Curves for the Extraction of Aspect-Evaluation Pairs.

corpus. To perform this measurement we used 75% of the 2000 sentences as training data for the relation extraction component and 25% for testing. As a result of this process we obtained (via ten-fold cross-validation) the precision-recall curves of figure 3, one for the extraction of aspect-evaluation pairs involving known aspects and one for pairs involving unknown aspects. Each point on a curve corresponds to a different combination of precision and recall, obtained by using a different threshold for the frequency of the extracted pairs.

As one can easily see, the system achieves a precision as high as 80% for known aspects and 72% for unknown ones, though in the latter case recall is relatively small. The achieved levels of recall, we believe, are primarily a consequence of the corpus we used and the learning difficulty of the aspect-evaluation patterns it contained. To verify this, we used another pattern-based relation extraction tool, namely LEILA[11], and we got very similar results. Further study of this issue in order to improve the system's effectiveness is left as future work.

In any case, as our initial intuition was, higher frequency thresholds were found to result in higher levels of precision. This means that the particular parameter can be used by the system's user to control the size and quality of the system's output; if correct but few results are returned then he/she can lower the threshold in order to increase their number, whereas if many but incorrect results are returned, he/she can increase the threshold to increase their precision.

B. Evaluation of Aspect-Evaluation Pair Polarity Detection

To evaluate the accuracy of the pair polarity detection we considered 56 (distinct) aspect-evaluation pairs from our dataset along with the sentences they were found in. The number of sentences per pair was between 3 and 10. We then determined the polarity of these pairs both manually (using 2 human judges) and automatically (using the context polarity calculation approach we described in the previous section). The comparison between the manual and automatic assignments indicated an 80% accuracy of the automatic approach.

Among others, the system managed to correctly identify "high price" or "tiny portion" as negative concepts in the particular domain while "fast service" or "affordable food" as positive ones. This is important as, for example, the term "high" has a positive polarity in SentiWordnet or the Thelwall-Lexicon, while the concept "high price" has a positive polarity in SenticNet. Similarly, the term "fast" (in the sense of acting or moving quickly) has a neutral polarity in SentiWordnet.

V. Related Work

In terms of the kind of knowledge our proposed lexicon is designed to provide, the most similar resources are SenticNet [7] and the Visual Sentiment Ontology [8]. The former associates sentiment scores to 30,000 common and common-sense concepts, including some that could be seen as aspect-evaluation pairs, such as *"tasty_food"* or *"busy_street"*. Yet the resource does not explicitly state which concepts can have this role in a given domain, neither the associated polarity scores vary according to the domain. On the other hand, the Visual Sentiment Ontology, contains 1172 noun-adjective- pairs, such as *"beautiful girl"* or *"little house"*, accompanied by a (positive or negative) sentiment score. The pairs have been extracted from image tags and their sentiment scores have been automatically calculated by adding the SentiWordNet polarities of the noun and the adjective. This, nevertheless, means that polarity values are aspect-independent since, for example, *"high price"* and *"high standards"* would have a similar if not identical score.

Regarding the automatic extraction of aspect-evaluation pairs from text, relevant works that explicitly address this task are limited. In [11] the authors utilize co-occurrence patterns of subject, attributes and value expressions to extract such pairs from Japanese texts. However, their approach assumes that there are already available domain specific dictionaries of evaluation and aspect phrases, limiting the task to only the detection of pairs within the texts. On the other hand, in [12], the extraction of pairs is performed on

[11] http://www.mpi-inf.mpg.de/yago-naga/leila/index.html

semi-structured online reviews that have the form of pros and cons lists. This simple structure makes the task much easier than if it were to be executed on more complex texts as we have done in this paper.

VI. Conclusions and Future Work

In this paper we have defined a framework for representing and populating semantic lexicons that contain domain-specific aspect-evaluation-polarity relations and which may be used in the context of aspect-based sentiment analysis. The framework consists of a SKOS-based ontological schema for representing such relations as well as a semi-automatic process for extracting these relations from texts. The population process has been evaluated in the restaurant domain, showing a a good level of accuracy, especially for the subtask of evaluation extraction of known aspects.

The key difference of our approach from other relevant approaches is that the generated lexicons consider the polarity of an evaluation expression as something domain and aspect-dependent. In our future work, we intend to improve the effectiveness of our lexicon population pipeline, mainly in terms of recall, by contemplating methods that may detect more complex patterns of aspect-evaluation pairs.

References

[1] B. Liu, *Sentiment Analysis and Opinion Mining*, ser. Synthesis Lectures on Human Language Technologies. Morgan & Claypool Publishers, 2012.

[2] M. Hu and B. Liu, "Mining and summarizing customer reviews," in *Proceedings of the Tenth ACM SIGKDD International Conference on Knowledge Discovery and Data Mining*, ser. KDD '04, 2004, pp. 168–177.

[3] I. Titov and R. McDonald, "Modeling online reviews with multi-grain topic models," in *Proceedings of the 17th International Conference on World Wide Web*, ser. WWW '08, 2008, pp. 111–120.

[4] S. Brody and N. Elhadad, "An unsupervised aspect-sentiment model for online reviews," in *Human Language Technologies: The 2010 Annual Conference of the North American Chapter of the Association for Computational Linguistics*, ser. HLT '10, 2010, pp. 804–812.

[5] Y. Jo and A. H. Oh, "Aspect and sentiment unification model for online review analysis," in *Proceedings of the Fourth ACM International Conference on Web Search and Data Mining*, ser. WSDM '11, 2011, pp. 815–824.

[6] M. Pontiki, D. Galanis, J. Pavlopoulos, H. Papageorgiou, I. Androutsopoulos, and S. Manandhar, "Semeval-2014 task 4: Aspect based sentiment analysis," in *Proceedings of the 8th International Workshop on Semantic Evaluation (SemEval 2014)*. Dublin, Ireland: Association for Computational Linguistics and Dublin City University, August 2014, pp. 27–35.

[7] E. Cambria, D. Olsher, and D. Rajagopal, "Senticnet 3: A common and common-sense knowledge base for cognition-driven sentiment analysis," 2014.

[8] D. Borth, R. Ji, T. Chen, T. Breuel, and S.-F. Chang, "Large-scale visual sentiment ontology and detectors using adjective noun pairs," in *Proceedings of the 21st ACM International Conference on Multimedia*, ser. MM '13, 2013, pp. 223–232.

[9] P. Alexopoulos, B. Villazon-Terrazas, and J. M. Gomez-Perez, "Knowledge tagger: Customizable semantic entity resolution using ontological evidence." in *I-SEMANTICS (Posters & Demos)*, ser. CEUR Workshop Proceedings, S. Lohmann, Ed., vol. 1026. CEUR-WS.org, 2013, pp. 16–19.

[10] I. Tiddi, N. Mustapha, Y. Vanrompay, and M.-A. Aufaure, "Ontology learning from open linked data and web snippets," in *On the Move to Meaningful Internet Systems: OTM 2012 Workshops*, ser. Lecture Notes in Computer Science, P. Herrero, H. Panetto, R. Meersman, and T. Dillon, Eds. Springer Berlin Heidelberg, 2012, vol. 7567, pp. 434–443.

[11] N. Kobayashi, K. Inui, Y. Matsumoto, K. Tateishi, and T. Fukushima, "Collecting evaluative expressions for opinion extraction," in *Proceedings of the First International Joint Conference on Natural Language Processing*, ser. IJCNLP'04. Berlin, Heidelberg: Springer-Verlag, 2005, pp. 596–605.

[12] J. Bro and H. Ehrig, "Generating a context-aware sentiment lexicon for aspect-based product review mining." in *Web Intelligence*, J. X. Huang, I. King, V. V. Raghavan, and S. Rueger, Eds. IEEE, 2010, pp. 435–439.