Spanish Corpus for Sentiment Analysis Towards Brands

María Navas-Loro[1(✉)], Víctor Rodríguez-Doncel[1], Idafen Santana-Perez[1], and Alberto Sánchez[2]

[1] Ontology Engineering Group, Universidad Politécnica de Madrid, Madrid, Spain
mnavas@fi.upm.es
[2] Havas Media, Madrid, Spain

Abstract. Posts published in the social media are a good source of feedback to assess the impact of advertising campaigns. Whereas most of the published corpora of messages in the Sentiment Analysis domain tag posts with polarity labels, this paper presents a corpus in Spanish language where tagging has been made using 8 predefined emotions: love-hate, happiness-sadness, trust-fear, satisfaction-dissatisfaction. In every post, extracted from Twitter, sentiments have been annotated towards each specific brand under study. The corpus is published as a collection of RDF resources with links to external entities. Also a vocabulary describing this emotion classification along with other relevant aspects of customer's opinion is provided.

Keywords: Corpus · Sentiment analysis · NLP · Opinion mining

1 Introduction

Emotions, rather than cognitive thinking, determine our purchase decisions. Modern marketing campaigns strive to link brands to specific emotions and the success of these campaigns can be evaluated with more complex instruments than the mere figures of sales. Emotions aroused by brands can be found in posts in the social media, and computer algorithms can, to some extent, automatically evaluate the impact of the marketing campaigns. These messages are important *per se*, as a large percent of social media users (up to 70% according to Nielsen [33]) take into account the product experience published by other users.

Even if Sentiment Analysis has progressed fast in the last few years, there is not much research on other aspects of the message besides polarity that might be useful for commercial companies and the image of their brands. One of the objectives of the LPS BIGGER [25] project is to cover this gap. The intended analysis goes deeper in more complex aspects and nuances of opinions, such as the feelings arisen in customers by different brands or the stage in the whole shopping process the client is in at the moment of giving their opinion. Once combined, all these analyses result into remarkably rich information that opens up rich potential exploitation opportunities, such as automatically personalized

A. Karpov et al. (Eds.): SPECOM 2017, LNAI 10458, pp. 680–689, 2017.
DOI: 10.1007/978-3-319-66429-3_68

offering generation or immediate reactions to events related to brands. Focusing mainly on Twitter as source of opinions, at least the following four aspects (see Fig. 1) are of interest with respect to the brands:

- **Sentiment Analysis** identifies emotions towards a brand in a post beyond polarity. Several classification of human sentiments have been proposed in Psychology, such as Plutchik's [36] or Ekman's [14]; the one used in the project is based in the taxonomy stated by the latter in conjunction with those by Shaver [45] and Richins [39], distinguishing between four non-exclusive sentiments and their direct opposite: *love* and *hate*, *happiness* and *sadness*, *trust* and *fear*, and *satisfaction* and *dissatisfaction*. This new taxonomy has been proposed by industrial partners in the LPS BIGGER project as a response of an uncovered necessity of an emotion taxonomy thought specifically for marketing purposes.
- **Purchase Funnel** places the opinion within a five-staged consumer decision journey: *Awareness*, *Evaluation*, *Purchase*, *Postpurchase* and *Review*.
- **Marketing Mix** comprises the different marketing strategies the customers can evaluate, known as the four Ps: *product*, *price*, *promotion* and *place* [3].
- **Meaningful Brands™** is a metric proposed by Havas Media [23] that measures the value of the brand, based on the customer's wellbeing. It is divided in *marketplace* (relating the product to performance such as quality and price), *personal wellbeing* (such as self-esteem) and *collective wellbeing* (the role brands play in communities).

Given that the available corpora, identified in Sect. 2, are only of tangential interest for classifying Twitter messages related to brands, we have built a simple vocabulary and a new corpus to fill the gap: the Sentiment Analysis towards Brands (SAB). Whereas the vocabulary covers the four aspects described above, our first release covers only the brand and emotional tagging, focusing on the Sentiment Analysis aspects. This corpus is published both as an spreadsheet and as linked RDF [41], using vocabulary terms defined by well-known ontologies and mapping some of the resources with external datasets.

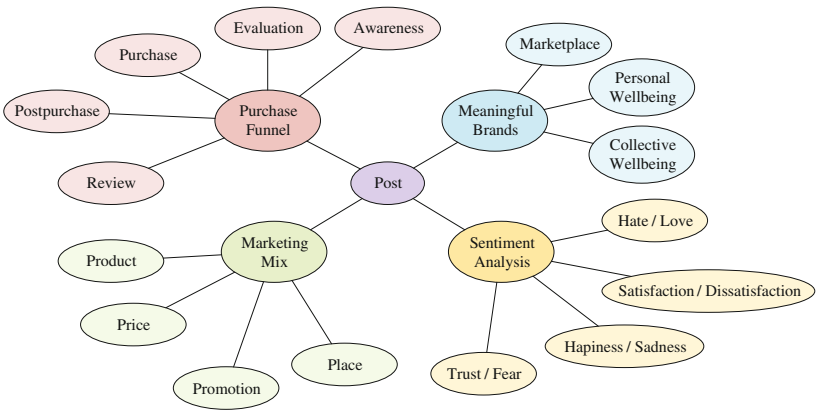

Fig. 1. The four aspects of interest for each tweet.

2 Related Work

2.1 Corpora for Sentiment Analysis

Even when most of the available corpora for Sentiment Analysis is English oriented, several Spanish corpora have also been described. Opinions in fields as different as Medicine [35] or Tourism [29] have been collected from diverse sources, such as social networks [11,26,38] or specialized opinion websites [10,29,35].

The publication style of corpora of tweets varies. Many researchers collect by themselves a set of tweets and keep it as a private corpus; others share the tweet IDs and instructions to retrieve them from Twitter, or share them preprocessed [52]. This behavior often responds to Twitter policies on text dissemination [30]. However, since Twitter periodically deletes tweets from their servers (making the text eventually irretrievable), the lifetime of a corpus with no text but only ID is randomly short, hindering its re-usability. Some of the few corpora available with text are only distributed on demand for private use, such as happens with the TASS [11] and the COST [26] corpora (see Table 1 for a review of representative corpora). In addition to this lack of appropriate public Twitter corpora, we also find that mainstream Opinion Mining annotation provides just the notion of polarity, determining if an opinion is positive or negative, sometimes expanded with intensity [11] or a rating scale [10,29,35]. The only available sentiment corpus the authors are aware of in Spanish is EmIroGeFb [38]; however, since it does not refer to brands and uses a different sentiment classification, none of the requirements for the project are fulfilled. The Spanish corpus for Sentiment Analysis towards Brands (SAB) we present covers therefore a gap in Spanish Sentiment Analysis, providing emotion tags toward brands, even if it is inevitably subordinated to Twitter policies[1].

2.2 Ontologies for Sentiment Analysis Towards Brands

Not many published ontologies are of use for supporting post classification within the coordinates of interest for brands. The marketing mix is supported by Sam's ontology [42], whereas, up to the authors' knowledge, there is total lack of vocabulary to represent the stages in the purchase funnel. However, several computer ontologies have been proposed for supporting the knowledge representation needs in Sentiment Analysis tasks. We can distinguish among them ontologies about emotions [17,19–21,34,37,56], usually based in previous emotion classifications such as those of Plutchik [36], Ekman [14] or Nakamura [32], and ontologies dealing with opinion representation [43,55]; due to our concrete case, we also include in this last group those representing Twitter services [50] and Sentiment Analysis on tweets [9,40].

One of the most referred ontologies in the context of sentiment representation is Ontoemotion [17], an ontology developed by Universidad Complutense

[1] The SAB corpus is available online offering only the ID of the tweets.

Table 1. Spanish Corpora available for sentiment analysis.

Corpus	Sector	Source	Annotation	Amount	Access	Text
HOPINION [8]	Tourism	TripAdvisor [51]	Rating (1-5)	17934	Registration	Yes
COST [26]	General	Twitter	Polarity (0/1)	34634	On request	Yes
COPOS [35]	Patient opinions	MasQueMedicos [27]	Rating (0-5)	743	On request	Yes
COAH [29]	Tourism	TripAdvisor	Rating (1-5)	1816	Registration	Yes
COAR [6]	Restaurants	TripAdvisor	Rating (1-5)	2202	Registration	Yes
Spanish Movie Reviews [10]	Cinema	MuchoCine [31]	Rating (1-5)	3878	Free	Yes
TASS General [11]	Personalities	Twitter	Level of agreement, polarity (P+,P,NEU,N, N+,NONE)	>68000*	On Request	Yes*
TASS Social-TV [11]	Sports	Twitter	Polarity (P,NEU,N), aspects	2773*	On Request	Yes*
TASS STOMPOL [11]	Politics	Twitter	Polarity (P,NEU,N), aspects	1284*	On Request	Yes*
SFU Spanish Corpus [46]	Several items	Ciao [5]	Rating (1-2,4-5)	400	Registration	Yes
EmIroGeFb [38]	Politics, Football, Celebrities	Facebook	Ekman emotions, gender, topic, presence of irony	1200	Free	No(IDs)

(*) As of 2015 [11]. TASS corpora change every year.

de Madrid for Emotional Voice Synthesis and later extended for its use in Italian texts [2]. In this ontology, emotions are defined in a space of three emotional dimensions (*Power*, *Activation* and *Evaluation*), having as one of the root classes the concept of *Emotion*. Also in the media context we find the Ontology of Emotional Cues [34], that models emotional cues linking them with the media properties that reveal them and classifying the different types of cues (e.g. verbal or psychological); in this ontology, the concrete emotions could both be expressed as categories or dimensions. For representing emotional responses, the EmotionsOnto ontology [19] (and its later version [48]) offers an easy integration both with FrameNet [1] and the DOLCE upper ontology [18]. Another proposal, the Emotion Ontology [21,22], represents emotions related to mental diseases, linking to other ontologies of the field but being also usable just for emotional purposes.

Finally, it must be noted that also cultural differences intervene in emotion classification. Such is the case of the Chinese ontology [56] built from the Chinese knowledge-base HowNet [13], or the one built from Japanese corpora [37] by using among others the EmotionML markup language [15] and Nakamura's emotion classification with only two binary dimensions: *Valence* (*positive* or *negative*) and *Activation* (*activated* or *deactivated*). In the case of dealing with different sentiment classifications, the high level Human Emotion Ontology (HEO) [20] covers different sentiment taxonomies and supports different dimensions, using as default *Arousal*, *Valence* and *Dominance* but admitting also other emotional spaces (such as the previously presented for Ontoemotion).

Related to opinion representation, the Marl [55] and Onyx [43] ontologies allow to represent opinions in RDF, being the former oriented to polarity and the latter to emotions. They are aligned with previous efforts and ontologies, such as WordNet-Affect [47] and linguistic linked data resources. Focusing on tweet representation, the Twitter API Rest Ontology [50] allows to represent the whole REST process but can also be used partially for expressing the opinion text and related information (such as the user and if the tweet is a retweet or not); other proposals include TwO [53] (Twitter Ontology) or SIOC [4]. Reusing some of these options we find ontologies directly dealing with Sentiment Analysis in tweets, such as EmpaTweets [40] and TweetOntoSense [9].

3 Building the Corpus

In existing corpora we find different ways to obtain tweets: some of them are built from concrete sentiment seeds, looking for polarized hashtags [28] or emoticons [26]. Even when this leads to corpora richer in actual sentiment-expressing messages, most of posts with just non-explicit emotions are lost in the process. Since the LPS BIGGER project demands a system capable of detecting also tweets with no emotions, ambiguous or without a context, we don't want to lose in our corpus this kind of messages. What we search are therefore the names of the brands we want to analyze, just imposing a constraint in the language of the tweets. Since not all the users directly refer to the brand by its official profile or the complete name of the brand, we also searched for names commonly used. The steps in the corpus building process included therefore preprocessing, and were the following:

1. Selection of the brands to analyze: we need to know the official names but also the Twitter profiles and the commonly used names for each brand in order to retrieve related tweets. The final list of analyzed brands (derived from the LPS BIGGER project) can be found in the website of the corpus.
2. Acquisition of tweets: the data collection took place between the 1st and the 7th February of 2017, having different capture processes (with different keywords searched) for each sector. The only filters used were the language ("es" for Spanish) and the brand keywords; tweets marked as retweets were not retrieved.
3. Sifting: The collected tweets were screened, searching for repeated tweets. Also messages where there was no real brand (in case the brand name was polysemous, or might appear as a part of other words) were deleted; so was done for tweets in other languages (even with the Twitter language filter, some messages in different languages managed to pass) and repeated tweets where the only difference was a URL (since Twitter automatically shortens them, the message would be in this case exactly the same).
4. Tagging: Three different people intervened in the tagging process, which consisted in determining if each message showed or not each of the emotions on the taxonomy (being possible several for the same message), or if on the

contrary it was a neutral tweet. For this task, a document explaining the criteria to follow was given to the taggers. These criteria (provided with the corpus) include for instance that a recurrent purchase should be tagged as *trust* and that *happiness* can only be inspired by an already acquired product or service; also usual combinations of emotions are explained: not finding a desired product would be for example *satisfaction* and *sadness*.

5. Transformation: treatment of the data to link it to other resources. Main ontologies and resources used, as well as a sample record, can be found in Sect. 4.

4 Spanish Corpus for Sentiment Analysis Towards Brands

The corpus comprises 4548 tagged messages, covering 7 sectors and 8 emotions; Table 2 shows percentages of appearance per sector and emotion, becoming apparent how neutral tweets (e.g. coming from community managers or news) must be also identified, since they often mention brands and usually contain emotional words that may mislead classifiers. It must be also noted how the occurrence of emotions is linked to the sector; *fear*, for instance, makes sense in BANKING but not in SPORTS or BEVERAGES.

Table 2. Column ANY shows the percentage of posts with any emotion (non neutral posts); remaining show the percentage for each emotion among these non neutral posts.

	ANY	HAT	SAD	FEA	DIS	SAT	TRU	HAP	LOV
FOOD	50.68	2.69	2.15	0	15.05	82.26	80.11	26.88	23.66
AUTOMOTIVE	7.80	0	2.33	11.63	25.58	76.74	39.53	13.95	11.63
BANKING	21.21	21.19	3.97	60.93	96.69	5.30	1.99	0	0
BEVERAGES	63.12	3.46	1.85	1.15	30.25	69.75	51.96	11.78	12.24
SPORTS	34.15	7.17	7.62	0.90	39.01	55.16	34.98	14.35	33.18
RETAIL	20.41	9.96	3.69	4.80	36.90	43.91	43.54	11.44	10.33
TELECOM	38.96	32.99	2.06	0.00	75.26	21.65	15.46	8.25	3.09
TOTAL	30.88	9.05	3.42	8.33	41.03	54.06	43.09	12.68	14.74

In order to evidence the subjectivity of emotion tagging, one of the sectors (BEVERAGES) has been completely tagged by two additional people, as it is the most expressive one in terms of expressed sentiments. We calculated the inter-annotator agreement using both the Fleiss' kappa [16] between the three taggers (all of them Spanish native speakers) for each emotion and the Cohen's kappa [7] for pairwise inter-agreement. As shown in Table 2 for BEVERAGES, several emotions appear scarcely in the corpus, being therefore statistically insignificant and leading to unrealistic kappas. Conversely, well-represented emotions such as

```
lps:826812979421257730 a sioc:Post ;
  sioc:id "826812979421257730" ;
  sioc:content "Ya me quede sin credito?? Hace 3 dias tengo
  credito nomas... Movistar y la concha de tu hermana"@es ;
  marl:describesObject lps:Movistar ;
  lps:isInPurchaseFunnel lps:postPurchase;
  lps:hasMarketingMix lps:price;
  lps:hasMeaningfulBrand lps:marketplace;
  onyx:hasEmotion lps:hate, lps:dissatisfaccion ;
  marl:hasPolarity marl:negative ;
  marl:forDomain "TELCO" .

lps:hate a onyx:Emotion ;
  rdfs:label "odio"@es, "hate"@en .

lps:dissatisfaction a onyx:Emotion ;
  rdfs:label "insatisfaccion"@es, "dissatisfaction"@en .


lps:Movistar a gr:Brand ;
  rdfs:seeAlso <http://dbpedia.org/resource/Movistar> ;
  rdfs:label "Movistar" .

lps:1-5000062703 a gr:Business ;
  rdfs:label "Telefonica de Espana, S.A.U.";
  rdfs:seeAlso <https://opencorporates.com/companies/es/82018474> ;
  owl:sameAs permid:1-5000062703 .
```

Fig. 2. Sample tagged post, and extra information on its brand and company.

dissatisfaction get Fleiss' kappa of 0.372 and average Cohen's kappa of 0.354; detailed results and extensive information on distribution in the corpus can be found with it.

Pursuing a richer representation, the dataset is published in RDF with extensive links to other datasets. Different vocabularies and ontologies have been used, such as Marl and Onyx [44], SIOC [4] or GoodRelations [24]. In addition, our own vocabulary to cover the purchase funnel and the marketing mix has been published [54]; extended information on the brands and companies, such as links to external databases like Thomson Reuters' PermID [49] or DBpedia [12], is also provided whenever possible. A sample post in RDF referring to a given brand (*Movistar*) and tagged as *hate* and *dissatisfaction* is shown in Fig. 2.

5 Conclusions and Future Work

The SAB corpus presented in this paper is the first one in Spanish containing tagged tweets related to brands. The corpus has been published not only as a spreadsheet but also as an RDF graph linked to external resources like DBpedia and Thomson Reuters's PermID. After this first publication, the corpus will be enlarged in size and tagging labels in forthcoming versions. Results of common classification algorithms and strategies will follow for each of the categories: whether a tweet bears an emotion, at which stage of the purchase funnel the Twitter user is and whether the post is related to the marketing mix or with a meaningful brands dimension. We dare to make the reasonable guess that using

the information obtained from external entities (easily retrievable as the corpus is already linked) will enhance the results when comparing with the information in the bare text. This will hopefully support the publication of linked corpora, as it will evidence the advantages of using linked data for classification tasks.

Acknowledgments. This work has been partially supported by LPS-BIGGER (IDI-20141259, Ministerio de Economía y Competitividad), a research assistant grant by the Consejería de Educación, Juventud y Deporte de la Comunidad de Madrid partially founded by the European Social Fund (PEJ16/TIC/AI-1984) and a Juan de la Cierva contract.

References

1. Baker, C.F., Fillmore, C.J., Lowe, J.B.: The Berkeley FrameNet project. In: Proceedings of the COLING-ACL, vol. 1, pp. 86–90. ACL (1998)
2. Baldoni, M., Baroglio, C., et al.: ArsEmotica: emotions in the social semantic web. In: Proceedings of the 7th International Conference on Semantic Systems, pp. 171–174 (2011)
3. Borden, N.H.: The concept of the marketing mix. J. Advertising Res. **4**(2), 2–7 (1964)
4. Breslin, J.G., Decker, S., et al.: Sioc: an approach to connect web-based communities. Int. J. Web Based Communities **2**(2), 133–142 (2006)
5. Ciao, website with opinions on several topics. http://www.ciao.es/
6. Corpus COAR, with opinions about restaurants. http://sinai.ujaen.es/coar/
7. Cohen, J.: A coefficient of agreement for nominal scales. Educ. Psychol. Meas. **20**(1), 37–46 (1960)
8. Corpus HOpinion. http://clic.ub.edu/corpus/es/node/106
9. Cotfas, L.-A., Delcea, C., Roxin, I., Paun, R.: Twitter ontology-driven sentiment analysis. In: Barbucha, D., Nguyen, N.T., Batubara, J. (eds.) New Trends in Intelligent Information and Database Systems. SCI, vol. 598, pp. 131–139. Springer, Cham (2015). doi:10.1007/978-3-319-16211-9_14
10. Cruz, F.L., Troyano, J.A., et al.: Clasificación de documentos basada en la opinión: experimentos con un corpus de crıticas de cine en espanol. Procesamiento Lenguaje Nat. **41**, 73–80 (2008)
11. Cumbreras, M.Á.G., Cámara, E.M., et al.: TASS 2015 - The evolution of the Spanish opinion mining systems. Procesamiento Lenguaje Nat. **56**, 33–40 (2016)
12. DBPedia website. http://dbpedia.org/
13. Dong, Z., Dong, Q., Hao, C.: Hownet and its computation of meaning. In: Proceedings of COLING 2010: Demonstrations, pp. 53–56. ACL (2010)
14. Ekman, P., Friesen, W.V., Ellsworth, P.: Emotion in the Human Face: Guidelines for Research and an Integration of Findings. Pergamon Press (1972)
15. Emotion ML. https://www.w3.org/TR/emotionml/
16. Fleiss, J.L., Cohen, J.: The equivalence of weighted kappa and the intraclass correlation coefficient as measures of reliability. Educ. Psychol. Meas. **33**(3), 613–619 (1973)
17. Francisco, V., Gervás, P., Peinado, F.: Ontological reasoning to configure emotional voice synthesis. In: Marchiori, M., Pan, J.Z., Marie, C.S. (eds.) RR 2007. LNCS, vol. 4524, pp. 88–102. Springer, Heidelberg (2007). doi:10.1007/978-3-540-72982-2_7

18. Gangemi, A., Guarino, N., Masolo, C., Oltramari, A., Schneider, L.: Sweetening ontologies with DOLCE. In: Gómez-Pérez, A., Benjamins, V.R. (eds.) EKAW 2002. LNCS, vol. 2473, pp. 166–181. Springer, Heidelberg (2002). doi:10.1007/3-540-45810-7_18
19. Gil, R., Virgili-Gomá, J., et al.: Emotions ontology for collaborative modelling and learning of emotional responses. Comput. Hum. Behav. **51**, 610–617 (2015)
20. Grassi, M.: Developing HEO human emotions ontology. In: Fierrez, J., Ortega-Garcia, J., Esposito, A., Drygajlo, A., Faundez-Zanuy, M. (eds.) BioID 2009. LNCS, vol. 5707, pp. 244–251. Springer, Heidelberg (2009). doi:10.1007/978-3-642-04391-8_32
21. Hastings, J., Ceusters, W., et al.: The emotion ontology: enabling interdisciplinary research in the affective sciences. In: International and Interdisciplinary Conference on Modeling and Using Context, pp. 119–123 (2011)
22. Hastings, J., Ceusters, W., et al.: Annotating affective neuroscience data with the emotion ontology. In: Proceedings of the Workshop Towards an Ontology of Mental Functioning at ICBO, pp. 1–5 (2012)
23. Havas Media website. http://www.havasmedia.com/
24. Hepp, M.: GoodRelations: an ontology for describing products and services offers on the web. In: Gangemi, A., Euzenat, J. (eds.) EKAW 2008. LNCS, vol. 5268, pp. 329–346. Springer, Heidelberg (2008). doi:10.1007/978-3-540-87696-0_29
25. LPS BIGGER project website. http://www.cienlpsbigger.es/
26. Martínez-Cámara, E., Martín-Valdivia, M.T., et al.: Polarity classification for Spanish tweets using the COST corpus. J. Inf. Sci. **41**(3), 263–272 (2015)
27. Masquemedicos, with opinions in the medical domain. http://masquemedicos.com
28. Mohammad, S.M., Kiritchenko, S.: Using hashtags to capture fine emotion categories from tweets. Comput. Intell. **31**(2), 301–326 (2015)
29. Molina-González, M.D., Martínez-Cámara, E., et al.: Cross-domain sentiment analysis using Spanish opinionated words. In: Métais, E., Roche, M., Teisseire, M. (eds.) NLDB 2014. LNCS, pp. 214–219. Springer, Cham (2014). doi:10.1007/978-3-319-07983-7_28
30. Montejo-Ráez, A., Díaz-Galiano, M.C., et al.: Crowd explicit sentiment analysis. Knowl. Based Syst. **69**(1), 134–139 (2014)
31. MuchoCine, Spanish website with reviews about films. www.muchocine.net
32. Nakamura, A.: Kanjo Hyogen Jiten. Tokyodo Publishing (1993)
33. Nielsen: The social media report. http://blog.nielsen.com/nielsenwire/social/2012/
34. Obrenovic, Z., Garay, N., López, J.M., Fajardo, I., Cearreta, I.: An ontology for description of emotional cues. In: Tao, J., Tan, T., Picard, R.W. (eds.) ACII 2005. LNCS, vol. 3784, pp. 505–512. Springer, Heidelberg (2005). doi:10.1007/11573548_65
35. Plaza-Del-Arco, F.M., Martín-Valdivia, M.T., et al.: COPOS: Corpus of patient opinions in Spanish. Application of sentiment analysis techniques. Procesamiento Lenguaje Nat. **57**, 83–90 (2016)
36. Plutchik, R.: The nature of emotions: Human emotions have deep evolutionary roots (2001)
37. Ptaszynski, M., Rzepka, R., et al.: A robust ontology of emotion objects. In: Proceedings of the 18th Annual Meeting of the Association for Natural Language Processing, pp. 719–722 (2012)
38. Rangel, F., Rosso, P., Reyes, A.: Emotions and irony per gender in facebook. In: Proceedings of Workshop ES3LOD, LREC-2014, pp. 1–6 (2014)
39. Richins, M.L.: Measuring emotions in the consumption experience. J. Consum. Res. **24**(2), 127–146 (1997)

40. Roberts, K., Roach, M., Johnson, J.: EmpaTweet: annotating and detecting emotions on twitter. In: Proceedings of LREC 2012, pp. 3806–3813 (2012)
41. SAB corpus website. http://sabcorpus.linkeddata.es
42. Sam, K.M., Lei, P., Chatwin, C.: Ontology development for e-marketing mix model mapping with internet consumers' decision-making styles. In: Sobh, T. (ed.) Innovations and Advanced Techniques in Computer and Information Sciences and Engineering, pp. 279–282. Springer, Dordrecht (2007). doi:10.1007/978-1-4020-6268-1_50
43. Sánchez-Rada, J.F., Iglesias, C.A.: Onyx: a linked data approach to emotion representation. Inf. Process. Manag. **52**(1), 99–114 (2016)
44. Sánchez Rada, J.F., Torres, M., et al.: A linked data approach to sentiment and emotion analysis of twitter in the financial domain. In: 2nd International Workshop on Finance and Economics on the Semantic Web (2014)
45. Shaver, P., Schwartz, J., et al.: Emotion knowledge: further exploration of a prototype approach. J. Pers. Soc. Psychol. **52**(6), 1061–1086 (1987)
46. Spanish Corpus of reviews about films. http://www.sfu.ca/~mtaboada/research/SFU_Review_Corpus.html
47. Strapparava, C., Valitutti, A.: WordNet-Affect: an affective extension of WordNet. In: Proceedings of LREC, pp. 1083–1086 (2004)
48. The Emotions & Cognition Ontology. http://rhizomik.net/html/ontologies/emotions&cognitionontology/
49. Thomson Reuter's PermID website. https://permid.org/
50. Togias, K., Kameas, A.: An ontology-based representation of the twitter REST API. In: Proceedings of the IEEE 24th ICTAI, vol. 1, pp. 998–1003 (2012)
51. Tripadvisor website, with opinions on tourism. https://www.tripadvisor.es/
52. Twitter preprocessed datasets available at the TU Eindhoven. http://www.win.tue.nl/~mpechen/projects/smm/#Datasets
53. TwO, the Twitter Ontology. https://github.com/joshhanna/Twitter-Ontology
54. Vocabulary of the SAB corpus. http://sabcorpus.linkeddata.es/vocab
55. Westerski, A., Iglesias, C.A., Rico, F.T.: Linked opinions: describing sentiments on the structured web of data. In: Proceedings of the 4th International Workshop Social Data on the Web, vol. 830 (2011)
56. Yan, J., Bracewell, D.B., et al.: The creation of a Chinese emotion ontology based on HowNet. Eng. Lett. **16**(1), 166–171 (2008)