

# Do Enterprises Have Emotions?

Sven Buechel and Udo Hahn

Jena University Language & Information Engineering (JULIE) Lab

Friedrich-Schiller-Universität Jena

Jena, Germany

<http://www.julielab.de>

Jan Goldenstein and Sebastian G. M. Händschke and Peter Walgenbach

School of Economics and Business Administration

Friedrich-Schiller-Universität Jena

Jena, Germany

<http://www.orga.uni-jena.de>

## Abstract

Emotional language of human individuals has been studied for quite a while dealing with opinions and value judgments people have and share with others. In our work, we take a different stance and investigate whether large organizations, such as major industrial players, have and communicate emotions, as well. Such an anthropomorphic perspective has recently been advocated in management and organization studies which consider organizations as social actors. We studied this assumption by analyzing 1,676 annual business and sustainability reports from 90 top-performing enterprises in the United States, Great Britain and Germany. We compared the measurements of emotions in this homogeneous corporate text corpus with those from RCV1, a heterogeneous Reuters newswire corpus. From this, we gathered empirical evidence that business reports compare well with typical emotion-neutral economic news, whereas sustainability reports are much more emotionally loaded, similar to emotion-heavy sports and fashion news from Reuters. Furthermore, our data suggest that these emotions are distinctive and relatively stable over time per organization, thus constituting an emotional profile for enterprises.

## 1 Introduction

In the past years, we have witnessed an enormous upsurge of research activities in the field of NLP related to affective language use in social networks.

This work has mostly focused on subjective, often evaluative language of individual or informal ad-hoc groups of human actors in a multitude of social media platforms (Pang and Lee, 2008; Liu, 2015). Overall, this research analyzes language *about* organizations (e.g., opinions about the products they offer) rather than the language *of* organizations.

Quite recently, some areas of management and organization studies started modeling formal organizations in an anthropomorphic way, as social actors with human-like traits (King et al., 2010). Management and organization researchers following this approach claim that organizations have their own, persistent, human-like identity, play social roles and assume responsibility for their doings in the societies they are embedded in (Whetten, 2006). One of the far-reaching implications of granting organizations the status of social actors is, by default, their unequivocal submission to juridical standards as legal entities liable for violations of law in the same way as individual citizens are.

It is exactly this shift in the modeling of organizations (and thus business corporations, such as large enterprises) as social actors, which led us to our research questions. Within the new paradigm of computational social science (DiMaggio, 2015), we investigate whether organizations can also be attributed behavioral traits and properties typically associated with humans, such as attitudes, affects, emotions, etc.—a question which is, to the best of our knowledge, so far untackled despite its theoretical and empirical relevance (King et al., 2010).

## 2 Related Work

*Subjectivity analysis* is typically used as an umbrella term for NLP approaches concerned with all sorts of affective language use in which speakers' emotions play a crucial role. The most widespread subtask of subjectivity analysis is sentiment analysis or opinion mining (both terms are used interchangeably) (Pang and Lee, 2008). Most often, *sentiment* refers to the semantic orientation (or polarity), the positiveness or negativeness, of a sentence or a document. More recently, another subtask has attracted a lot of attention, namely emotion detection. Unlike (bipolar) sentiments, *emotion* describes a much more complex type of affective state typically associated with phenomena such as sadness, fear or joy. Yet its exact definition and distinction from other affective phenomena is an open issue (Munezero et al., 2014).

The more complex nature of emotion implies that both subtasks—sentiment analysis and emotion detection—need distinct analytic resources, especially lexicons. The number of general-language sentiment lexicons is, compared to the number of emotion lexicons, relatively large, including well-known resources such as SENTIWORDNET (Esuli and Sebastiani, 2006; Baccianella et al., 2010). Another notable resource is WORDNET-AFFECT which contains both, sentiment assessments (positive, negative, neutral and ambiguous) and a hierarchy of various emotion categories (Strapparava and Valitutti, 2004; Strapparava et al., 2006). Lately, however, an increasing number of emotion lexicons have been developed—within the fields of NLP (Mohammad and Turney, 2013; Staiano and Guerini, 2014), as well as cognitive psychology (Bestgen and Vincze, 2012; Warriner et al., 2013). As far as studies of sentiment in the economic area are concerned, Loughran and McDonald (2011) adapted the *Harvard Psychosociological Dictionary* to better fit the word usage of the finance domain. The resulting resource comprises six word lists (two of which refer to positive and negative words), thus forming a finance-specific sentiment lexicon.

Researchers in NLP and cognitive psychology have devised a multitude of different models of emotion which can be roughly subdivided into categorical and dimensional models (Scherer, 2000; Calvo and Kim, 2013). In computational studies, categor-

ical models most often employ Ekman's (1992) six basic emotions or a derivative therefrom. According to this psychological theory, all human beings share a common set of (basic) emotions so that each emotional state of an individual can be unambiguously classified as one of these. Dimensional approaches, on the other hand, often refer to Russell and Mehrabian's (1977) Valence-Arousal-Dominance (VAD) model. According to this model, emotional states can be described relative to three fundamental emotional dimensions: valence (the degree of pleasure of an emotion), arousal (level of mental activity, ranging from low engagement to ecstasy) and dominance (extent of control felt in a given situation). Accordingly, emotions are characterized on three dimensions, each of which spans an interval of real-valued numbers indicating the strength and orientation of each dimension. Hence, other than in categorical approaches with a usually small (up to nine) and finite number of states, an infinite number of emotional states can be represented in dimensional approaches.

Subjective language use has also been a topic of interest in the business and economy domain. A common theme is here whether the linguistic 'tone' of public mass media's coverage of enterprises or publications authored by individual enterprises themselves are indicative of the companies' future economic performance, and can thus, e.g., be useful for stock trading strategies (for a recent survey, cf. Nassirtoussi et al. (2014)).

Early work on subjective language use in newspapers as predictors for companies' accounting earnings and stock returns (Tetlock et al., 2008) or in quarterly company reports for stock market performance prediction (Kloptchenko et al., 2004) reveals that such linguistic analysis uncover otherwise hard-to-quantify aspects of firms' fundamentals. The potential of news-driven sentiment analysis is elucidated in a recent study by Uhl (2014) who finds evidence that polarity measurements on Reuters news can explain and predict changes in stock returns better than macroeconomic factors, in isolation.

Bollen et al. (2011) examine to what extent subjective language on the social media platform TWITTER can be interpreted as a predictor of stock market prizes. They find that the predictive power of certain emotion categories exceeds that of the semantic orientation, leading to a significant improvement of

basic stock market forecasting models. Their study reveals the possible benefits of the additional information emotion detection might contribute in contrast to (less expressive) sentiment analysis.

In a similar vein, Génereux et al. (2011) investigate the impact of financial news items on the stock price of companies. They treat short financial news snippets about companies as if they were carrying implicit sentiment about the future market direction made explicit by the vocabulary they employ. They investigate how this sentiment vocabulary can be automatically extracted from texts and subsequently be used for classification. This resembles previous work by Devitt and Ahmad (2007) who explored a computable metric of positive or negative polarity in financial news text which is consistent with human judgments and can be used in a quantitative analysis of news sentiment impact on financial markets.

Unlike sentiment analysis of informal communication pieces from social or mass media material, Kogan et al. (2009) were the first in company-centric text analysis to focus on official statements from enterprises on a larger scale (pioneering work was conducted by Kloptchenko et al. (2004), as mentioned above). They built up the *10-K Corpus*, a collection of 54,379 annual business reports (from 10,492 different publicly traded companies) published over the period from 1996 to 2006. They exploited this corpus to predict the volatility of stock returns, an established empirical measure of financial risk, using regression models. The economic assessments are derived from the distribution of unigrams and bigrams in the reports incorporating TF and TF-IDF-based measures for a bag-of-words (BOW) model.

Wang et al. (2013), in a follow-up study, reused the 10-K Corpus by incorporating Loughran and McDonald's (2011) finance-specific sentiment dictionary. The models they learn suggest strong correlations between financial sentiment words and risk in terms of stock return volatility. Using basically the same experimental set-up as the former study, Hájek et al. (2013) go one step further and demonstrate that by combining qualitative sentiment information of annual reports with quantitative financial indicators (e.g. market capitalization, profitability, etc.) the resulting stock price forecasting model is more accurate than using quantitative indicators alone. These findings are further supported by

Kazemian et al. (2014) who present empirical data which indicate to act cautiously with respect to stock trading strategies based on sentiment analysis of linguistic data only (they propose to consider actual market returns, in addition). Interestingly, the study by Hájek et al. (2013) points out that the *change* in sentiment (rather than its specific value in some point of time) seems to be an important determinant of stock price development in the long run.

### 3 Experimental Set-up

In order to test the assumption whether enterprises can be attributed an emotional status as part of their identity as social actors, we, first, compiled a text corpus composed of annual business and sustainability reports of the thirty top-performing corporations in the United States, Great Britain and Germany. These enterprises were selected based on their listing in the Dow Jones, the FTSE 100, and the DAX Index, respectively. The global economic power of these corporations is quite impressive since, in 2014, their revenues comprised about 7% of the worldwide GDP, an equivalent of 5.4 trillion USD.

For these 90 corporations, we collected the English versions of all annual business and sustainability reports (also called corporate social responsibility (CSR) reports (Matten and Moon, 2008)) available online. This corpus amounts to 1,676 reports (3,798,070 sentences; 128,145,063 text tokens) divided into 1,087 annual and 589 sustainability reports covering the time span from 1992 to 2015.

We compared business reports (which we thought to be emotionally mostly neutral) to sustainability reports (which we thought to be much more emotionally loaded). This comparison is based on a simple vectorial BOW text analysis engine whose novel feature is an attached emotion word lexicon composed of 13,915 lexical entries (Warriner et al., 2013). Each of its lexical entries carries empirically determined (crowd-sourced) real-valued assessments for the three fundamental emotion dimensions of the VAD model (cf. Table 1 for illustrative lexicon entries and associated VAD values).

So far, this model—though among the most popular models of emotion in cognitive psychology—has found only very scant attention in the field of text-based subjectivity analysis (Calvo and Kim, 2013).

Lemma	V	A	D
sunshine	3.14	0.32	0.43
leukemia	-3.53	0.75	-2.17
terrorism	-3.40	2.42	-2.31
calm	1.89	-3.33	2.44
successful	2.76	0.08	2.71
uncontrollable	-1.16	0.81	-2.82

**Table 1:** Lexicon entries from the VAD lexicon (adapted via scale transformation from Warriner et al. (2013)) with considerably high (low) valence, arousal or dominance values.

In the lexicon used in our experiments, valence, arousal and dominance are balanced in the interval of  $[-4, 4]$  (in the original VAD lexicon (Warriner et al., 2013), values range in the interval of  $[1, 9]$ ).

The basic processing cycle is as follows: A term-vector representation (stop word-free, lemmatized) is built up for each of the 1,676 reports. Each term component of the 1,676 vectors is matched with all of the 13,915 entries of the VAD lexicon. If a match occurs, i.e., the lexical item at the  $i$ -th vector component matches an entry in the VAD lexicon, the 3D emotion values of that item (cf. Table 1) are multiplied with the number of occurrences of this item in the underlying document. Finally, the emotion values for all matched entries per document are summed up and divided by the number of matches. This average yields the *emotion value* of a given document on the three principle emotional dimensions of the VAD scheme.

## 4 Results of the Experiments

In general, we found valence and dominance values around 0.6 and arousal values at  $-1.0$ . The fact that the latter is considerably lower than the former two is partially derived from the statistical properties of the emotion word lexicon we employ. Another noteworthy characteristic is the strong correlation between the valence and the dominance values ( $r = 0.87$ ) as shown in Figure 1(b). This correlation can also be observed in the lexicon as well as in other corpora (see below) although it is pronouncedly higher in the enterprise corpus. The correlation’s strength might be a suitable measure for the appreciation of control as expressed in the individual documents. This is so because high correlation hints at frequent co-occurrences of words having high (low) valence ratings and words having

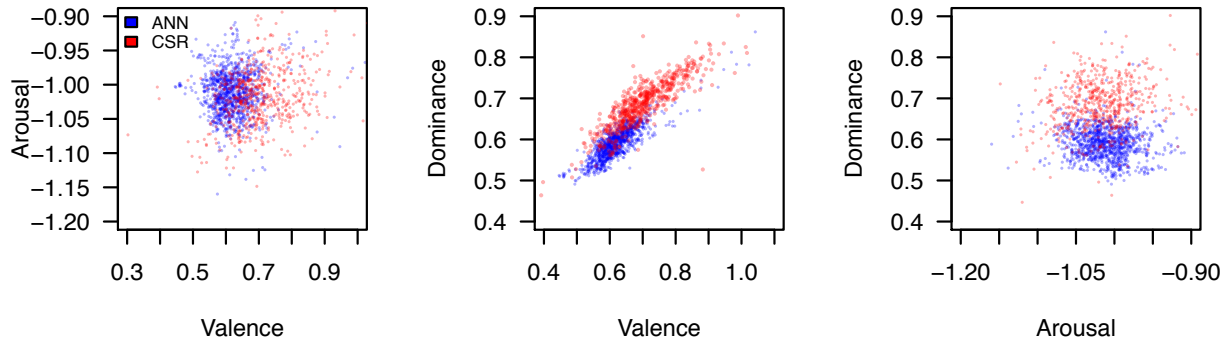
high (low) dominance ratings or a high frequency of words having both, high (low) valence and high (low) dominance ratings. According to this, the results can be interpreted as pointing out the great concernment corporations place on control.

Next, we investigated the role of the two text genres. First of all, in Figure 1, both genres can clearly be distinguished, in particular in Figures 1(b) and 1(c). We find that sustainability reports are on average way more positive (valence: 0.70) and dominant (0.68) than annual reports (0.62 and 0.59, respectively). The differences in both emotional dimensions are significant (two-tailed t-test:  $p < 0.001$ ) and show large effect sizes (Cohen’s  $d > |0.7|$ ). This tendency may be due to the fact that annual reports have to comply with numerous external constraints corporations have to respect (e.g., themes and data required by legal prescriptions, etc.), whereas in sustainability reports corporations have full control over phrasing and contents and may focus on their voluntary activities.

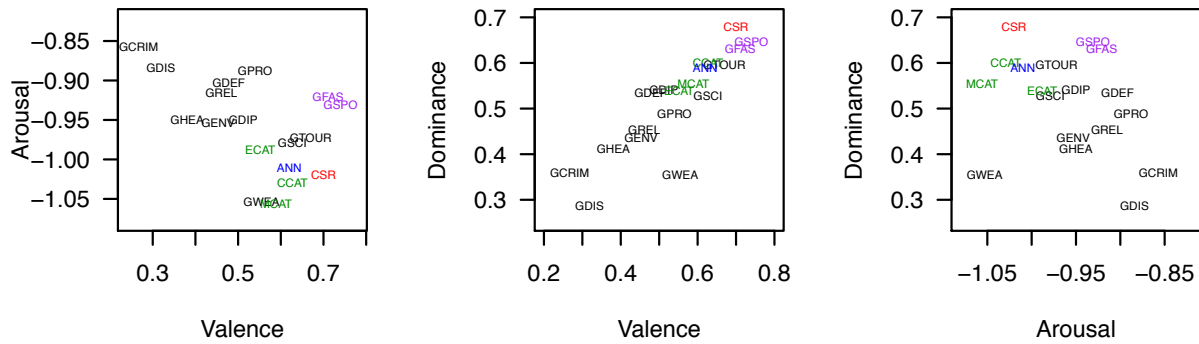
To relate these findings to external resources and thus gather evidence for the validity of our approach, we compared our results from the enterprise data set to those from a Reuters corpus which contains newswire material (RCV1 (Lewis et al., 2004)). We found slightly lower valence and dominance values than in the enterprise data but standard deviation in the corporation corpus was half as high as in RCV1—this points out that the enterprise corpus is emotionally more homogeneous than Reuters.

Figure 2 depicts the localization of the two enterprise text genres relative to several newswire categories from the Reuters corpus in the VAD space. As can be seen, the enterprise documents compare well with Reuters business- and economy-related categories (in green), CCAT in particular, which contains corporate and industrial topics. The emotional values of business reports are thus shown to come close to that of newspaper articles of the corresponding content categories—these similarities seem to indicate that the measures we use are valid.

On the arousal dimension, sustainability reports reveal only a minor distance relative to Reuters business and economic categories and the enterprises’ annual reports. Yet, on the valence and dominance dimensions, they correspond very closely to highly emotional fashion and sports news (GFAS



**Figure 1:** Scatter plots of the emotional values of the 1,676 enterprise reports—1,087 annual reports (ANN: blue) and 589 sustainability reports (CSR: red)—of the corporation corpus in the VAD space.



**Figure 2:** Average emotional values for the categories of the RCV1 (business- and economy-related categories (CCAT, ECAT, and MCAT) in green, sports and fashion category (GSPO and GFAS, respectively) in purple) as well as both genres from the enterprise corpus (annual business reports (ANN) in blue, sustainability reports (CSR) in red).

and GSPO, respectively; both in purple). This can be taken as further evidence for the dissimilarity of the two corporation genres mentioned before. The dominance value of sustainability reports is on average greater than that of *all* news categories.

Furthermore, we could show that the reports (in each case with respect to the authoring company) all share a specific tendency in their emotion value which is even relatively constant over time. This is especially true when examining the two subcorpora of annual and sustainability reports separately. In this case, the proportion of explained variance with consideration of the corporation which authored a report reaches values of about 70% (for arousal in annual reports, data are available in Büchel (2016)).

## 5 Conclusion

In summary, our research provides one of the first attempts to study emotional factors in documents representing large corporations—as reflected in the enterprises’ annual and sustainability reports—rather

than individuals. In comparison with economic newswire material from the RCV1 corpus, we located enterprise documents on three fundamental emotional dimensions, namely valence, arousal and dominance (according to the VAD model), and found strong evidence for particularly high dominance in sustainability reports. Furthermore, the data indicate that organizations exhibit a distinctive and persistent emotional profile. So, indeed, we have reasons to believe that—in the light of their reporting—enterprises have emotions in the sense of an anthropomorphic model and that this profile contributes to a unique organizational identity.

Technically, to the best of our knowledge, the VAD-based emotion lexicon (with >10k entries) we employed for our study has never been used for text analytics tasks before. Note that this lexicon exceeds well-known resources with a comparable emotion model (Bradley and Lang, 1999) by an order of magnitude.

## References

- Stefano Baccianella, Andrea Esuli, and Fabrizio Sebastiani. 2010. SENTIWORDNET 3.0: An enhanced lexical resource for sentiment analysis and opinion mining. In Nicoletta Calzolari, Khalid Choukri, Bente Maegaard, Joseph Mariani, Jan Odijk, Stelios Piperidis, and Daniel Tapias, editors, *LREC 2010 — Proceedings of the 7th International Conference on Language Resources and Evaluation. La Valletta, Malta, May 17-23, 2010*, pages 2200–2204. European Language Resources Association (ELRA).
- Yves Bestgen and Nadja Vincze. 2012. Checking and bootstrapping lexical norms by means of word similarity indexes. *Behavior Research Methods*, 44(4):998–1006.
- Johan Bollen, Huina Mao, and Xiaojun Zeng. 2011. Twitter mood predicts the stock market. *Journal of Computational Science*, 2(1):1–8.
- Margaret M. Bradley and Peter J. Lang. 1999. Affective Norms for English Words (ANEW): Stimuli, Instruction Manual and Affective Ratings. Technical Report Technical Report C-1, The Center for Research in Psychophysiology, University of Florida, Gainesville, FL.
- Sven Eric Büchel. 2016. Automatische Analyse von Emotionen in Geschäfts- und Nachhaltigkeitsberichten. Bachelor thesis, Friedrich-Schiller-Universität Jena, Jena, Germany.
- Rafael A Calvo and Sunghwan Mac Kim. 2013. Emotions in text: Dimensional and categorical models. *Computational Intelligence*, 29(3):527–543.
- Ann Devitt and Khurshid Ahmad. 2007. Sentiment polarity identification in financial news: A cohesion-based approach. In *ACL 2007 — Proceedings of the 45th Annual Meeting of the Association for Computational Linguistics. Prague, Czech Republic, June 23-30, 2007*, pages 984–991. Association for Computational Linguistics (ACL).
- Paul DiMaggio. 2015. Adapting computational text analysis to social science (and vice versa). *Big Data & Society*, 2(2):1–5.
- Paul Ekman. 1992. An argument for basic emotions. *Cognition and Emotion*, 6(3-4):169–200.
- Andrea Esuli and Fabrizio Sebastiani. 2006. SENTIWORDNET: A publicly available lexical resource for opinion mining. In Nicoletta Calzolari, Aldo Gangemi, Bente Maegaard, Joseph Mariani, Jan Odijk, and Daniel Tapias, editors, *LREC 2006 — Proceedings of the 5th International Conference on Language Resources and Evaluation. Genoa, Italy, 22-28 May, 2006*, pages 417–422. European Language Resources Association (ELRA).
- Michel Génèreux, Thierry Poibeau, and Moshe Koppel. 2011. Sentiment analysis using automatically labelled financial news. In Khurshid Ahmad, editor, *Affective Computing and Sentiment Analysis. Emotion, Metaphor and Terminology*, number 45 in Text, Speech and Language Technology, pages 111–126. Springer.
- Petr Hájek, Vladimír Olej, and Renáta Myšková. 2013. Forecasting stock prices using sentiment information in annual reports: A neural network and support vector regression approach. *WSEAS Transactions on Business and Economics*, 10(4):293–305.
- Siavash Kazemian, Shunan Zhao, and Gerald Penn. 2014. Evaluating sentiment analysis evaluation: A case study in securities trading. In *WASSA 2014 — Proceedings of the 5th Workshop on Computational Approaches to Subjectivity, Sentiment and Social Media Analysis @ ACL 2014. Baltimore, Maryland, USA, June 27, 2014*, pages 119–127.
- Brayden G. King, Teppo Felin, and David A. Whetten. 2010. Finding the organization in organizational theory. A meta-theory of the organization as a social actor. *Organization Science*, 21(1):290–305.
- Antonina Kloptchenko, Tomas Eklund, Barbro Back, Jonas Karlsson, Hannu Vanharanta, and Ari Visa. 2004. Combining data and text mining techniques for analysing financial reports. *Intelligent Systems in Accounting, Finance and Management*, 12(1):29–41.
- Shimon Kogan, Dimitry Levin, Bryan R. Routledge, Jacob S. Sagi, and Noah A. Smith. 2009. Predicting risk from financial reports with regression. In *NAACL-HLT 2009 — Proceedings of Human Language Technologies: The 2009 Annual Conference of the North American Chapter of the Association for Computational Linguistics. Boulder, CO, USA, May 31 - June 5, 2009*, volume 1, pages 272–280, Stroudsburg/PA. Association for Computational Linguistics (ACL).
- David D. Lewis, Yiming Yang, Tony G. Rose, and Fan Li. 2004. RCV1: A new benchmark collection for text categorization research. *Journal of Machine Learning Research*, 5:361–397, April.
- Bing Liu. 2015. *Sentiment Analysis: Mining Opinions, Sentiments and Emotions*. Cambridge University Press, New York, NY.
- Tim Loughran and Bill McDonald. 2011. When is a liability not a liability? Textual analysis, dictionaries, and 10-Ks. *The Journal of Finance*, 66(1):35–65.
- Dirk Matten and Jeremy Moon. 2008. ‘Implicit’ and ‘explicit’ CSR: A conceptual framework for a comparative understanding of Corporate Social Responsibility. *Academy of Management Review*, 33(2):404–424.
- Saif M. Mohammad and Peter D. Turney. 2013. Crowdsourcing a word-emotion association lexicon. *Computational Intelligence*, 29(3):436–465.
- Myriam D. Munezero, Calkin Suero Montero, Erkki Sutinen, and John Pajunen. 2014. Are they different?

- Affect, feeling, emotion, sentiment, and opinion detection in text. *IEEE Transactions on Affective Computing*, 5(2):101–111.
- Arman Khadjeh Nassirtoussi, Saeed Aghabozorgi, Teh Ying Wah, and David Chek Ling Ngo. 2014. Text mining for market prediction: A systematic review. *Expert Systems with Applications*, 41(16):7653–7670.
- Bo Pang and Lillian Lee. 2008. Opinion mining and sentiment analysis. *Foundation of Trends in Information Retrieval*, 2(1-2):1–135.
- James A Russell and Albert Mehrabian. 1977. Evidence for a three-factor theory of emotions. *Journal of Research in Personality*, 11(3):273–294.
- Klaus R. Scherer. 2000. Psychological models of emotion. In Joan C. Borod, editor, *The Neuropsychology of Emotion*, pages 137–162. Oxford University Press, Oxford, U.K.; New York, N.Y.
- Jacopo Staiano and Marco Guerini. 2014. DEPECHE MOOD: A lexicon for emotion analysis from crowd annotated news. In *ACL 2014 — Proceedings of the 52nd Annual Meeting of the Association for Computational Linguistics. Baltimore, Maryland, USA, June 22-27, 2014*, volume 2: Short Papers, pages 427–433, Stroudsburg/PA. Association for Computational Linguistics (ACL).
- Carlo Strapparava and Alessandro Valitutti. 2004. WORDNET-AFFECT: An affective extension of WORDNET. In Maria Teresa Lino, Maria Francisca Xavier, Fátima Ferreira, Rute Costa, and Raquel Silva, editors, *LREC 2004 — Proceedings of the 4th International Conference on Language Resources and Evaluation. In Memory of Antonio Zampolli. Lisbon, Portugal, 24-30 May, 2004*, pages 1083–1086. European Language Resources Association (ELRA).
- Carlo Strapparava, Alessandro Valitutti, and Oliviero Stock. 2006. The affective weight of lexicon. In Nicoletta Calzolari, Aldo Gangemi, Bente Maegaard, Joseph Mariani, Jan Odijk, and Daniel Tapias, editors, *LREC 2006 — Proceedings of the 5th International Conference on Language Resources and Evaluation. Genoa, Italy, 22-28 May, 2006*, pages 423–426. European Language Resources Association (ELRA).
- Paul C. Tetlock, Maytal Saar-Tsechansky, and Sofus Macskassy. 2008. More than words: Quantifying language to measure firms’ fundamentals. *The Journal of Finance*, 63(3):1437–1467.
- Matthias W. Uhl. 2014. Reuters sentiment and stock returns. *Journal of Behavioral Finance*, 15(4):287–298.
- Chuan-Ju Wang, Ming-Feng Tsai, Tse Liu, and Chin-Ting Chang. 2013. Financial sentiment analysis for risk prediction. In *IJCNLP 2013 — Proceedings of the 6th International Joint Conference on Natural Language Processing. Nagoya, Japan, 14-18 October 2013*, pages 802–808. Asian Federation of Natural Language Processing (AFNLP).
- Amy Beth Warriner, Victor Kuperman, and Marc Brysbaert. 2013. Norms of valence, arousal, and dominance for 13,915 English lemmas. *Behavior Research Methods*, 45(4):1191–1207.
- David A. Whetten. 2006. Albert and Whetten revisited: Strengthening the concept of organizational identity. *Journal of Management Inquiry*, 15(3):219–234.