# SenticNet 2: A Semantic and Affective Resource for Opinion Mining and Sentiment Analysis

**Erik Cambria**
National University of Singapore
Singapore, 117411 Singapore
cambria@nus.edu.sg

**Catherine Havasi**
MIT Media Lab
Cambridge, 02139-4307 USA
havasi@media.mit.edu

**Amir Hussain**
University of Stirling
Stirling, FK4 9LA UK
ahu@cs.stir.ac.uk

### Abstract

Web 2.0 has changed the ways people communicate, collaborate, and express their opinions and sentiments. But despite social data on the Web being perfectly suitable for human consumption, they remain hardly accessible to machines. To bridge the cognitive and affective gap between word-level natural language data and the concept-level sentiments conveyed by them, we developed SenticNet 2, a publicly available semantic and affective resource for opinion mining and sentiment analysis. SenticNet 2 is built by means of sentic computing, a new paradigm that exploits both AI and Semantic Web techniques to better recognize, interpret, and process natural language opinions. By providing the semantics and sentics (that is, the cognitive and affective information) associated with over 14,000 concepts, SenticNet 2 represents one of the most comprehensive semantic resources for the development of affect-sensitive applications in fields such as social data mining, multimodal affective HCI, and social media marketing.

## Introduction

As the Web plays a more and more significant role in people's social lives, it contains more and more information concerning their opinions and sentiments. The distillation of knowledge from this huge amount of unstructured information, also known as opinion mining and sentiment analysis, is a task that has recently raised more and more interest for purposes such as marketing, customer service, and financial market prediction. Such a task, however, is of an extremely difficult nature as web-contents (and social media contents in particular) today are perfectly suitable for human consumption but they remain hardly accessible to machines.

The Web, in fact, mostly owes its success to the development of search engines like Google and Yahoo, which represent the starting point for information retrieval. Such engines, which base their searches on keyword-based algorithms relying on the textual representation of the web-page, are very good in retrieving texts, splitting them into parts, checking the spelling, and counting their words. But when it comes to interpreting sentences and extracting useful information for users, they still have to face a lot of limitations.



Current attempts to perform automatic understanding of text, e.g., textual entailment (Dagan, Glickman, and Magnini 2006) and machine reading (Etzioni, Banko, and Cafarella 2006), still suffer from numerous problems including inconsistencies, synonymy, polysemy, and entity duplication, as they focus on a mere syntactical analysis of text. To help natural language processing (NLP) researchers, human computer interaction (HCI) designers, and marketers bridge the cognitive and affective gap between word-level natural language data and the concept-level opinions and sentiments conveyed by them, we developed SenticNet 2, a publicly available semantic and affective resource for opinion mining and sentiment analysis.

SenticNet 2 is built by means of sentic computing (Cambria and Hussain 2012), a novel paradigm that exploits both computer and social sciences to better recognize, interpret, and process opinions and sentiments over the Web. Differently from SenticNet (Cambria et al. 2010c), which simply associates a polarity value to about 5,700 concepts from the Open Mind corpus, SenticNet 2 provides the semantics and sentics (that is, the cognitive and affective information) associated with over 14,000 concepts, allowing a deeper and more multi-faceted analysis of natural language text.

The rest of this paper is organized as follows: the first section is a brief overview of main approaches to opinion mining; the following section describes techniques and models employed in this work; the next two sections explain how SenticNet 2 is built and how to work with it, respectively; after a section on evaluation, finally, some concluding remarks and future work recommendations are made.

## Online Opinions and Sentiments

Existing approaches to automatic identification and extraction of opinions and sentiments from text can be grouped into three main categories: keyword spotting, in which text is classified into categories based on the presence of fairly unambiguous affect words (Elliott 1992; Wiebe, Wilson, and Cardie 2005), lexical affinity, which assigns arbitrary words a probabilistic affinity for a particular topic or emotion (Wilson, Wiebe, and Hoffmann 2005; Somasundaran, Wiebe, and Ruppenhofer 2008; Rao and Ravichandran 2009; Stevenson, Mikels, and James 2007; Bradley and Lang 1999), and statistical methods, which calculate the valence of keywords, punctuation, and word co-occurrence frequencies on the base of a large training corpus (Turney and Littman 2003; Hu and Liu 2004; Pang and Lee 2005; Abbasi, Chen, and Salem 2008; Velikovich et al. 2010). These approaches mainly rely on parts of text in which opinions and sentiments are explicitly expressed such as polarity terms (e.g., good, bad, nice, nasty, excellent, poor) and affect words (e.g., happy, sad, calm, angry, interested, bored).

More generally, however, opinions and sentiments are expressed implicitly through context and domain dependent concepts, which make purely syntactical approaches ineffective. To overcome such a problem, we need to use NLP techniques that rely on semantics and sentics, rather than syntactics. Within this work, in particular, we exploit sentic computing to infer the cognitive and affective information associated with common sense concepts and, hence, build one of the most comprehensive semantic resources for opinion mining and sentiment analysis.

## Sentic Computing

Sentic computing is a multi-disciplinary approach to opinion mining and sentiment analysis at the crossroads between affective computing (Picard 1997) and common sense computing (Cambria et al. 2009b), which exploits both computer and social sciences to better recognize, interpret, and process opinions and sentiments over the Web. In particular, sentic computing involves the use of AI and Semantic Web techniques, for knowledge representation and inference; mathematics, for carrying out tasks such as graph mining and multi-dimensionality reduction; linguistics, for discourse analysis and pragmatics; psychology, for cognitive and affective modeling; sociology, for understanding social network dynamics and social influence; finally ethics, for understanding related issues about the nature of mind and the creation of emotional machines. Unlike statistical classification, which generally requires large inputs and thus cannot appraise texts with satisfactory granularity, sentic computing enables the analysis of documents not only at page and paragraph-level but also at sentence and clause-level.

In this work, we exploit the ensemble application of spectral association (Havasi, Speer, and Holmgren 2010), an approximation of many steps of spreading activation, and CF-IOF (Cambria et al. 2010a), an approach similar to TF-IDF weighting, to extract semantics from ConceptNet (Havasi, Speer, and Alonso 2007), a semantic network of common sense knowledge. The extraction of sentics, in turn, is performed through the combined use of AffectiveSpace (Cambria et al. 2009a), a multi-dimensional vector space representation of affective common sense knowledge, and the Hourglass of Emotions (Cambria et al. 2010b), a brain-inspired emotion categorization model.

### Spectral Association

Spectral association is a technique that involves assigning activations to 'seed concepts' and applying an operation that spreads their values across the graph structure of ConceptNet. This operation transfers the most activation to concepts that are connected to the key concepts by short paths or many different paths in common sense knowledge.

In particular, we build a matrix $C$ that relates concepts to other concepts, instead of their features, and add up the scores over all relations that relate one concept to another, disregarding direction. Applying $C$ to a vector containing a single concept spreads that concept's value to its connected concepts. Applying $C^2$ spreads that value to concepts connected by two links (including back to the concept itself).

As we aim to spread the activation through any number of links, with diminishing returns, the operator we want is:

$$1 + C + \frac{C^2}{2!} + \frac{C^3}{3!} + ... = e^C$$

We can calculate this odd operator, $e^C$, because we can factor $C$. $C$ is already symmetric, so instead of applying Lanczos' method to $CC^T$ and getting the singular value decomposition (SVD), we can apply it directly to $C$ and get the spectral decomposition $C = V\Lambda V^T$. As before, we can raise this expression to any power and cancel everything but the power of $\Lambda$. Therefore, $e^C = Ve^{\Lambda}V^T$. This simple twist on the SVD lets us calculate spreading activation over the whole matrix instantly. We can truncate this matrix to $k$ axes and therefore save space while generalizing from similar concepts. We can also rescale the matrix, so that activation values have a maximum of 1 and do not tend to collect in highly-connected concepts, by normalizing the truncated rows of $Ve^{\Lambda/2}$ to unit vectors, and multiplying that matrix by its transpose to get a rescaled version of $Ve^{\Lambda}V^T$.

### CF-IOF Weighting

CF-IOF (concept frequency – inverse opinion frequency) is a technique that identifies common topic-dependent semantics in order to evaluate how important a concept is to a set of opinions concerning the same topic. It is hereby used to feed spectral association with 'seed concepts'.

Firstly, the frequency of a concept $c$ for a given domain $d$ is calculated by counting the occurrences of the concept $c$ in the set of available $d$-tagged opinions and dividing the result by the sum of number of occurrences of all concepts in the set of opinions concerning $d$. This frequency is then multiplied by the logarithm of the inverse frequency of the concept in the whole collection of opinions, that is:

$$CF\text{-}IOF_{c,d} = \frac{n_{c,d}}{\sum_k n_{k,d}} \log \sum_k \frac{n_k}{n_c}$$

where $n_{c,d}$ is the number of occurrences of concept $c$ in the set of opinions tagged as $d$, $n_k$ is the total number of concept occurrences and $n_c$ is the number of occurrences of $c$ in the whole set of opinions. A high weight in CF-IOF is reached by a high concept frequency (in the given opinions) and a low opinion frequency of the concept in the whole collection of opinions. Therefore, thanks to CF-IOF weights, it is possible to filter out common concepts and detect relevant topic-dependent semantics.

### AffectiveSpace

To extract sentics from natural language text, we use AffectiveSpace, a multi-dimensional vector space built upon ConceptNet and WordNet-Affect (WNA) (Strapparava and

Valitutti 2004), a linguistic resource for the lexical representation of affective knowledge. The alignment operation operated over ConceptNet and WNA yields a matrix, $A$, in which common sense and affective knowledge coexist, i.e., a matrix 14,301 $\times$ 117,365 whose rows are concepts (e.g., 'dog' or 'bake cake'), whose columns are either common sense and affective features (e.g., 'isA-pet' or 'hasEmotion-joy'), and whose values indicate truth values of assertions.

Therefore, in $A$, each concept is represented by a vector in the space of possible features whose values are positive for features that produce an assertion of positive valence (e.g., 'a penguin is a bird'), negative for features that produce an assertion of negative valence (e.g., 'a penguin cannot fly') and zero when nothing is known about the assertion. The degree of similarity between two concepts, then, is the dot product between their rows in $A$. The value of such a dot product increases whenever two concepts are described with the same feature and decreases when they are described by features that are negations of each other. In particular, we use truncated SVD (Wall, Rechtsteiner, and Rocha 2003) in order to obtain a new matrix containing both hierarchical affective knowledge and common sense.

The resulting matrix has the form $\tilde{A} = U_k \Sigma_k V_k^T$ and is a low-rank approximation of $A$, the original data. This approximation is based on minimizing the Frobenius norm of the difference between $A$ and $\tilde{A}$ under the constraint $rank(\tilde{A}) = k$. For the Eckart–Young theorem (Eckart and Young 1936) it represents the best approximation of $A$ in the least-square sense, in fact:

$$\begin{aligned} \min_{\tilde{A}|rank(\tilde{A})=k} |A - \tilde{A}| &= \min_{\tilde{A}|rank(\tilde{A})=k} |\Sigma - U^* \tilde{A} V| \\ &= \min_{\tilde{A}|rank(\tilde{A})=k} |\Sigma - S| \end{aligned}$$

assuming that $\tilde{A}$ has the form $\tilde{A} = USV^*$, where $S$ is diagonal. From the rank constraint, i.e., $S$ has $k$ non-zero diagonal entries, the minimum of the above statement is obtained as follows:

$$\min_{\tilde{A}|rank(\tilde{A})=k} |\Sigma - S| = \min_{s_i} \sqrt{\sum_{i=1}^{n} (\sigma_i - s_i)^2} =$$

$$= \min_{s_i} \sqrt{\sum_{i=1}^{k} (\sigma_i - s_i)^2 + \sum_{i=k+1}^{n} \sigma_i^2} = \sqrt{\sum_{i=k+1}^{n} \sigma_i^2}$$

Therefore, $\tilde{A}$ of rank $k$ is the best approximation of $A$ in the Frobenius norm sense when $\sigma_i = s_i$ $(i = 1, ..., k)$ and the corresponding singular vectors are the same as those of $A$. If we choose to discard all but the first $k$ principal components, common sense concepts and emotions are represented by vectors of $k$ coordinates: these coordinates can be seen as describing concepts in terms of 'eigenmoods' that form the axes of AffectiveSpace, i.e., the basis $e_0$,...,$e_{k-1}$ of the vector space. For example, the most significant eigenmood, $e_0$, represents concepts with positive affective valence. That is, the larger a concept's component in the $e_0$ direction is, the more affectively positive it is likely to be.

Concepts with negative $e_0$ components, then, are likely to have negative affective valence. Thus, by exploiting the information sharing property of truncated SVD, concepts with the same affective valence are likely to have similar features – that is, concepts conveying the same emotion tend to fall near each other in AffectiveSpace. Concept similarity does not depend on their absolute positions in the vector space, but rather on the angle they make with the origin. For example we can find concepts such as 'beautiful day', 'birthday party', 'laugh' and 'make person happy' very close in direction in the vector space, while concepts like 'sick', 'feel guilty', 'be laid off' and 'shed tear' are found in a completely different direction (nearly opposite with respect to the centre of the space).

### The Hourglass of Emotions

To reason on the disposition of concepts in AffectiveSpace, we use the Hourglass of Emotions, an affective categorization model developed starting from Plutchik's studies on human emotions (Plutchik 2001). In the model, sentiments are reorganized around four independent dimensions whose different levels of activation make up the total emotional state of the mind. The Hourglass of Emotions, in fact, is based on the idea that the mind is made of different independent resources and that emotional states result from turning some set of these resources on and turning another set of them off (Minsky 2006). Each such selection changes how we think by changing our brain's activities: the state of 'anger', for example, appears to select a set of resources that help us react with more speed and strength while also suppressing some other resources that usually make us act prudently.

The primary quantity we can measure about an emotion we feel is its strength. But when we feel a strong emotion it is because we feel a very specific emotion. And, conversely, we cannot feel a specific emotion like 'fear' or 'amazement' without that emotion being reasonably strong. Mapping this space of possible emotions leads to an hourglass shape. In the model, affective states are not classified, as often happens in the field of emotion analysis, into basic emotional categories, but rather into four concomitant but independent dimensions, characterized by six levels of activation, which determine the intensity of the expressed/perceived emotion as a $float \in$ [-1,+1]. Such levels are also labeled as a set of 24 basic emotions (six for each of the affective dimensions) in a way that allows the model to specify the affective information associated with text both in a dimensional and in a discrete form.

## Building SenticNet 2

Currently available lexical resources for opinion polarity and affect recognition such as SentiWordNet (Esuli and Sebastiani 2006) or WNA are known to be pretty noisy and limited. These resources, in fact, mainly provide opinion polarity and affective information at syntactical level, leaving out polarity and affective information for common sense knowledge concepts such as 'accomplish goal', 'bad feeling', 'celebrate special occasion', 'lose temper' or 'be on cloud nine', which are usually found in natural language text to express viewpoints and affect.

In a previous work we developed SenticNet, a publicly available semantic resource for opinion mining that associates a polarity value to about 5,700 common sense knowledge concepts from the Open Mind corpus. Evaluation with patient opinions confirmed the superiority of SenticNet with respect to other currently available lexical resources for opinion mining but the system is still limited to a relatively small number of concepts and simply provides polarity values associated with these.

In order to build a comprehensive resource for opinion mining and sentiment analysis, we use sentic computing to extract both cognitive and affective information from natural language text in a way that it is possible to map it into a fixed structure. In particular, we propose to bridge the cognitive and affective gap between word-level natural language data and their relative concept-level opinions and sentiments, by building semantics and sentics on top of them. Hence, SenticNet 2 provides, for each concept in the Open Mind corpus, not only its polarity but also its most probable domain of pertinence together with its top-ten semantically related concepts, its sentic values (i.e., its affective valence in terms of Pleasantness, Attention, Sensitivity and Aptitude) and its top-ten affectively related concepts. This information is encoded in RDF/XML using the descriptors defined by Human Emotion Ontology (HEO) (Grassi 2009).

### Extracting Semantics

The extraction of semantics associated with common sense knowledge concepts is performed through the ensemble application of spectral association and CF-IOF on the graph structure of ConceptNet. In particular, we apply CF-IOF on a set of 10,000 topic-tagged posts extracted from LiveJournal[1], a virtual community of more than 23 million who are allowed to label their posts not only with a topic tag but also with a mood label, by choosing from more than 130 predefined moods or by creating custom mood themes.

Thanks to CF-IOF weights, it is possible to filter out common concepts and detect domain-dependent concepts that individualize topics typically found in online opinions such as art, food, music, politics, family, entertainment, photography, travel, and technology. These concepts represent seed concepts for spectral association, which spreads their values across the ConceptNet graph. In particular, in order to accordingly limit the spreading activation of ConceptNet nodes, the rest of the concepts detected via CF-IOF are given as negative inputs to spectral association so that just domain-specific concepts are selected.

### Extracting Sentics

The extraction of sentics associated with common sense knowledge concepts is performed through the combined use of AffectiveSpace and the Hourglass model. In particular, we discard all but the first 100 singular values of the SVD and organize the resulting vector space using a k-medoids clustering approach (Park and Jun 2009), with respect to the Hourglass of Emotions (i.e., by using the model's labels as 'centroid concepts').

[1]http://livejournal.com

By calculating the relative distances (dot product) of each concept from the different centroids, it is possible to calculate its affective valence in terms of Pleasantness, Attention, Sensitivity and Aptitude, which is stored in the form of a four-dimensional vector, called sentic vector. The detection of the affectively related concepts, eventually, is given by simply selecting the first ten concepts that, in a normalized AffectiveSpace, have dot product with the given concept closest to 1.

### Encoding Semantics and Sentics

In order to represent SenticNet in a machine-accessible and machine-processable way, results are encoded in RDF triples using a XML syntax. In particular, concepts are identified using the ConceptNet Web API and statements are encoded in RDF/XML format on the base of HEO.

Statements have forms such as *concept – hasPlesantness – pleasantnessValue, concept – hasPolarity – polarityValue, concept – hasDomain – DomainName, concept – isSemanticallyRelated – concept* and *concept – hasPrimaryMood – PrimaryMoodName*. Given the concept 'birthday party', for example, SenticNet 2 provides 'events' as high-level domain of pertinence (which can be useful for tasks such as document auto-categorization) and a set of semantically related concepts, e.g., 'sweet', 'surprise friend' or 'clown' (which can be exploited as extra/contextual information to improve search results). The resource also provides a sentic vector specifying Pleasantness, Attention, Sensitivity and Aptitude associated with the concept (for tasks such as emotion recognition), a polarity value (for tasks such as polarity detection), a primary and secondary mood (for tasks such as HCI), and a set of affectively related concepts, e.g., 'celebration' or 'special occasion' (for tasks such as opinion classification).

Encoding semantics and sentics in RDF/XML using the descriptors defined by HEO allows cognitive and affective information to be stored in a Sesame triple-store, a purpose-built database for the storage and retrieval of RDF metadata. Sesame can be embedded in applications and used to conduct a wide range of inferences on the information stored, based on RDFS and OWL type relations between data. In addition, it can also be used in a standalone server mode, much like a traditional database with multiple applications connecting to it.

## Working with SenticNet 2

SenticNet 2 is freely available for download[2]. Thanks to its Semantic Web aware format, it is very easy to interface the resource with any real-world application that needs to extract semantics and sentics from natural language. This cognitive and affective information is supplied both at category-level (through domain and sentic labels) and dimensional-level (through polarity values and sentic vectors). Labels, in particular, are useful in case we deal with real-time adaptive applications (in which, for example, the style of an interface or the expression of an avatar has to quickly change according to user's input).

[2]http://sentic.net

Polarity values and sentic vectors, in turn, are useful for tasks such as information retrieval, opinion mining and sentiment analysis (in which it is needed to process batches of documents and, hence, perform calculations, such as addition, subtraction, and average, on both cognitive and affective information). Averaging results obtained at category-level is also possible by using a continuous 2D space whose dimensions are evaluation and activation, but the best strategy is usually to consider the opinionated document as composed of small bags of concepts (SBoCs) and feed these into SenticNet 2 to perform statistical analysis of the resulting sentic vectors.

In particular, we use a pre-processing module to interpret all the affective valence indicators usually contained in text such as special punctuation, complete upper-case words, onomatopoeic repetitions, exclamation words, negations, degree adverbs and emoticons, and eventually lemmatize text. A semantic parser then deconstructs text into concepts using a lexicon based on 'sentic n-grams', i.e., sequences of lexemes which represent multiple-word common sense and affective concepts extracted from the Open Mind corpus, WNA and other linguistic resources. The module also provides, for each retrieved concept, the relative frequency, valence, and status, that is, the concept's occurrence in the text, its positive or negative connotation, and the degree of intensity with which the concept is expressed.

We then use the resulting SBoC as input for SenticNet 2 and look up in it to obtain the relative sentic vectors, which we average in order to detect primary and secondary moods conveyed by the analyzed text and/or its polarity, given by the formula (Cambria et al. 2010c):

$$p = \sum_{i=1}^{N} \frac{Plsnt(c_i) + |Attnt(c_i)| - |Snst(c_i)| + Aptit(c_i)}{3N}$$

where $N$ is the size of the SBoC.

## Use Case Evaluation

As a use case evaluation of the system, we select the problem of crowd validation of the UK national health service (NHS), that is, the exploitation of the wisdom of the patient to efficiently validate the official UK hospital ratings provided by the health-care providers and NHS Choices.

To validate such data, we exploit patient stories extracted from PatientOpinion[3], a social enterprise providing an online feedback service for users of the UK NHS. The problem is that this social information is often stored in natural language text and hence intrinsically unstructured, which makes comparison with the structured information supplied by health-care providers very difficult. To bridge the gap between these data, which are different at structure-level yet similar at concept-level, we need to extract both the semantics and sentics associated with patient opinions. We exploit SenticNet 2 to marshal PatientOpinion's social information in a machine-accessible and machine-processable format and, hence, compare it with the official hospital ratings provided by NHS Choices and each NHS trust.

[3] http://patientopinion.org.uk

In particular, we use SenticNet 2 inferred ratings to validate the information declared by the relevant health-care providers, crawled separately from each NHS trust website, and the official NHS ranks, extracted using NHS Choices API. This kind of data usually consists of ratings that associate a polarity value to specific features of health-care providers such as communication, food, parking, service, staff, and timeliness. The polarity can be either a number in a fixed range or simply a flag (positive/negative).

Since each patient opinion can regard more than one topic and the polarity values associated with each topic are often independent from each other, in order to efficiently perform the mapping, we need to extract, from each opinion, a set of topics and then, from each topic detected, the polarity associated with it. In particular, after deconstructing each opinion into a set of SBoCs (one SBoC for each sentence), we analyze these through SenticNet 2 in order to tag each SBoC with one of the relevant topics (if any) and calculate a polarity value. We ran this process on a set of 2000 topic- and polarity-tagged stories extracted from PatientOpinion database and computed recall and precision rates as evaluation metrics. On average, each post contained around 140 words, from which about 12 affective valence indicators and 60 concepts were extracted.

As for the SBoC categorization, results showed that SenticNet 2 can detect topics in patient stories with satisfactory accuracy. In particular, the classification of 'food' and 'communication' sentences was performed with a precision of 75.1% and 69.3% and recall rates of 65.5% and 58.2%, respectively. The total F-measure rates, hence, were considerably good (70.8% for sentences about 'food' and 63.1% for sentences about 'communication'), particularly if compared to the corresponding F-measure rates calculated by using the extracted SBoCs as bags of words (BoWs) for the baseline methods (44.5% and 35.8% for keyword spotting, 53.2% and 39.1% for lexical affinity, 61.9% and 52.4% for statistical methods). As for the polarity detection, in turn, positiveness and negativeness of patient opinions were identified with particularly high precision (89.3% and 83.1%, respectively) and good recall rates (78.2% and 70.8%), for a total F-measure of 83.6% and 76.3%, respectively.

## Conclusion and Future Efforts

Today web-contents are perfectly suitable for human consumption but they remain hardly accessible to machines. Currently available information retrieval tools, in fact, still have to face a lot of limitations. To bridge the cognitive and affective gap between word-level natural language data and the concept-level opinions and sentiments conveyed by them, we developed SenticNet 2, a publicly available resource for opinion mining and sentiment analysis that associates semantics and sentics to every common sense concept from the Open Mind corpus. We showed how SenticNet 2 can easily be embedded in real-world applications, specifically in the field of social data mining, in order to effectively combine and compare structured and unstructured information. We are keeping on developing SenticNet 2 in a way that it can be enhanced with more common sense concepts from the always-growing Open Mind corpus but also from other semantic resources. We are also working on techniques and tools that will allow the resource to be easily merged with external domain-dependent knowledge bases, in order to improve the extraction of semantics and sentics from many different types of media. Finally, we like to see SenticNet 2 as a first step towards the development of sentic interfaces, i.e., next-generation intelligent applications capable of perceiving, interpreting, and expressing the cognitive and affective information associated with user interaction.

## References


Abbasi, A.; Chen, H.; and Salem, A. 2008. Sentiment analysis in multiple languages: Feature selection for opinion classification in web forums. *ACM Transactions on Information Systems* 26(3):1–34.

Bradley, M., and Lang, P. 1999. Affective norms for english words (ANEW): Stimuli, instruction manual and affective ratings. Technical report, The Center for Research in Psychophysiology, University of Florida.

Cambria, E., and Hussain, A. 2012. *Sentic Computing: Techniques, Tools, and Applications*. Berlin Heidelberg: Springer.

Cambria, E.; Hussain, A.; Havasi, C.; and Eckl, C. 2009a. AffectiveSpace: Blending common sense and affective knowledge to perform emotive reasoning. In *Proceedings of CAEPIA*, 32–41.

Cambria, E.; Hussain, A.; Havasi, C.; and Eckl, C. 2009b. Common sense computing: From the society of mind to digital intuition and beyond. volume 5707 of *Lecture Notes in Computer Science*. Berlin Heidelberg: Springer. 252–259.

Cambria, E.; Hussain, A.; Durrani, T.; Havasi, C.; Eckl, C.; and Munro, J. 2010a. Sentic computing for patient centered application. In *Proceedings of IEEE ICSP*, 1279–1282.

Cambria, E.; Hussain, A.; Havasi, C.; and Eckl, C. 2010b. SenticSpace: Visualizing Opinions and Sentiments in a Multi-Dimensional Vector Space. volume 6279 of *Lecture Notes in Artificial Intelligence*. Berlin Heidelberg: Springer. 385–393.

Cambria, E.; Speer, R.; Havasi, C.; and Hussain, A. 2010c. SenticNet: A publicly available semantic resource for opinion mining. In *Proceedings of AAAI CSK*, 14–18.

Dagan, I.; Glickman, O.; and Magnini, B. 2006. The pascal recognising textual entailment challenge. In Quiñonero-Candela, J.; Dagan, I.; Magnini, B.; and d'Alché Buc, F., eds., *Machine Learning Challenges*, volume 3944 of *Lecture Notes in Computer Science*. Berlin Heidelberg: Springer. 177–190.

Eckart, C., and Young, G. 1936. The approximation of one matrix by another of lower rank. *Psychometrika* 1(3):211–218.

Elliott, C. D. 1992. *The Affective Reasoner: A Process Model of Emotions in a Multi-Agent System*. Ph.D. Dissertation, Northwestern University, Evanston.

Esuli, A., and Sebastiani, F. 2006. SentiWordnet: A publicly available lexical resource for opinion mining. In *Proceedings of LREC*.

Etzioni, O.; Banko, M.; and Cafarella, M. 2006. Machine reading. In *Proceedings of AAAI06*.

Grassi, M. 2009. Developing HEO human emotions ontology. volume 5707 of *Lecture Notes in Computer Science*. Berlin Heidelberg: Springer. 244–251.

Havasi, C.; Speer, R.; and Alonso, J. 2007. ConceptNet 3: a flexible, multilingual semantic network for common sense knowledge. In *Proceedings of RANLP*.

Havasi, C.; Speer, R.; and Holmgren, J. 2010. Automated color selection using semantic knowledge. In *Proceedings of AAAI CSK*.

Hu, M., and Liu, B. 2004. Mining opinion features in customer reviews. In *Proceedings of AAAI*.

Minsky, M. 2006. *The Emotion Machine: Commonsense Thinking, Artificial Intelligence, and the Future of the Human Mind*. Simon & Schuster.

Pang, B., and Lee, L. 2005. Seeing stars: Exploiting class relationships for sentiment categorization with respect to rating scales. In *Proceedings of ACL*, 115–124.

Park, H., and Jun, C. 2009. A simple and fast algorithm for k-medoids clustering. *Expert Systems with Applications* 36(2):3336–3341.

Picard, R. 1997. *Affective Computing*. The MIT Press.

Plutchik, R. 2001. The nature of emotions. *American Scientist* 89(4):344–350.

Rao, D., and Ravichandran, D. 2009. Semi-supervised polarity lexicon induction. In *Proceedings of EACL*, 675–682.

Somasundaran, S.; Wiebe, J.; and Ruppenhofer, J. 2008. Discourse level opinion interpretation. In *Proceedings of COLING*.

Stevenson, R.; Mikels, J.; and James, T. 2007. Characterization of the affective norms for english words by discrete emotional categories. *Behavior Research Methods* 39:1020–1024.

Strapparava, C., and Valitutti, A. 2004. WordNet-Affect: an affective extension of WordNet. In *Proceedings of LREC*.

Turney, P., and Littman, M. 2003. Measuring praise and criticism: Inference of semantic orientation from association. *ACM Transactions on Information Systems* 21(4):315–346.

Velikovich, L.; Goldensohn, S.; Hannan, K.; and McDonald, R. 2010. The viability of web-derived polarity lexicons. In *Proceedings of NAACL*, 777–785.

Wall, M.; Rechtsteiner, A.; and Rocha, L. 2003. Singular value decomposition and principal component analysis. In Berrar, D.; Dubitzky, W.; and Granzow, M., eds., *A Practical Approach to Microarray Data Analysis*. Springer US. 91–109.

Wiebe, J.; Wilson, T.; and Cardie, C. 2005. Annotating expressions of opinions and emotions in language. *Language Resources and Evaluation* 39(2):165–210.

Wilson, T.; Wiebe, J.; and Hoffmann, P. 2005. Recognizing contextual polarity in phrase-level sentiment analysis. In *Proceedings of HLT/EMNLP*.