Towards Crowd Validation of the UK National Health Service

Erik Cambria
Dept. of Computing Science and Mathematics
University of Stirling, UK
eca@cs.stir.ac.uk

Amir Hussain
Dept. of Computing Science and Mathematics
University of Stirling, UK
ahu@cs.stir.ac.uk

Catherine Havasi
MIT Media Lab
Massachussets Institute of Technology, USA
havasi@media.mit.edu

Chris Eckl
Sitekit Labs, UK
chris.eckl@sitekit.net

James Munro
Patient Opinion, UK
james.munro@patientopinion.org.uk

ABSTRACT

Online patient opinions are a very important instrument for the effective evaluation of local hospitals, hospices, and mental health services. The distillation of knowledge from this unstructured information, however, remains a difficult and complex task. In this paper, we aim to effectively mine and analyze this social information to make a comprehensive and dynamic evaluation of the UK National Health Service. To this end we use SenticNet, a new opinion mining and sentiment analysis resource which exploits AI and Semantic Web techniques to better recognize, interpret, and process opinions and sentiments in natural language text.

Keywords

AI, Knowledge Base Management, NLP, Opinion Mining and Sentiment Analysis, E-Health

1. INTRODUCTION

The transformation of the Web from its humble read-only origins to the current interactive/read-write medium has made its users ever more keen to interact, share and collaborate through blogs, wikis, fora, chats, social networks, etc., giving birth to a sort of collective intelligence that emerges from the integration, competition and cooperation of many individuals or agents.

In recent years, this novel approach to information sharing has increasingly found its way, under various guises, into many fields such as encyclopedic knowledge, commerce, education, tourism and health. Online patient opinions, in particular, are a very important instrument for the effective evaluation of local hospitals, hospices and mental health services but the distillation of knowledge from this unstructured information remains a difficult and complex task.

Patient Opinion[1] is a social enterprise pioneering an online feedback service for users of the UK National Health Service (NHS) to enable people to share their recent experience of local health services on-line, and, ultimately, help citizens improve their NHS.

[1]`http://patientopinion.org.uk`

Web Science Conf. 2010, April 26–27, 2010, Raleigh, NC, USA.
.

Patients, relatives and even health-care professionals have been telling their stories in support forums and on personal sites since the Web began but the information they have been providing is narrative, loosely structured and – like most of the Web – hard to identify and retrieve. Patient Opinion gives patients and carers a new and powerful voice, allowing stories and experiences to be shared, feedback to be offered and new kinds of social reputation to be created.

In this work, to effectively distill, manage and process this social information in an automated way, we propose to use SenticNet [1], a novel publicly available semantic resource for opinion mining.

2. SENTICNET

Existing approaches to automatic identification and extraction of opinions and sentiments from text can be grouped into three main categories: keyword spotting [2][3], in which text is classified into categories based on the presence of fairly unambiguous affect words, lexical affinity [4][5], which assigns arbitrary words a probabilistic affinity for a particular opinion or emotion, and statistical methods [6][7], which consist in calculating the valence of keywords, punctuation and word co-occurrence frequencies on the base of a large training corpus.

The problem with these approaches is that they mainly rely on parts of text in which opinions are explicitly expressed such as positive terms (e.g., good, nice, excellent, fortunate, correct, superior, best) and negative terms (e.g., bad, nasty, poor, unfortunate, wrong, inferior, worst). In general, in fact, opinions are expressed implicitly through context and domain dependent concepts, which makes keyword-based approaches extremely ineffective.

SenticNet overcomes this problem by exploiting a novel emotion categorization model (section 2.1), an affective semantic network (section 2.2) and a language visualization and analysis system (section 2.3).

2.1 The Hourglass of Emotions

This model is a variant of Plutchik's emotion categorization [8] and constitutes an attempt to emulate Marvin Minsky's conception of emotions. Minsky sees the mind as made of thousands of different resources and believes that our emotional states result from turning some set of these resources on and turning another set of them off [9].

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 Hourglass of Emotions (Fig. 1) is specifically designed to recognize, understand and express emotions in the contexts of user profiling [10] and human-computer interaction (HCI).

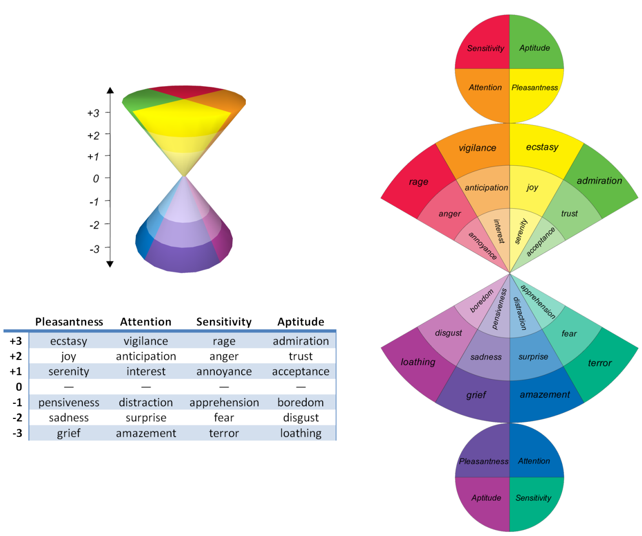
In the model, in fact, affective states are not classified, as often happens in the field of emotion analysis, into basic emotional categories, but rather into four concomitant but independent dimensions – Pleasantness, Attention, Sensitivity and Aptitude – in order to understand how much respectively:

1. the user is happy with the service provided
2. the user is interested in the information supplied
3. the user is comfortable with the interface
4. the user is disposed to use the application

Each affective dimension is characterized by six levels of activation, called 'sentic levels', which determine the intensity of the expressed/perceived emotion. The concomitance of the different affective dimensions makes possible the generation of compound emotions such as 'love', which is given by the sum of positive values of Pleasantness and Aptitude, or 'aggressiveness', given by the concomitance of Attention and Sensitivity.

2.2 AffectNet

When people communicate with each other, they rely on similar background knowledge, e.g., the way objects relate to each other in the world, people's goals in their daily lives, the emotional content of events or situations. This 'taken for granted' information is what we call common-sense [11] – obvious things people normally know and usually leave unstated. The Open Mind Common Sense project has been collecting this kind of knowledge from volunteers on the Internet since 2000 to provide intuition to AI systems.



	Pleasantness	Attention	Sensitivity	Aptitude
+3	ecstasy	vigilance	rage	admiration
+2	joy	anticipation	anger	trust
+1	serenity	interest	annoyance	acceptance
0	—	—	—	—
-1	pensiveness	distraction	apprehension	boredom
-2	sadness	surprise	fear	disgust
-3	grief	amazement	terror	loathing

Figure 1: The Hourglass of Emotions

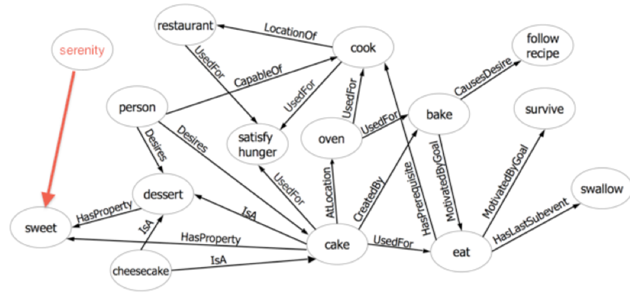


Figure 2: A sketch of AffectNet

ConceptNet [12] represents the information in the Open Mind corpus as a directed graph in which the nodes are concepts and the labeled edges are assertions of common-sense that interconnect them. WordNet-Affect [13] is a linguistic resource for the lexical representation of affective knowledge, developed starting from WordNet [14]. The ontology is built by assigning to a number of WordNet synsets one or more affective labels (a-labels) and then by extending the core with the relations defined in WordNet.

In particular, the affective concepts representing emotional states are identified by synsets marked with the a-label 'emotion', but there are also other a-labels for concepts representing moods, situations eliciting emotions or emotional responses. The blend of ConceptNet and WordNet-Affect is what we call AffectNet – an affective semantic network in which common-sense concepts are linked to a hierarchy of affective domain labels (Fig. 2).

Blending [15] is a technique that performs inference over multiple sources of data simultaneously, taking advantage of the overlap between them. It basically combines two sparse matrices linearly into a single matrix in which the information between the two initial sources is shared. The first step to create a blend is to transform the input data so that it can all be represented in the same matrix.

To this end, we align the lemma forms of ConceptNet concepts with the lemma forms of the words in WordNet-Affect and map the most common relations in the affective knowledge base into ConceptNet's set of relations, e.g., Hypernym into IsA and Holonym into PartOf. This alignment operation yields a new semantic network in which common-sense and affective knowledge are in fact combined, not just concomitant, i.e., a network in which everyday life concepts like 'have breakfast', 'meet people' or 'watch tv' are linked to affective domain labels like 'joy', 'anger' or 'surprise'.

2.3 SenticSpace

SenticSpace [16] is the vector space representation of AffectNet, obtained by applying principal component analysis (PCA) [17] on its matrix representation. It is a language visualization and analysis system which transforms natural language from a linguistic form into a real-time navigable space. After performing truncated singular value decomposition (TSVD) on AffectNet, we obtain a new matrix $\tilde{A} = U_k * \Sigma_k * V_k^T$, which forms 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 [18], it represents the best approximation of A in the mean-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:

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

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 same as those of A. We use a trial and error approach to set $k = 50$ and hence obtain SenticSpace (Fig. 3), a 50-dimensional space in which different vectors represent different ways of making binary distinctions among concepts and emotions.

By exploiting the information sharing property of TSVD, concepts with the same affective valence are likely to have similar features, i.e., concepts concerning the same emotion tend to fall near each other in the vector space. In SenticSpace, in fact, concepts are represented by vectors of 50 coordinates which can be seen as describing concepts and emotions in terms of 'eigenmoods' that form the axes, i.e., the bases e_0,...,e_{49}, 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. Consequently concepts with negative e_0 components have negative affective valence.

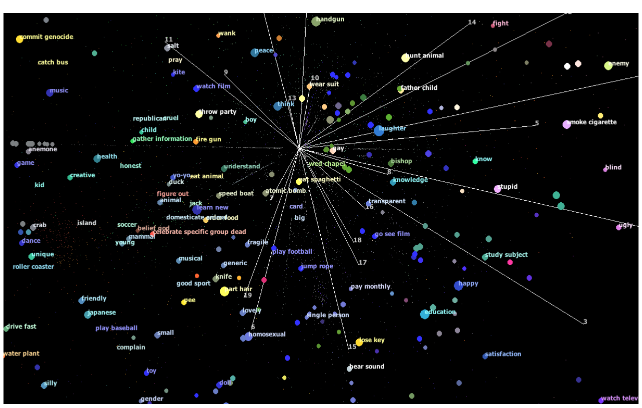


Figure 3: A sketch of SenticSpace

3. PATIENT OPINION ANALYSIS

The lack of power of patients, in comparison to the health-care system they navigate and the professionals with whom they negotiate, is nothing new. The word 'patient', in fact, derives from the Latin 'to suffer' and embodies meanings like 'being passive' and 'waiting without complaint'. By definition, patients are expected to be passive, vulnerable and gratefully receptive to the prodding and poking of health-care tests and treatments.

The families and carers of patients and service users are not much better placed. They may support their loved ones on their journey through health-care, advocate for them, ask questions, argue for services. But fundamentally, they are similarly powerless in relation to the system of care and the professionals who inhabit it.

In some cultures, patients are redefined as consumers of health-care in attempt to restore to them some of the apparent power of the customer in a marketplace. The movement from being passive to being active is a significant one, yet in practice the power of a consumer seldom goes beyond the choice either to consume what is being offered or to go elsewhere.

In this section we analyze the importance for patients to evolve from passive to active entities in the health-care system for an effective improvement of the NHS (section 3.1), describe the process for mining, managing and analyzing online patient opinions (section 3.2) and provide an evaluation of it (section 3.3).

3.1 Patient 2.0

The advent of Web 2.0 has already unleashed a wave of 'creative destruction' upon traditional business models, throwing into disarray sectors such as music, journalism, advertising and retailing. And it threatens to have a similarly transformative effect on health-care. Web 2.0, in fact, drops the cost of voice, of finding others 'like me', of obtaining and republishing information, and of forming groups, to zero. As a result, it becomes easy and rewarding for patients, service users, carers and families to research conditions and treatments, give feedback, ask questions, demand answers and press for openness [19].

In March 2010 two UK hospitals hit the headlines when it emerged that hundreds of patients may have died unnecessarily, despite one hospital being rated as good and the other fair. Panorama investigated how many more hospitals may be failing the public while being allowed to assess themselves as good. On its own, in fact, self assessment is an incomplete measure of hospital performance. NHS trusts need to be validated by somebody outside the NHS but, at the same time, inside the health-care system. Neither the doctors, nor the nurses, nor the therapists but the real end-users of health-care – the patients.

Our vision of next-generation patient is far from being peripheral to health-care. Patient 2.0 is central to understanding the effectiveness and efficiency of services and how they can be improved. He/she is not just a consumer of the health-care system but a quality control manager – his/her opinions are not just reviews of a product/service but more like small donations of experience, digital gifts which, once given, can be shared, copied, moved around the world and directed to just the right people who can use them to improve health-care locally, regionally or nationally.

3.2 Opinion Analysis Process

Patient Opinion sees the patient feedback gathered from on-line users as as a public resource which can and should be re-used for wider public benefit. To this end it provides read-only access to published postings, with their linked tags, services and responses, via an API, which we use to mine patient opinions and all the relevant information related to them.

Given a textual resource about a set of topics T containing a set of opinions O with different polarity $p \in \{-1, 1\}$, we define opinion mining as the process which aims to extract, for each $t \in T$, the subset of opinions $o \subseteq O$ concerning t and determine p. We therefore attack the problem of opinion mining firstly as a topic spotting problem and secondly as a polarity classification one.

In particular we process each patient opinion through a NLP module, which performs a first skim of the document, a Semantic Parser, whose aim is to extract concepts from the stemmed text, AnalogySpace [20], a process to detect topics in text, and eventually SenticSpace, for the opinion polarity detection.

The NLP module interprets 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.

The Semantic Parser then deconstructs text into concepts and provides, for each of them, the relative frequency, valence and status, i.e., the concept's occurrence in the text, its positive or negative connotation, and the degree of intensity with which the concept is expressed.

AnalogySpace is the analogical closure of ConceptNet, i.e., the vector space resulting from applying TSVD on the common-sense knowledge base. We use a k-means clustering approach to find concepts semantically related to topics such as clinical service, communication, food, parking, staff or timeliness, and then scan each opinion for these concepts in order to classify it.

Likewise we organize SenticSpace according to the sentic levels of the Hourglass model and then project each retrieved concept into the vector space. This way concepts' affective valence is calculated, according to their relative positions in SenticSpace, and opportunely converted into $float \in \{-3, 3\}$ to express it in terms of Pleasantness, Attention, Sensitivity and Aptitude.

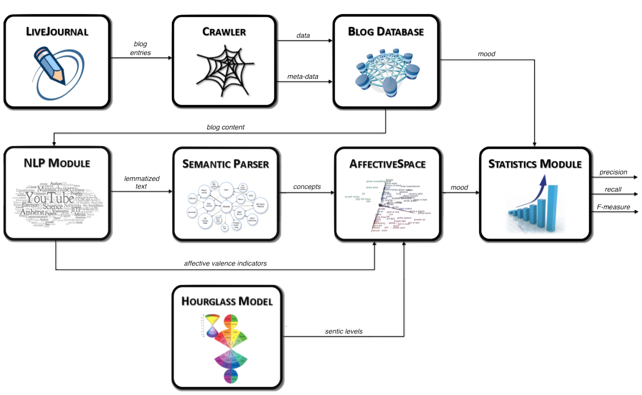


Figure 4: SenticNet Evaluation Process

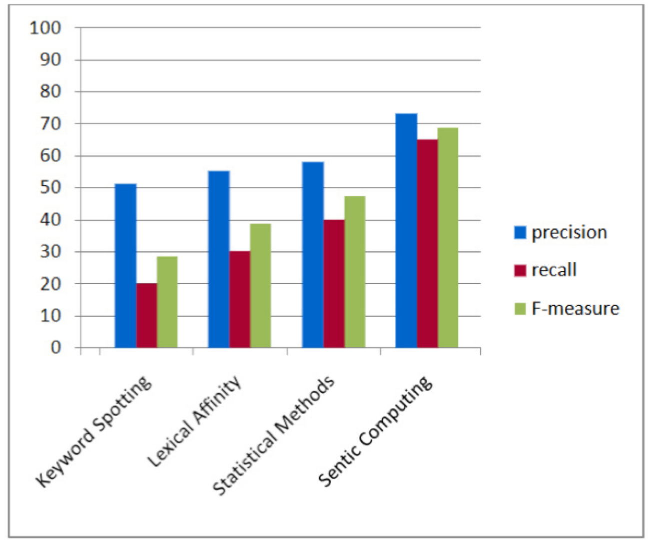


Figure 5: Comparison of different opinion mining and sentiment analysis approaches

For each opinion we finally average out all the obtained sentic values and calculate the overall polarity, which we define as:

$$p = \sum_{i=1}^{N} \frac{Plsnt(o_i) + |Attnt(o_i)| - |Snsit(o_i)| + Aptit(o_i)}{9N}$$

3.3 Evaluation

An evaluation of the polarity classification process was performed by considering a corpus of blog-posts from LiveJournal (LJ), a virtual community where web-users can keep a blog, journal or diary.

One of the interesting features of this web-site is that LJ bloggers, who number over 23 millions, are allowed to label their posts with a mood tag, by choosing from more than 130 predefined moods or even by creating custom mood themes. Since the indication of the affective status is optional, the mood-tagged posts are likely to reflect the true mood of the authors, and hence form a good test-set for our polarity classification process.

However, since LJ mood themes do not perfectly match the sentic levels, we had to consider a relatively small set of moods, i.e., 'ecstatic', 'cheerful', 'pensive', 'surprised', 'enraged', 'sad', 'angry', 'annoyed', 'scared' and 'bored'. Moreover we could not consider non-affective web-posts since untagged blog entries do not necessarily lack emotions.

All LJ accounts have Atom 1.0, RSS 2.0, and other data feeds, which show recent public entries, friend relationships and interests. Unfortunately there is no possibility to get mood-tagged blog-posts via data feeds so we had to design our own crawler.

After selecting the most popular LJ blogs, the crawler retrieved the mood-tagged posts and stored their relevant data and meta-data in an ad-hoc database. For each post in the database we then extracted the content and gave it as input to the polarity classification process. The output of the process was finally compared with the mood retrieved from the database to calculate statistical classifications such as precision and recall (Fig. 4).

In particular we crawled and processed 5,000 web-posts: on average the post length was 242 words, the NLP Module detected 7 affective valence indicators per post, the Semantic Parser extracted 53 concepts and SenticSpace produced 48 sentic vectors (i.e., the concepts' relative distances in the vector space) per each post.

According to the values of the sentic vectors, we categorized each post and compared the results with the corresponding LJ tags: the system showed a very high precision (73%) and significantly good recall and F-measure rates (65% and 68% respectively), outperforming the baseline methods (Fig. 5).

4. CONCLUSION AND FUTURE WORK

In this paper we proposed a new opinion mining and sentiment analysis paradigm to automatically analyze patient opinions about the UK NHS. By exploiting a novel emotion categorization model, an affective semantic network and a language visualization and analysis system, we extracted from Patient Opinion on-line contributions useful information about the perceived quality of many UK hospitals, hospices and mental health services.

The next step of this research work is to merge and compare this social information with the official hospital ratings provided by NHS Choices[2] and each NHS trust, to make a comprehensive and dynamic evaluation of the UK NHS (Fig. 6). We call this process 'Crowd Validation', because of the feedback coming the masses, and we believe it will be the next frontier of health-care since patient opinions are crucial to understand the effectiveness and efficiency of health services and how they can be improved.

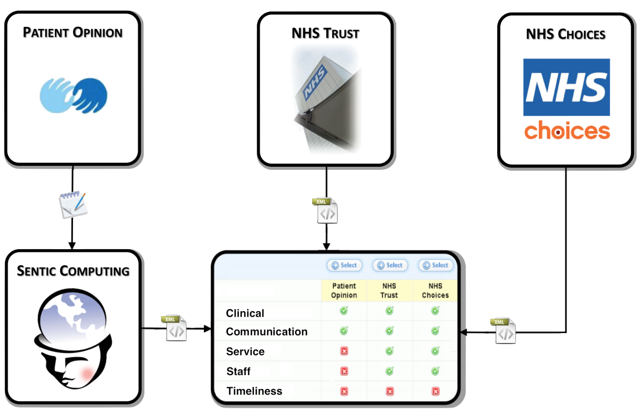


Figure 6: Crowd Validation of the UK NHS

5. REFERENCES

[1] E. Cambria, R. Speer, C. Havasi, and A. Hussain. *SenticNet: A publicly available semantic resource for opinion mining.* In: AAAI CSK, pp. 14-18, Arlington (2010)

[2] C. Elliott. *The Affective Reasoner: A Process Model of Emotions in a Multi-Agent System.* The Institute for the Learning Sciences, Technical Report No. 32 (1992)

[3] J. Wiebe, T. Wilson, and C. Claire. *Annotating Expressions of Opinions and Emotions in Language.* Language Resources and Evaluation 2(3), 165–210 (2005)

[4] S. Somasundaran, J. Wiebe, and J. Ruppenhofer. *Discourse Level Opinion Interpretation.* In: COLING, Manchester (2008)

[5] T. Wilson, J. Wiebe, and P. Hoffmann. *Recognizing Contextual Polarity in Phrase-Level Sentiment Analysis.* In: HLT-EMNLP, Vancouver (2005)

[6] A. Abbasi, H. Chen, and A. Salem. *Sentiment Analysis in Multiple Languages: Feature Selection for Opinion Classification in Web Forums.* ACM Transactions on Information Systems 26(3), 1–34 (2008)

[7] M. Gamon. *Sentiment Classification on Customer Feedback Data: Noisy Data, Large Feature Vectors, and the Role of Linguistic Analysis.* In: COLING, Geneva (2004)

[8] R. Plutchik. *The Nature of Emotions.* American Scientist 89(4), 344–350 (2001)

[9] M. Minsky. *The Emotion Machine.* Simon and Schuster, New York (2006)

[10] E. Cambria, P. Chandra, A. Sharma, and A. Hussain. Do not feel the trolls. In: ISWC, Shanghai (2010)

[11] E. Cambria, A. Hussain, C. Havasi, and C. Eckl. *Common Sense Computing: From the Society of Mind to Digital Intuition and Beyond.* In: LNCS, vol. 5707, pp. 252–259. Springer (2009)

[12] C. Havasi, R. Speer, and J. Alonso. *ConceptNet 3: a Flexible, Multilingual Semantic Network for Common Sense Knowledge.* In: RANLP, Borovets (2007)

[13] C. Strapparava and A. Valitutti. *WordNet-Affect: an Affective Extension of WordNet.* In: LREC, Lisbon (2004)

[14] C. Fellbaum. *WordNet: An Electronic Lexical Database.* MIT Press, Cambridge (1998)

[15] C. Havasi, R. Speer, J. Pustejovsky, and H. Lieberman. *Digital Intuition: Applying Common Sense Using Dimensionality Reduction.* IEEE Intelligent Systems 24(4), 24–35 (2009)

[16] E. Cambria, A. Hussain, C. Havasi, and C. Eckl. *SenticSpace: Visualizing opinions and sentiments in a multi-dimensional vector space.* In: LNAI, vol. 6279, pp. 385–393, Springer (2010)

[17] M. Wall, A. Rechtsteiner, and L. Rocha. *Singular Value Decomposition and Principal Component Analysis.* In: Berrar, D. et al. (eds.) A Practical Approach to Microarray Data Analysis, pp. 91–109. Kluwer, Norwell (2003)

[18] C. Eckart and G. Young. *The Approximation of One Matrix by Another of Lower Rank.* Psychometrika 1(3), 211–218 (1936)

[19] C. Shirky. *Here Comes Everybody: the Power of Organizing Without Organizations.* Allen Lane, London (2008)

[20] R. Speer, C. Havasi, and H. Lieberman. *AnalogySpace: Reducing the Dimensionality of Common Sense Knowledge.* In: AAAI, Chicago (2008)

[2] `http://www.nhs.uk`