



HAL
open science

Beyond sentiments and opinions: exploring social media with appraisal categories

Valentina Dragos, Delphine Battistelli, Emmanuelle Kelodjoue

► To cite this version:

Valentina Dragos, Delphine Battistelli, Emmanuelle Kelodjoue. Beyond sentiments and opinions: exploring social media with appraisal categories. FUSION 2018, Jul 2018, CAMBRIDGE, United Kingdom. pp.1851-1858, 10.23919/ICIF.2018.8455751 . hal-01961227

HAL Id: hal-01961227

<https://hal.science/hal-01961227v1>

Submitted on 19 Dec 2018

HAL is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.

L'archive ouverte pluridisciplinaire **HAL**, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d'enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.

Beyond sentiments and opinions: exploring social media with appraisal categories

Valentina Dragos

ONERA - The French Aerospace Lab
Chemin de la Hunière
Palaiseau, France
valentina.dragos@onera.fr

Delphine Battistelli

MoDyCo-
Université Paris- Nanterre – CNRS
Nanterre, France
delphine.battistelli@parisnanterre.fr

Emmanuelle Kelodjue

ONERA - The French Aerospace Lab
Chemin de la Hunière
Palaiseau, France
e.kelodjue@onera.fr

Abstract - The digital era arrives with a whole set of disruptive technologies that creates both risk and opportunity for open sources analysis. Although the sheer quantity of online conversations makes social media a huge source of information, their analysis is still a challenging task and many of traditional methods and research methodologies for data mining are not fit for purpose. Social data mining revolves around subjective content analysis, which deals with the computational processing of texts conveying people's evaluations, beliefs, attitudes and emotions. Opinion mining and sentiment analysis are the main paradigm of social media exploration and both concepts are often interchangeable. This paper investigates the use of appraisal categories to explore data gleaned for social media, going beyond the limitations of traditional sentiment and opinion-oriented approaches. Categories of appraisal are grounded on cognitive foundations of the appraisal theory, according to which people's emotional response are based on their own evaluative judgments or appraisals of situations, events or objects. A formal model is developed to describe and explain the way language is used in the cyberspace to evaluate, express mood and subjective states, construct personal standpoints and manage interpersonal interactions and relationships. A general processing framework is implemented to illustrate how the model is used to analyze a collection of tweets related to extremist attitudes.

Keywords—social media, open source, sentiments, opinions, appraisal theory, big data

I. INTRODUCTION

With millions of users connected through Twitter, Facebook, Google+ and other online platforms, Internet dramatically changed the way people interact, exchange ideas and influence each other in the cyberspace and real life. Communication and social interactions, used to describe how people act and react to those around them are now transposed in virtual environments, giving rise to virtual communities without physical borders and strong links. Ideas, views and impressions are spread through blog posts, reviews, forums,

augmented with specific mechanisms for online propagation such as like and retweet.

From content analysis perspective, social media users are rather vocal, often turning to online platforms to express their tastes. Exploring online media comes with unprecedented opportunities for a broad category of applications, among which intelligence analysis and homeland security can explore the huge amount of data for insights allowing them to understand and to some degree predict new phenomena. Having abounding data becomes an advantage only through careful analysis performed with robust methods able to turn the huge volume of text into relevant patterns.

Analysis of social data comes with challenges for traditional data mining approaches that are often too slow and expensive, rely on sample sizes, and come with biases leading to errors. Social media data are vast, noisy, unstructured, inherently dynamic and heterogeneous in nature. Moreover, they convey author's personal points of view and processing solutions fall under the umbrella of sentiment analysis and opinion mining [12], a natural language processing task dealing with automatic processing of people's evaluations, attitudes, and emotions as expressed in written language. While sometimes those concepts are interchangeable, it is generally accepted that opinion mining considers not only the sentiment conveyed by written stances, but also the topics driving that sentiment.

This paper tackles the use of appraisal categories to explore social data. The approach is grounded on cognitive foundations of the appraisal theory developed by White and Martin [1] and providing finer categories to analyse subjective content and making the distinction between affect, appreciation or judgement. The appraisal framework investigates the meanings by which text convey positive and negative attitudes, but also considers the way such utterances are strengthened or weakened by author's thanks to linguistic clues of intensity. Adopting appraisal categories allows create richer descriptions of social data, going beyond limitations of opinion and

sentiment, and presents author's feelings, tastes and opinions with greater or lesser intensity.

The remainder of the paper is organized as follows: the next section discussed related approaches for social data analysis. Section III introduces main notions of the appraisal theory while section IV presents the construction of an ontology offering a formal representation of those notions. An experimentation carried out to illustrate the use of the ontology to analyse a collection of tweets is discussed in section V. Conclusion and perspectives for future work end this paper.

II. RELATED WORK

The sudden eruption of social media and its associated volumes of data triggered a dynamic activity in the area of opinion mining and sentiment analysis and methods developed can be roughly divided into lexicon-based methods [18] and machine-learning methods [25].

Early machine learning approaches constructed a basic binary classifier which used n-grams and part-of-speech features, to assign positive or negative labels to text. Among them, Pak and Paroubek [27] classified tweets as objective, positive and negative by using a sentiment classifier based on the multinomial Naive Bayes, and using a combination of syntactic and linguistic features such as n-gram and POS-tags.

Barbosa and colleagues [28] implemented a two phase classifier, that detected first subjective and objective tweets, and then classified subjective tweets were classified as positive or negative. The set of feature also included platform-specific input, in the form of retweets and hashtags.

Liang et al. [4] used a basic unigram Naive Bayes model to classify tweets as positive, negative, and non-opinions. The overall classification approach was improved by using the Mutual Information and Chi square test to eliminate useless, irrelevant features. Another solution based on Bayesian classifiers augmented with linguistic inputs is presented by Gamallo and colleagues in [6]. The authors designed two variants of Naive Bayes classifiers were built Baseline, trained to classify tweets as positive, negative and neutral, and Binary classifying tweets as positive and negative while neutral tweets were ignored. The set of features consisted of Lemmas of nouns, verbs, adjectives and adverbs and results are also tuned by using a polarity lexicon.

Xia et al. [11] analysed the association of various feature sets and classification techniques. The authors used two types of feature sets (Part-of-speech information and lexical relations) and three basic classifiers - Naive Bayes, Maximum Entropy and Support Vector Machines. Then, they achieved a better accuracy for sentiment classification by using different combination strategies such as weighted and Meta-classifier combination.

Although approaches above show promising results, opinion mining techniques making use of machine learning become problematic for social data exploration, which involves several different domains, multi languages and distinct text types, because models have to be trained for each one, and large sets of training data are required to achieve good results.

Generally, most classifiers built using supervised methods [31] perform well on polarity detection tasks, but their accuracy decreases drastically when used in new domains [30].

Li and colleagues adopted active learning to tackle the problem of cross-domain opinion mining [14]. Their approach implements two individual classifiers, trained with samples from two distinct domains called source and target respectively and then classify unlabeled data in the target domain with the label propagation algorithm. The overall solution takes advantage of a selection strategy designed to dynamically select samples of labelled data to be used as training sets for both classifiers.

Adaptive SVM models are described in [16] by Liu et al., who build a topic-adaptive classifier. The adaptation is carried out by alternatively performing three steps: optimization of classification, selection of unlabeled data sets and adaptive feature expansion with respect to newly added training sets. The authors show that, compared to supervised sentiment classifiers and semi-supervised approaches, their solution increases accuracy on 6 topics when applied to process a publicly available tweet corpus.

Mejova and Srinivasan [7] addressed domain adaptation by using three media types of texts, namely, blogs, reviews, and Twitter, and showed that models used to classify texts in one domain (also called source domain), provide good results when used on a distinct domain, called target domain. The authors used 37 topics to illustrate the transposition of classification models created for a domain to another.

From a different perspective, several approaches adapted lexicon-based methods, which rely on a lexicon, or a collection of known and precompiled terms, augmented with sentiment-specific attributes, such as polarity and orientation. Among resources created by different teams, SenticNet 2.0 [26] offers a collection of around 100,000 natural language concepts, described in terms of four affective dimensions (Pleasantness, Attention, Sensitivity, and Aptitude) and also having a polarity assignment, as a floating number between -1 and +1, where -1 is negative polarity and +1 is positive polarity.

A Sentiment Treebank is used in [17] to provide fine grained sentiment labels for around 215 000 phrases and to allow sentiment compositionality. The Treebank is used to train a recursive neural tensor network, and the authors show that the model outperforms previous opinion detection methods on several metrics, while being able to accurately capture the effect of contrastive conjunctions and negations.

Domain adaptation is still a challenge for lexicon-based approaches, and Bollgala et al. describe in [9] a solution using a distributional thesaurus to expand feature vectors during training and testing phases of a binary classifier. The lexicon provides a set of labelled data for the source domain and unlabeled data for both source and target domains, and sensitivity attributes are added for each word by measuring their distributional similarity.

Although based on lexical features, several approaches analyse not only the word, but also wordarounds, in an effort to harness sentiment expression forms unique to data to be processed. Among them, SentiStrength [13] is an algorithm developed by to detect the intensity of sentiments by using a list

III. OVERVIEW OF THE APPRAISAL THEORY

of 2,489 positive and negative sentiment stems and strengths, and combination rules to estimate the overall sentiment and strength at phrase level by combining their values. They also take into account linguistic boosters and downtoners, and the approach is suitable to analyse short text gleaned on social media, which often lacks standard grammar and spelling. In the same line, the solution described in [3] learns word embeddings to classify Twitter corpora. The approach models the syntactic context of word and also encodes sentiment information in the continuous representation of words. The overall solution shows good accuracy but requires a large scale training corpora.

However, such relatively robust techniques, grounded on linguistic bases, come with a main drawback, as they are inflexible regarding the ambiguity of terms conveying sentiments, orientations and polarity. The context in which a term occurs slightly modifies its meaning, and this impacts in particular adjectives [32]. Several studies have shown that adding contextual information [33], [34] and can have a high impact on the polarity of ambiguous [35] and improve the estimation of their meaning and polarity.

In addition to those main classes, hybrid approaches are also developed, with sentiment lexicons playing a key role in the implementation of learning strategies [20]. Among them, Agarwal and colleagues [29] carried experimentations with three different approaches - unigrams, a feature based model and a tree kernel based model - to classify tweets as positive, neutral or negative. The experiment showed that the set of features combining prior polarity of words and parts-of-speech (pos) tags are the most relevant and plays a major role in the classification task.

A detailed review of sentiment detection techniques developed to analyse Twitter data is presented in [15], including many sub-components and examples of tools.

Approaches above investigate sentiment and opinion analysis, but in several domains such as homeland security and defence, analysts are still struggling to understand emergent online phenomena - online hate, propagation of extreme attitudes - which are grounded on citizen's feelings towards today's stories and events. For those domains the practical need is to investigating subjective content above traditional notions of opinion and sentiment. This paper describes an approach designed to explore social with appraisal categories data for security and defence purposes. This in an emergent topic and there are few approaches currently using the appraisal theory for social data analysis, and they addressed the construction of a semantic resource [23] to detect sentiments based on the appraisal theory and the use of appraisal categories to analyse microblogs [22] and news [24].

The approach we take for social data exploration focuses on building a semantic resource having a number of subcomponents which all are able to highlight various aspects of subjective content. The main body of content analysis involves a set of natural language processing steps which create annotations on segments of text gleaned online.

The Appraisal Theory is a cognitive frame claiming that people's emotions are elicited by their personal and continuous interpretations, evaluations or appraisals of objects, events and situations. The most important aspect of the appraisal theory is that it provides a way to express how humans interpret some particular event -positive, negative- as well as their position, support and engagement with respect to their own interpretation and report – confidence, support, agreement, disagreement.

From a linguistic standpoint, the appraisal theory describes how authors use linguistic expressions to communicate their emotional states and engagement. The appraisal framework goes beyond limitations of traditional concepts of sentiment and opinions and considers the appraisal expression, which is a linguistic unit by which a personal appreciation is conveyed, whether it is an opinion, sentiment or supportive statement.

Appraisal theory structures appraisal expressions under three main basic systems describing attitudes, engagement and graduation, see fig. 1.

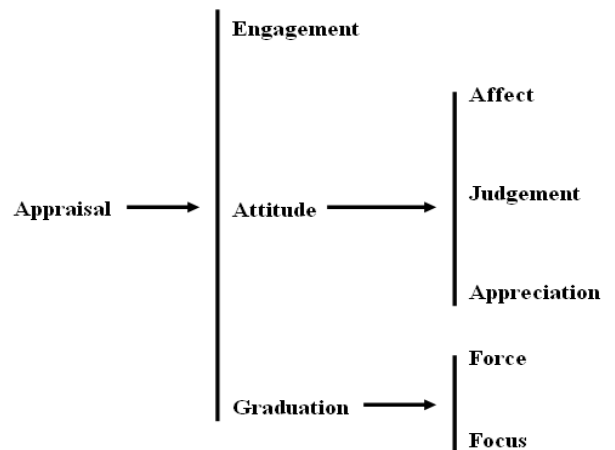


Fig. 1. Main systems of the appraisal theory

The attitude system is related to linguistic expressions conveying the current of authors at the time they write the text and. This system covers three main subcategories: affect, appreciation and judgement, discussed hereafter.

- Affect is related to linguistic expressions of author's feelings such as happiness, joy, sadness, grief, etc..., as shown in sentences hereafter:

S1: This movie bores me.

S2: Helping others makes me happy.

- Judgment highlights linguistic expressions conveying characterization of persons and behaviors by the author. Generally it conveys opinions and personal tastes about objects, such as nice, ugly, beautiful, shy but also about interactions and behaviors in the social context: heroic,

brave, open-minded, feeble-minded, see sentences below:

S4: *Paul is **rude**, but **honest**.*

S5: *They are a **kind** nation.*

- Appreciation is related to assessment and evaluations of entities, objects, events and scenes, as shown in examples:

S5: *The painting is **beautiful**.*

S5: *My phone is **useless**.*

The engagement system gathers linguistic expression specifying the author's position with respect to his own statements. When reporting, writers often embed clues as to how strongly they support the content being conveyed and may indicate confidence, doubt, skepticism, conviction, etc., about the information reported. The engagement system is closely related to the notions of trust, confidence, probability or possibility.

Categories under this system see fig. 2 encompass aspects related to denial (S6), concession (S7), confirmation (S8), endorsement (S9), acknowledgement (S10) and distance (S11).

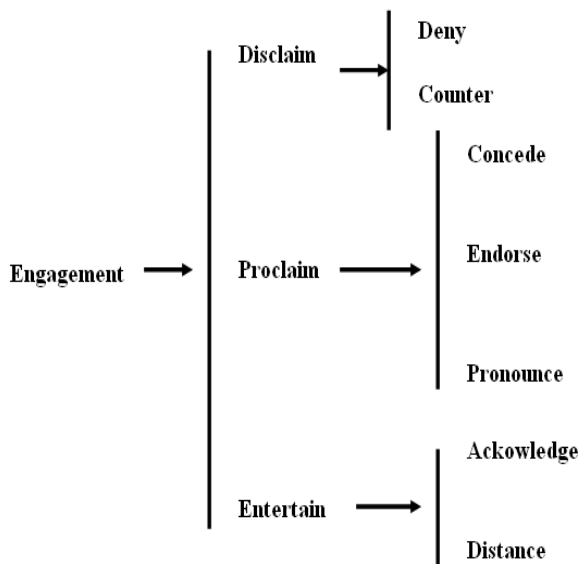


Fig. 2. The engagement system

S6: *You **don't need** to access the file.*

S7: ***Although** it was raining we went out.*

S8: ***Of course** Brexit was unexpected*

S9: *The reports **show** he was involved in the accident.*

S10: ***According to** Times, he is leading the election*

S11: ***Many are claiming** that he will not win.*

The graduation is the last system of the appraisal theory and it is introduced in order to provide means to measure or at least

estimate the orientation and various degrees of intensity associated to affect, appreciations, judgment and engagement. The system is composed of Force and Focus, and is conveyed by linguistic modifiers such as intensifiers (very, enough, etc.) and downtoners (few, low, etc.).

In order to implement automatic procedures to identify appraisal expressions, a formal ontology[36] was built to represent main concepts and their linguistic instances and make explicit relationships. The construction of this ontology is described in the next chapter.

IV. A FORMAL MODEL OF APPRAISAL CATEGORIES

The appraisal ontology was built in order to have a general description of appraisal categories and to provide a formal model of concepts and the set of relations. The model was created from scratch, as there are no appraisal experts and the construction process was guided when needed by two existing general resources: WordNetv[37] and SenticNet [26]. The main step of ontology construction is conceptualization, which identifies main concepts and relations. The model created with Protégé [38] and formalized using OWL DL language [39].

A. Modeling concepts

Starting with categories introduced by the appraisal theory, and adding additional concepts to characterise appraisal expressions, six main concepts were identified (fig. 3) to model: attitude, engagement, graduation, orientation, polarity and modifiers.

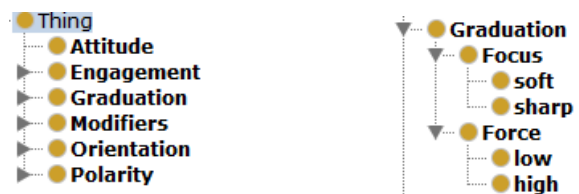


Fig. 3. Main concepts of the ontology Fig. 4 Graduation concept

Among appraisal categories, attitude and engagement are modeled as introduced by the appraisal theory, but with finer description of their sub-categories, as shown in tab. 1 and II.

TABLE I. SUBCLASSES OF ATTITUDE SYSTEM

Classes	Attitude System	
	Sub-classes	Instances
Affect	-	Happy, joyful, miserable
Appreciation	Composition	Detailed, consistent

Classes	Attitude System	
	Sub-classes	Instances
	Impact	Amazing, monotonous
Judgement	Valuation	Innovative, profound
	Social esteem	Brave, disloyal , clever
	Social sanction	Virtuous, corrupt , honest

TABLE II. SUBCLASSES OF ENGAGEMENT SYSTEM

Classes	Engagement System	
	Sub-classes	Instances
Disclaim	Deny	No, didn't, never
	Counter	Yet, although, but
Proclaim	Concur	Naturally, obviously
	Pronounce	indeed
	Endorse	Demonstrates, shows, proves
Entertain	-	Perhaps, probable, apparently
Attribute	Aknowledge	It's said that, many argues,
	Distance	He claimed to

Graduation concept was enriched by adding a several subclasses to describe Focus as sharp or soft and Force as high or low, as shown in fig. 4. For all those categories linguistic examples were also added at instance level.

Modifiers shown in fig. 5 is a concept intended to capture the ability of linguistic markers to increase (fig. 6), decrease or reverse the intensity of appraisal expressions.

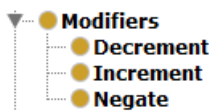


Fig. 5 Modifiers concept



Fig. 6 Incremental modifiers

Orientation is a concept used to capture positive and negative emotional states conveyed by appraisal expressions and polarity is used to make explicit the presence of negation markers (in this case polarity is marked otherwise it is considered as unmarked).

B. Modeling relations

Besides is-a relations used to create the hierarchical structure of concepts, 6 relations were added to connect concepts on the ontology, see tab. III.

TABLE III. ONTOLOGY RELATIONS

Name	Relation	
	Source	Target
hasGraduation	Engagement	Graduation
hasPolarity	Engagement	Polarity
hasOrientation	Engagement	Orientation
Increase	Modifier	Force
Decrease	Modifier	Force

C. Description of the ontology

The model is composed of 46 concepts with a 6 levels hierarchy; the model also has 4 Object Properties and 2 DataType Propertie, 50 nodes, 75 edges and 268 instances of concepts.

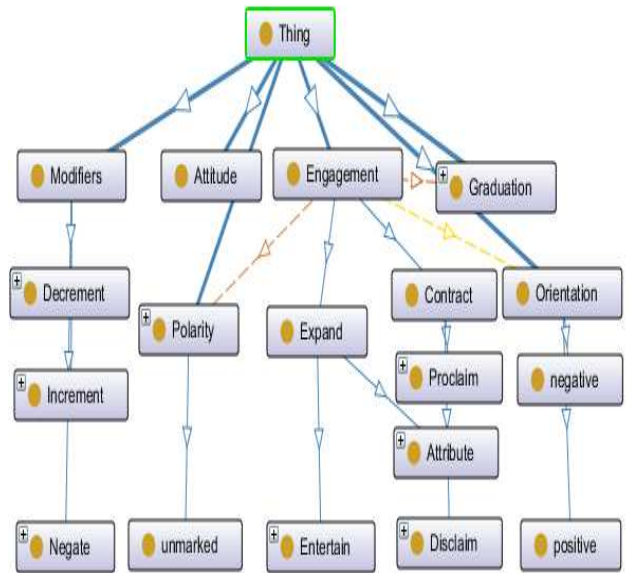


Fig. 7 Ontology of appraisal categories

Once concepts and relationships were modeled by ontology, linguistic instances were added by using WordNet, a lexical resource for English clustering similar terms into synsets and highlighting synonymy and antonymy relations. Association of concept to orientation and force was guided by SenticNet, a resource providing polarity associated with 100,000 natural language concepts that are also described in terms of four affective dimensions (Pleasantness, Attention, Sensitivity, and Aptitude).

V. ILLUSTRATION ON TWITTER ANALYSIS

A. Application context

The appraisal ontology was used to analyse a collection of tweets. The application context adopted is a social media exploration task that is important for homeland security and defence analysts, who are trying to achieve a clearer picture of emerging online phenomena. For those domains there is a practical need to investigate the way subjective content is expressed in social media, going above the limitations of opinion and sentiment notions. We argue that given the diversity of topics covered, the variety of user profiles, the huge volume of data submitted on a daily basis and the difficulties of social data processing [5], [21] an effective solution should be model-driven and we describe hereafter a practical illustration on using the appraisal ontology to explore social data.

B. Methodology and experimental datasets

In order to analyse expressions of subjective content on social media, we adopt a semantic annotation approach augmented with processing methods capable to perform a quantitative analysis of data gleaned on social media. The approach consists of several phases implemented to gather, process and analyse social data, as shown in fig. 8.

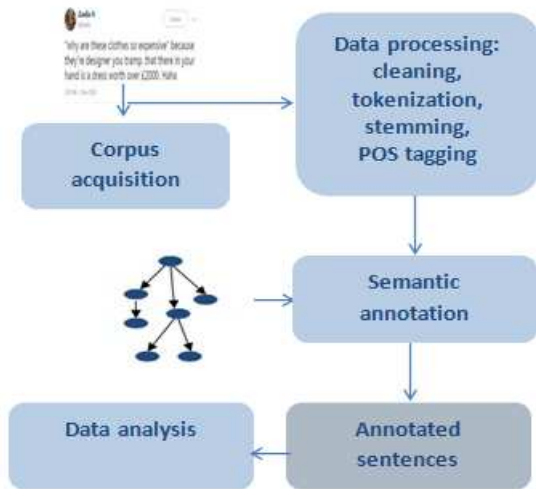


Fig. 8 Data processing and analysis

- **Data acquisition:** is done by crawling several sets of tweets with specific keywords and additional constraints to select only posts written in English. This phase was carried out as a straightforward step and relies only on data content of both keywords and posts and the ability of API used to mine the social network.

- **Data processing:** performs first a cleaning step, by removing URLs, hashtags and any information considered as irrelevant. While keeping tweets individually, each tweet is considered as a text paragraph. Processing at paragraph level includes sentences identification according to punctuation marks, tokenizing, part-of-speech (POS) tagging along with identification of words stems.
- **Semantic annotation:** attaches additional information to various text paragraphs based on their content analysis. Semantic annotation is performed automatically, by using lexical similarities measures that associate a real number to a pair of words and offers a measure of the degree to which two words are similar. For this work, lexical similarities are used to label words to concepts of the appraisal ontology.
- **Data analysis:** was carried out based on previous annotations and analyses each set of data by estimating; the percentage of subjective. vs. objective tweets; the percentage of tweets having positive vs. negative orientations; the distribution of tweets according to their low, medium or high strength and also their distributions with respect to *Attitude* and *Engagement* systems, and also with a finer distribution according to their specific concepts.

The experimentation was carried out by using five data sets and results are described hereafter.

C. Analysis of results and discussion

The overall collection was crawled using five keywords from December 2017 to February 2018 and it is structured into five data sets:

- Set 1: keyword ‘western values’, number of tweets: 7956;
- Set 2: keyword ‘white supremacy’, number of tweets: 7542;
- Set 3: keyword ‘Iraq War’, number of tweets: 7001;
- Set 4: keyword: ‘western coalition’, number of tweets: 6971
- Set 5: keyword ‘security’, number of tweets: 7509.

TABLE IV. ANALYSIS OF DATA SETS

Value	Data set				
	Set1	Set2	Set3	Set4	Set5
Subjective tweets	80%	88%	93%	93%	80%
Attitude annotation	30%	33%	64%	47%	38%
Positive tweets	55%	51%	53%	76%	51%
High force	19%	23%	21%	9%	20%
Low force	11%	12%	22%	5%	11%

Table IV shows the analysis of data sets in terms of numerical distribution of subjective and objective tweets, positive and negative orientation, high and low force along with a finer analysis in the light of appraisal categories.

Numerical values show a high percentage of subjective tweets for all datasets analyzed, regardless of the keyword. Tweets conveying *Attitude* concepts are generally less than 50%, with an exception for data set3. Moreover there is a good representation of positive-oriented tweets that account for more than 50% of the collection, while tweets having low or high impact are underrepresented for all data sets.

TABLE V. ANALYSIS OF ANNOTATIONS

Value	Data set				
	Set 1	Set 2	Set 3	Set4	Set5
Max Att.	capacity	valuation	tenacity	capacity	tenacity
Min Att.	social esteem	impact	veracity	appreciation	impact
Max Eng.	counter	affirm	distance	affirm	distance
Min Eng.	concede	concede	Pro nounce	endorse	concede
Att Cat	10	10	10	8	10
Eng Cat	7	8	8	7	8

Table V illustrates the distribution of the most and least frequents sub-concepts of *Attitude* and *Engagement* specific within each data set. Thus, *western values* keyword is more related to *capacity* and *counter* while being and less related to ‘*Social esteem*’ and *concede*. *White supremacy* shows strong correlations with *valuation* and *affirm*, and low connections with *impact* and *concede*. *Irak war* is strongly connected to *tenacity* and *distance*, and low linked to *veracity* and *pronounce*. The keyword *Western coalition* shows high occurrences for *capacity* and *affirm* and low occurrences for *appreciation* and *endorse*. The last data set gathered with *security* keyword is strongly correlated to *tenacity* and *distance* and lousy linked to *impact* and *concede*.

At collection level, among *Attitude* concepts, *capacity* and *tenacity* are the two concepts that are more frequently strongly correlated to data, while the set of less correlated concepts is more diverse and includes *impact*, *social esteem*, *veracity*, and *appreciation*. Under *Engagement* system, both sets of most and least frequent concepts have rather homogeneous compositions, with *affirm* and *distance* as the most frequent classes and *concede* as the less frequent concept.

The analysis also shows a rather similar distribution of data sets into categories of appraisal systems, with around 10 classes for *Attitude* and 7 or 8 classes for *Engagement*.

VI. CONCLUSION

This paper investigates the use of appraisal categories to explore social data mining. The approach goes beyond the generally accepted definitions of sentiment and opinions, and focuses on appraisal expressions, as introduced by the appraisal framework to describe the way humans express their attitudes, appreciations, and engagement.

The application context of this work is a cyberspace exploration task designed to support defence and homeland security intelligence practitioners in their efforts to gather valuable data allowing them to understand emergent phenomena, such as online hate proliferation or online propaganda. Linguistic clues of appraisal categories are used as indicators of subjective content to be collected and further analysed in order to understand the way online users express their extreme altitudes, embrace or support extreme ideologies and ideas.

Starting with the three systems introduced by the appraisal theory, a semantic resource was build modeling finer categories under each system, and highlighting terms and expressions specific to each category. The resource was used to analyze a collection of Tweets gathered using several keywords related to hate, racism and far right extremism.

Future work focus on adopting new strategies to use the model in various practical contexts, such as online trend analysis, providing a practical solution to analyse the dynamics of subjective content and highlighting the way opinions on specific topics change over the course of time.

ACKNOWLEDGMENT

This document has been produced in the context of the PRACTICIES project. Research leading to these results has received funding from the European Union’s Horizon 2020 research and innovation program under grant agreement No 740072.

REFERENCES

- [1] J.R. Martin, and P. R. White, The language of evaluation(Vol. 2). Basingstoke: Palgrave Macmillan, 2003
- [2] Z., Luo, M., Osborne, and T., Wang . An effective approach to tweets opinion retrieval. World Wide Web, 18(3), 545-566, 2015
- [3] D., Tang, F., Wei, N., Yang, M.,Zhou, T., Liu, and B., Qin, B., Learning Sentiment-Specific Word Embedding for Twitter Sentiment Classification. 2014
- [4] P. W., Liang and B. R. Dai, Opinion mining on social media data. In Mobile Data Management (MDM), 2013 IEEE 14th International Conference on (Vol. 2, pp. 91-96). IEEE, 2013
- [5] D., Maynard, K.,Bontcheva, and D., Rout. Challenges in developing opinion mining tools for social media. Proceedings of the@ NLP can u tag# usergeneratedcontent, 15-22., 2012.
- [6] P., Gamallo, and M., Garcia, M. (2014, August). Citius: A Naive-Bayes Strategy for Sentiment Analysis on English Tweets. In SemEval@ COLING (pp. 171-175), 2014.

- [7] Y., Mejova, and P., Srinivasan. Crossing Media Streams with Sentiment: Domain Adaptation in Blogs, Reviews and Twitter, 2012. In ICWSM.
- [8] S. Stieglitz, L., Dang-Xuan, A., Bruns, and C. Neuberger, Social media analytics. *Business & Information Systems Engineering*, 6(2), 89-96, 2014.
- [9] D., Bollegala, D. Weir, and J., Carroll, Cross-domain sentiment classification using a sentiment sensitive thesaurus. *IEEE transactions on knowledge and data engineering*, 25(8), 1719-1731, 2013.
- [10] M., Tsytzarau, and T., Palpanas. Survey on mining subjective data on the web. *Data Mining and Knowledge Discovery*, 24(3), 478-514, 2012.
- [11] R., Xia, and C., Zong (2011). A POS-based Ensemble Model for Cross-domain Sentiment Classification. In *IJCNLP* (pp. 614-622), 2011.
- [12] B., Bing. Sentiment analysis and opinion mining. *Synthesis lectures on human language technologies*, vol. 5, no 1, p. 1-167, 2012.
- [13] M., Thelwall, M., K., Buckley, and G., Paltoglou, Sentiment strength detection for the social web. *Journal of the Association for Information Science and Technology*, 63(1), 163-173, 2012.
- [14] S. Li, Y., Xue, Z., Wang, and G. Zhou, . Active Learning for Cross-domain Sentiment Classification. In *IJCAI*(pp. 2127-2133), 2013.
- [15] S., Bhuta, A., Doshi, U., Doshi, and M., Narvekar, A review of techniques for sentiment analysis Of Twitter data. In *Issues and Challenges in Intelligent Computing Techniques (ICICT)*, 2014 International Conference on (pp. 583-591). IEEE., 2014.
- [16] S., Liu, F., Li, X., F., Li, Cheng, X., and H., Shen. Adaptive co-training SVM for sentiment classification on tweets. In *Proceedings of the 22nd ACM international conference on Information & Knowledge Management* (pp. 2079-2088). ACM, 2013.
- [17] R., Socher, A., Perelygin, J., Wu, J., Chuang, C. D., Manning, A., Ng, and C. Potts, Recursive deep models for semantic compositionality over a sentiment treebank. In *Proceedings of the 2013 conference on empirical methods in natural language processing* (pp. 1631-1642), 2013.
- [18] M. Taboada, J., Brooke, M., Tofiloski, K., Voll, and M., Stede. Lexicon-based methods for sentiment analysis. *Computational linguistics*, 37(2), 267-307., 2011.
- [19] E., Cambria, B., Schuller, Y., Xia and C., Havasi, New avenues in opinion mining and sentiment analysis. *IEEE Intelligent Systems*, 28(2), 15-21, 2013. F. H., Khan, S., Bashir, nand U., Qamar, TOM: Twitter opinion mining framework using hybrid classification scheme. *Decision Support Systems*, 57, 245-25, 2014.
- [20] E., Kouloumpis, T., Wilson, and J. D. Moore, Twitter sentiment analysis: The good the bad and the omg!. *Icwsn*, 11(538-541), 164, 2011.
- [21] P., Korenek, and M. Šimko, Sentiment analysis on microblog utilizing appraisal theory. *World Wide Web*, 17(4), 847-867, 2014.
- [22] A. Balahur, J.M., Hermida, and A., Montoyo., Building and exploiting emotinet, a knowledge base for emotion detection based on the appraisal theory model. *IEEE Transactions on Affective Computing*, 3(1), 88-101., 2012.
- [23] C., Soo-Guan Khoo, A., Nourbakhsh, and J. C., Na, Sentiment analysis of online news text: a case study of appraisal theory. *Online Information Review*, 36(6), 858-878, 2012.
- [24] E., Boiy, and M.F., Moens, A machine learning approach to sentiment analysis in multilingual Web texts. *Information retrieval*, 12(5), 526-558, 2009.
- [25] E., Cambria, C., Havasi, and A. Hussain, SenticNet 2: A Semantic and Affective Resource for Opinion Mining and Sentiment Analysis. In *FLAIRS*, 2012.
- [26] A. Pak and P. Paroubek. „Twitter as a Corpus for Sentiment Analysis and Opinion Mining". In *Proceedings of the Seventh Conference on International Language Resources and Evaluation*, pp.1320-1326, 2010.
- [27] L. Barbosa and J. Feng. "Robust Sentiment Detection on Twitter from Biased and Noisy Data". *COLING 2010: Poster Volume*, pp. 36-44., 2010.
- [28] B., Agarwal, I., Xie, O., Vovsha, R., Rambow, R. Passonneau, "Sentiment Analysis of Twitter Data", In *Proceedings of the ACL 2011 Workshop on Languages in Social Media*, 2011, pp. 30-38, 2011
- [29] A. Aue and M. Gamon, Customizing sentiment classifiers to new domains: a case study. In *Proc. of the International Conference on Recent Advances in Natural Language Processing*, Borovetz, Bulgaria, 2005;
- [30] E. Boiy, Pieter Hens, Koen Deschacht, and Marie-Francine Moens. 2007. Automatic sentiment analysis of on-line text. In *Proc. of the 11th International Conference on Electronic Publishing*, Vienna, Austria, 2007.
- [31] A.C. Mullaly, C.L. Gagné, T.L. Spalding, and K.A. Marchak. Examining ambiguous adjectives in adjective-noun phrases: Evidence for representation as a shared core-meaning. *The Mental Lexicon*, 5(1):87-114, 2007.
- [32] A. Weichselbraun, S. Gindl, and A. Scharl. . A context-dependent supervised learning approach to sentiment detection in large textual databases. *Journal of Information and Data Management*, 1(3):329-342., 2010.
- [33] T. Wilson, J. Wiebe, and P. Hoffmann.. Recognizing contextual polarity: An exploration of features for phrase-level sentiment analysis. *Computational Linguistics*, 35(3):399-433, 2009.
- [34] S. Gindl, A. Weichselbraun, and A. Scharl, Crossdomain contextualisation of sentiment lexicons. In *Proceedings of 19th European Conference on Artificial Intelligence (ECAI-2010)*, pages 771-776, 2010.
- [36] T. Gruber, A translation approach for portable ontology specification, *Knowledge Acquisition* 5(2), 199-220, 1993.
- [37] G. A. Miller, WordNet: a lexical database for English, *Communications of the ACM*, 38(11), 39-41, 1995.
- [38] N. F., Noy, M., Crubézy, R. W., Ferguson, H., Knublauch, S.W., Tu, J., Vendetti, and M. A., Musen, Protege-2000: an open-source ontology-development and knowledge-acquisition environment. In *AMIA Annu Symp Proc* (Vol. 953, p. 953), 2003.
- [39] S., Bechhofer, OWL: Web ontology language. In *Encyclopedia of Database Systems* (pp. 2008-2009). Springer US, 2009.