Embedded Human Computation for Knowledge Extraction and Evaluation (uComp)

D3.3: HC-Based Pattern Discovery

10 May 2016
Version: 1.0

Version History

Version	Date	Author	Comments
0.1	26/08/2015	Arno Scharl	Initial Draft
0.2	02/09/2015	Stefan Gindl	Major Revision
0.3	10/09/2015	Arno Scharl	Feedback
0.4	15/09/2015	Adrian Brasoveanu	Minor Update
0.5	18/04/2016	Arno Scharl	Minor Update
0.6	05/05/2016	Stefan Gindl	Major Revision
1.0	10/05/2016	Arno Scharl	Final Revision

Dissemination Level: PU – Public

uComp receives the funding support of EPSRC EP/K017896/1, FWF 1097-N23, and ANR-12-CHRI-0003-03, in the framework of the CHIST-ERA ERA-NET program line.

Table of Contents

Overview

The uComp deliverable D3.3 connects factual and affective knowledge to identify attitudes such as uncertainties vis-à-vis certain events. Recent trends in computational linguistics indicated a shift from the initially proposed research roadmap. Instead of using machine-learning based pattern extraction using R and representing the gathered knowledge with the *Predictive Model Markup Language* (PMML), the usage of scalable triple store technology accessible via SPARQL queries as well as a REST API (Application Programming Interface) proved more flexible to manage and retrieve concepts required for hybrid approaches to processing factual and affective knowledge - for example, to structure and extend sentiment lexicons, or to disambiguate named entities when assigning sentiment values (Cambria, 2016).

We tackle the goal in two different ways: (i) the creation of a comprehensive *semantic knowledge base*, that combines the knowledge of linguistic repositories in one structure and (ii) an algorithmic approach to enrich existing affective resources with factual knowledge to enhance their comprehensiveness.

The first approach, i.e. combining linguistic knowledge bases, culminated in the creation of a semantic knowledge base. The purpose of this structure is two-fold: (i) to store factual and affective knowledge and (ii) to allow reasoning across both types of knowledge. It bridges the gap between WP4 (Factual Knowledge Extraction and Evaluation) and WP5 (Affective Knowledge Extraction and Evaluation) and funnels the combined information for the HC-based evaluation developed in WP2 (Human Computation Framework).

The technology stack of the semantic knowledge base uses a small ontology for integrating the various resources designed with Protégé (Musen, 2015), Sesame[1] (renamed RDF4J) for storing triples and reasoning, and SPARQL (Harris and Seaborne, 2013) as query language. Sesame holds the knowledge of well-established sources, such as WordNet (Miller 1995, Fellbaum 1998), DBPedia (Lehmann et al. 2014), ConceptNet (Speer and Havasi, 2013), and SenticNet (Cambria et al. 2014). On top of that it embeds automatically extracted linguistic information such as document keywords, their associated sentiment and background information, in a customized namespace. This allows reasoning across the otherwise separate and disjunct knowledge resources. The flexibility of SPARQL facilitates access to this rich and diverse knowledge repository. Accessing factual and affective knowledge in this way helps to identify uncertainty towards facts, which can also be an indicator of fear towards that fact.

The backbone of the semantic knowledge base is *Sesame*, storing many million triples of information. Since reasoning in such an extensive knowledge base becomes computationally intensive (Aranda, Hogan, Umbrich and Vandenbussche, 2013), an

[1] Sesame is going to be forked into the Eclipse RDF4J project (www.rdf4j.org). At the time of implementation, Sesame was still available in its original form.

Elasticsearch cluster put in front of it caches frequent queries and absorbs heavy-duty queries. This architecture yields high throughput, facilitating its usage in real-time system which would otherwise suffer from long response time (the infamous query timeouts from SPARQL), turning their usage into a time-consuming process.

As mentioned, the semantic knowledge base is not a mere container for already existing knowledge bases via an ontology. It is a dynamic structure allowing to also store newly and automatically gathered knowledge. Specific learning algorithms connect keywords with affective statements. This *target sentiment analysis* leverages the *Recognyze* named entity recognition and resolution framework (Weichselbraun et al., 2015) to detect keywords and applies linguistic heuristics and machine learning to identify connections to affective terms. Such fine-grained sentiment analysis goes beyond widely used but simple algorithms for the classification of documents or sentences. Detailed insight into the conveyed opinion becomes possible.

The semantic knowledge base covers the upper part of the uComp architecture shown in Figure 1. It provides an effective mechanism to store and manage the metadata elements used to structure the *domain content repository* (T1.5), and provides linked data services. Furthermore, by connecting knowledge bases from different domains and serving as a space for knowledge extracted from free-text sources, it allows for advanced pattern discovery using the flexible SPARQL query language.

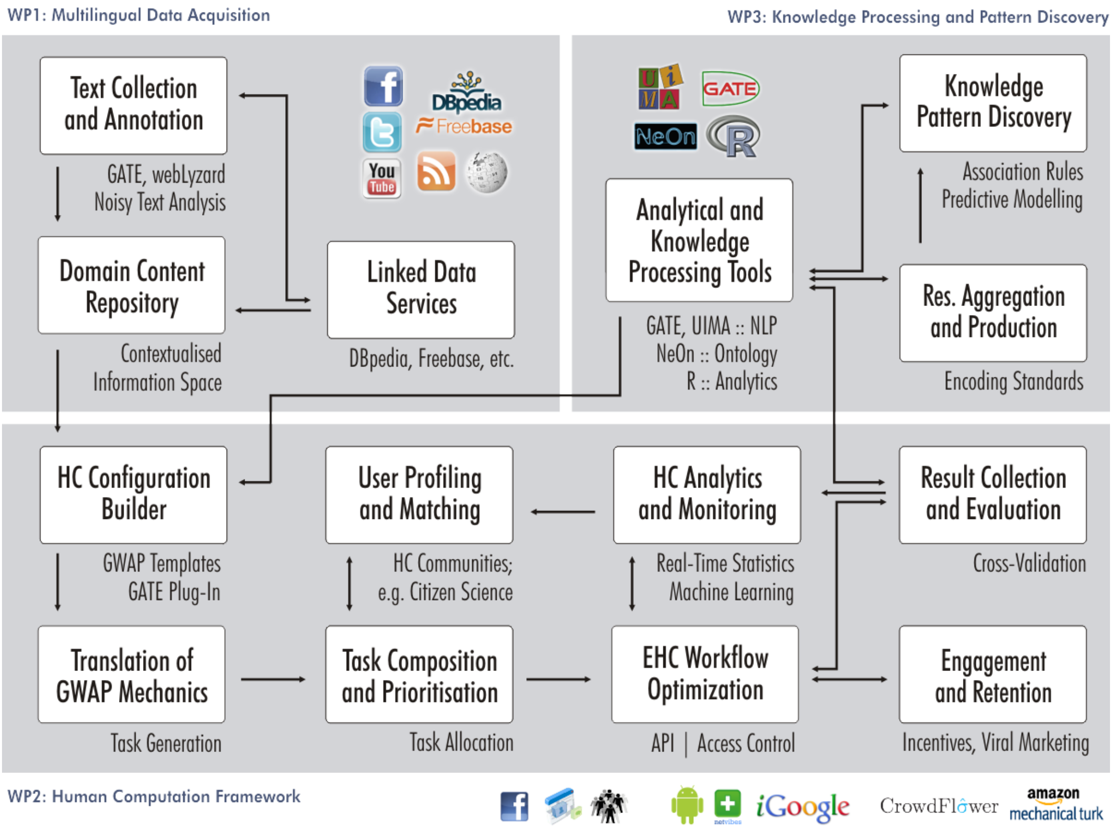

Figure 1. Overview of the uComp System Architecture

The second approach, i.e. algorithmically enriching factual knowledge bases, was driven by the need to complement existing resources with the knowledge contained in Web document collection, i.e. unstructured data. Analysing the distribution of sentiment terms in a document collection allows gathering hints for the inclusion of further terms helping to refine the quality of the affective resource. This approach has proven useful in previous research (Gindl et al., 2010). The subsequent lookup of the identified hints in factual resources further provides terms and concepts ready for the integration into the affective knowledge base. This yields an evolving affective resource with capabilities for application across domain and previously unknown topics.

An evaluation leveraging Human Computation (HC) helped to assess the efficacy of the approach while concurrently limiting the manual labour for the researchers. An evaluation of different crowd-sourcing methods (i.e. games-with-a-purpose vs. paid crowdsourcing marketplaces such as Amazon Mechanical Turk[2] and CrowdFlower[3]) helped to choose the most appropriate tool for a given task. On top of that, we analysed the rhetorical structure in the text to improve existing sentiment analysis methods. Particular rhetorical structures indicate high difficulty for the classification task and therefore should become candidates for special consideration.

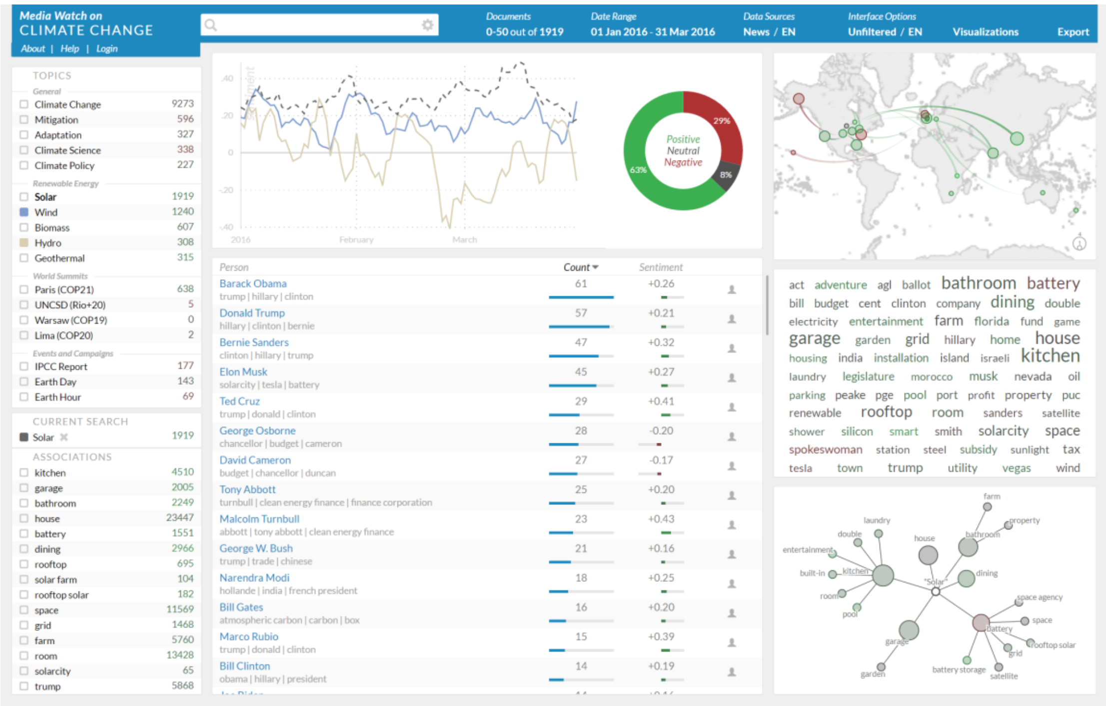

Figure 2. Screenshot of the Media Watch on Climate Change

Figure 2 shows a screenshot of the *Media Watch on Climate Change*,[4] analysing solar energy news media coverage between January and March 2016. By revealing patterns in the *domain content respository* (T1.5), the visual dashboard supports the integrated analysis of factual and affective knowledge. The shown example includes

[2] www.mturk.com
[3] www.crowdflower.com
[4] www.ecoresearch.net/climate

associated n-grams (list of associations, tag cloud, keyword graph), a trend chart reflecting the overall distribution of sentiment and its development of time, and a list of extracted named entities including the number of their mentions and the average sentiment expressed towards them.

Semantic Knowledge Base

The requirements for the semantic knowledge base were: (i) flexibility to store knowledge bases from different domains and namespaces, (ii) ability to handle extensive amounts of data, (iii) ease the reasoning across domains, and (iv) scalability to cater for the needs of automated media monitoring tools. The technology stack to achieve this consists of the following software tools:

- A minimal ontology for integrating resources from different Knowledge Bases
- A Sesame (RDF4J) triple store for data persistence and reasoning
- SPARQL as query language
- An Elasticsearch cluster to cache pre-defined and frequent queries.

Architecture

The architecture of the semantic knowledge base consists of three major parts: a serialization interface, tailored to the needs of the different and often diverse data sources, a triple store to permanently hold the data, and an query interface for communication with the triple store (see Figure 3).

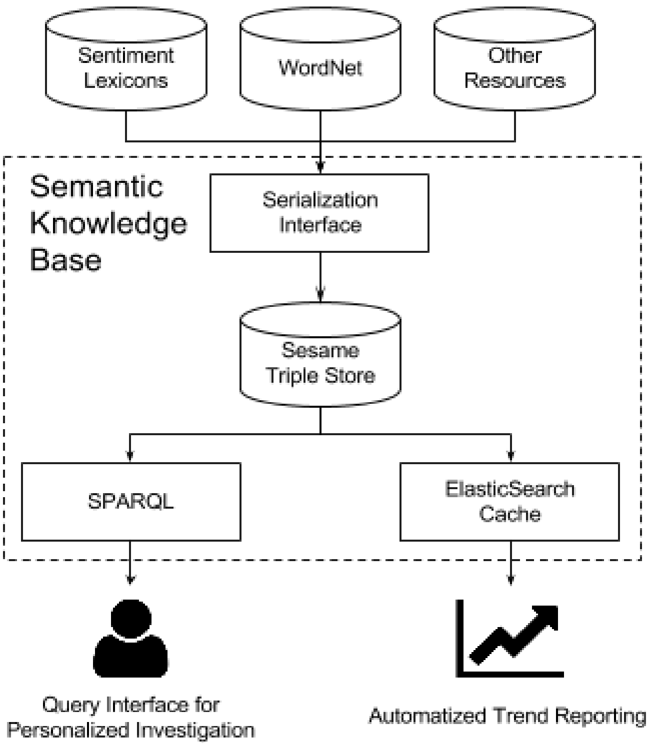

Figure 3. Architecture of the semantic knowledge base

The primary input is the serialization interface, bundling multiple services to turn external resources into triples and transfer them into the semantic knowledge base. The triples are stored in a Sesame (RDF4J) triple store. Sesame is either accessible directly via its built-in interface OpenRDF Workbench, useful for manual querying of the semantic knowledge base, and via a customized cache, providing scalability and high-throughput methods to cater for the needs of automatic trend monitoring tools.

Serialization

The diversity of the data sources requires the establishment of customized serialization services, bundled in the serialization interface (see Fig. 2). Each service is tailored to the specific setup of the data source. Data sources in RDF require the least effort, as they allow leveraging the built-in tools of the Sesame platform. A direct upload of the data without any additional effort is possible. Other data sources follow different standards. Two different approaches towards tackling this challenge are feasible: (i) investigating and identifying existing namespaces usable with only little need for adaptation, or (ii) defining and establishing a dedicated namespace for the resource. The latter option is labour-intensive. For that reason, we always intensively investigated for potential existing solutions that fit our needs.

Persistence

We use an RDF triple store as the container of for all the collected and utilized knowledge resources, dynamically updated based on collected evidence - including the results of crowdsourcing experiments as well user-generated content streams from Twitter and other platforms. The triple store architecture is realized with Sesame (now RDF4J), a Java-based framework for RDF. The choice of Sesame has been straightforward - it represents a well-known and tested triple store that scales well to extensive amounts of data, and easy to set up with standard server technology.

Querying

Querying the knowledge base happens two-fold. One way is manual access via the SPARQL query language. This allows flexible and personalized queries well-suited for research and investigation purposes. The second way is automatized querying, where queries are predefined and potentially executed repeatedly in short time frames, calling for high-throughput methods. This requires a different technological approach as compared to manual queries.

Ad-Hoc Queries

This type of query is highly customized and usually not time-critical. Sesame has built-in functionality to execute SPARQL queries, the so-called OpenRDF Workbench, which is the logical choice for manual queries.

Predefined Queries

Predefined queries are tailored to the requirements of an automatic trend reporting tool. This tool needs access to the same type of information repeatedly and in a timely manner. The semantic knowledge base fulfils this requirement by providing a cached version of query results. The backbone of the cache is an Elasticsearch cluster. This ensures high scalability and throughput.

Filling the cache happens either in timely intervals, for instance after the definition of new requirements. A scripted service derives the data from the triple store and puts it into the cache, ready to be consumed from the automatic tool. The second way of filling it is after the execution of a formerly unknown query. An identifier as well as the query result become persistent in the cache and are available for future retrieval.

Reasoning

Sesame (RDF4J) repositories can be configured to use various reasoning services depending on the use case and inference regimes. We generally prefer to use ontop/Quest (Calvanese et. al., 2015) as it is well-integrated with Sesame, supports both RDFS and OWL 2 QL profiles and is known to perform well when integrating multiple types of resources (Knowledge Bases, relational databases, NoSQL databases, etc.), or Pellet (Sirin et. al., 2007) as it is widely considered to be the fastest reasoner available. Our main use cases for reasoning are inferring new facts and/or attributes about existing concepts and entities, and aligning multiple types of resources.

Algorithmic Connection of Factual and Affective Knowledge

Bridging the gap between factual and affective knowledge is at the core of this deliverable. We achieve this goal in two ways: (i) by merging knowledge repositories from the factual/affective domain, leveraging the described semantic knowledge base, and (ii) by enriching affective knowledge bases with factual features. The advantage of the first way is the reusability of existing knowledge bases and the creation of new knowledge by conjoining them. The strength of the second approach is its applicability on unstructured data sources, i.e. document collections, and the potential to connect the extracted knowledge with existing repositories.

Merging Knowledge Repositories

This use case leverages the semantic knowledge base. It effectively stores multiple, different knowledge repositories, which follow different naming conventions, and makes their diverse knowledge available via the SPARQL query language. There is no need to standardize the namespaces of the repositories, as SPARQL provides the means align differentiating concepts. Retrieving knowledge about a car and connecting it to affective knowledge gathered from a repository such as SenticNet and real-world media coverage stored as keywords in the semantic knowledge base becomes possible. For an investigation it is sufficient to define the repositories needed. Subsequently, prefixing query triples with the respective identifiers takes care for the correct alignment.

For instance, investigating car companies and their contribution in the emission scandal on the media will deliver brand names, car names, car components, involved persons, associated topics, potentially affective terms, etc. Such an investigation in unstructured data, i.e. document collections, strongly benefits from a knowledge repository that interconnects these particles. Feeding the particles into the semantic knowledge base reveals details such as that a “gearshift” is a component of a car, or that a specific person is the CEO of a particular car company. Further keywords will be identified as emotional terms by leveraging the affective components of the semantic knowledge base. Via the formerly established connections, e.g. components or relevant persons, it becomes possible to propagate emotions along the connections. This helps to unveil, for instance, that a sentiment associated with a component also affects the car it belongs to as well as the brand producing the car.

Automatic Enrichment of Sentiment Lexicons with Concept Knowledge

To understand and identify emerging trends media monitoring tools require a deep know-how of language as well as the emotions transported by news coverage, be it in curated online media or self-authored social media texts. Factual knowledge bases give access to the factual part of media coverage, but neglect their emotional content. WP3 addressed this shortcoming by merging affective knowledge into the knowledge bases WordNet and ConceptNet. WordNet is a database grouping nouns, verbs, adjectives, etc. into semantically related clusters, so-called synsets, providing lexical definitions and semantic relations between these clusters. ConceptNet, on the other hand, provides common-sense knowledge to enable computer systems to further understand the semantic context of a concept. Combining factual and affective knowledge creates synergies - the factual knowledge provided through Recognyze annotations, for example, helps to contextualize the sentiment analysis process, to process ambiguous sentiment terms, and to detect opinion holders and opinion targets. Assigning and continuously updating confidence values support the processing of uncertain knowledge, both in terms of the precision of the named entity recognition process as well as the disambiguation of sentiment terms.

WP3 combined WordNet and ConceptNet with the affective knowledge base SenticNet. In a first step the contextualization framework of the webLyzard platform enhanced SenticNet with context knowledge. This knowledge helps to overcome a shortcoming of affective knowledge bases such as SenticNet, i.e. the fact that polarity values associated with opinionated terms, do not adapt to context changes given in a text. The contextualization procedure adds this knowledge and thus increases the accuracy of SenticNet. Subsequently, the grounding of SenticNet terms further increases the coverage of the knowledge base. Human computation, executed by CrowdFlower tasks, provides the means for the evaluation of the entire process. Figure 4 gives an overview of the entire procedure, starting with the contextualization and ending with a concept store holding positive and negative concepts.

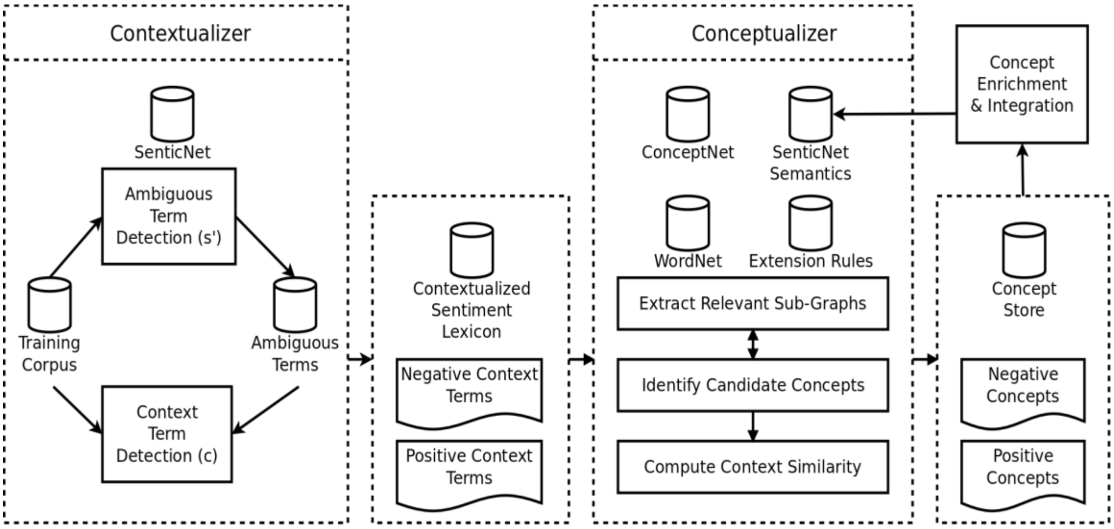

Figure 4. The conceptualization pipeline

Contextualization

Created in previous projects this procedure helps to disambiguate polysemous sentiment terms (Gindl et al., 2010). A polysemous sentiment term changes its polarity depending on the context it is used in, thus using a static polarity value is futile. The contextualization procedure identifies cues for disambiguation and stores them together with the respective sentiment term. The procedure comprises of three steps:

1. **Identification of ambiguous sentiment terms:** given a document collection as training input the system first identifies ambiguous sentiment terms to put them aside for the contextualization procedure.

2. **Integration of context terms:** The analysis of the used context is crucial information for the disambiguation of a polysemous sentiment term. Storing context terms, i.e. terms frequently co-occurring with this type of sentiment term, delivers the essential knowledge for disambiguation

3. **Improvement of document classification:** In the application phase the system matches the context of a given document with the stored context and determines the current polarity of ambiguous sentiment terms.

The data structure to hold contextual information is the so-called contextualized lexicon, which extends a regular sentiment lexicon with context terms and their probabilities to occur in positive/negative documents in conjunction with the ambiguous sentiment term.

Conceptualization

The conceptualization procedure enriches the contextualized lexicon created in the previous step with data from ConceptNet and WordNet (Weichselbraun et al., 2013, 2014). The common-sense knowledge of ConceptNet helps to level up the contextualized lexicon so it contains knowledge humans usually acquire by interacting with their environment, e.g. the knowledge that putting the hand above a burning candle results in a sensation of pain. Such knowledge is indispensable in next-generation sentiment analysis tools (Cambria et al., 2013). The conceptualization procedure follows a three-step process:

1. Extract positive and negative context terms from the contextualized lexicon
2. Query WordNet and ConceptNet for sub-graphs relevant for the respective ambiguous terms, context terms, their respective candidate concepts as well as their context terms
3. Similarity computation between the sub-graphs of the ambiguous terms and those for the candidate concepts

The procedure yields polarity information for the retrieved concepts for the integration into the underlying knowledge base, i.e. SenticNet. This further increases the expressiveness of the resource and improves its flexibility.

Table 1 contains context examples for qualified ambiguous terms as well as the concepts they have been grounded to by the explained procedure.

Table 1. Qualified ambiguous terms with their context terms as well as the ConceptNet and WordNet groundings

Term		**Context Term**	**ConceptNet**	**WordNet**
Adventure	+	during, nostradamus, diary	activity, magical journey, fun trip	wild and exciting undertaking
	-	educational, windvd, frame	software, band, video game	wild and exciting undertaking
Development	+	creating, dreamweaver, nduc	progression from simpler to more complex forms	growth
	-	onecare, paperport, auction	recent event that has some relevance for the present situation	development
God	+	reading, cuppa, hdd	one of greater rank or station or quality	deity
	-	folder, quicklaunch, netbook	an incorporal being believed to have powers to affect the course of human events	god

Analysis of Rhetorical Structure

Sentiment analysis is highly sensitive to language nuances and subtleties. Contextualization and conceptualization are two ways of overcoming these challenges and help to obtain a more holistic picture of an opinionated statement. However, there are further obstacles an algorithm can stumble over. These obstacles are not of semantic nature, i.e. they are not inherent to concept a term is referring to but rather to the linguistic purpose of the term. For instance, terms such as “but” or “in contrast” are used to depict contrasting statements. Identifying such patterns helps to further improve algorithms for sentiment analysis. WP3 covered this procedure by picking sentences which are hard to classify using the available sentiment analysis tool. In other words, the analysis takes sentences which are “unclassifiable”, i.e. classified incorrectly by the tool, and compares them with classifiable sentences. A statistical comparison, i.e. co-occurrence analysis, of bigrams special for each set of sentences delivers insight into language patterns requiring special treatment. The sentences come from the Sentiment Tree Bank (Socher et al., 2013) of the Stanford NLP Group.

The statistical comparison is completed by a manual assessment. Judging each bigram in its context, i.e. a set of example sentence, helps to identify those capable of changing the rhetorical course of a sentence in such a way that the sentence becomes impossible to classify by the algorithm.

Table 3. Bigram examples capable of changing the rhetoric course of a sentence.

Bigram	**Example**
never quite	however sincere it may be the rising place never quite justifies its own existence
isn’t nearly	there isn't nearly enough fun here despite the presence of some appealing ingredients
can’t help	the pivotal narrative point is so ripe the film can't help but go soft and stinky
falls short	falls short in explaining the music and its roots
even though	morton uses her face and her body language to bring us morvern's soul even though the character is almost completely deadpan
doesn’t quite	impostor has a handful of thrilling moments and a couple of good performances but the movie doesn't quite fly

The analysis revealed bigrams such as “feels like”, “look like”, “plays like” as highly common for unclassifiable sentences. The reason for that is the ambiguous usage of the term “like”, either as a verb expressing inclination towards a fact or as a comparator between two entities (“the car looks like a truck”). More subtle indicators are “good intentions” (e.g. “not an objectionable or dull film it merely lacks everything except good intentions”), “never really” (e.g. “it's fitfully funny but never really takes off”), or “yet another” (e.g. “yet another arnold vehicle that fails to make adequate use of his particular talents”). Table 3 contains a more comprehensive list of retrieved bigrams and example sentences.

Evaluation

The evaluation of a sentiment analysis application requires the availability of an annotated corpus, which serves as the ground truth. We used reviews from Amazon and IMDb as an annotated corpus. The advantage of such a strategy is manifold: it is easy to access the reviews in a sufficiently large number, they are per nature opinionated and supposed to express sentiment, and they already have a label attached to them, provided by the author of the reviews. These factors make them perfect for the evaluation.

We used five corpora in total, two from Amazon and three from IMDb. Each corpus contains 2.000 reviews. Table 4 shows detailed statistics about the corpora.

Table 4. Evaluation corpus statistics. Each corpus has 2.000 reviews.

Corpus	Total counts		Avg. per review	
	Sentences	Words	Sentences	Words
Amazon electronics	19.911	298.622	10	149
Amazon software	24.120	380.760	12	190
IMDb comedy	25.481	410.874	13	205
IMDb crime	30.155	494.686	15	247
IMDB drama	27.026	432.820	14	216

Enrichment Statistics

Table 5 summarizes the results of the domain adaptation and enrichment process. The contextualization yielded domain-specific positive and negative context terms, which can be used in context-aware sentiment analysis. The method extracted approximately four times more context terms for ambiguous sentiment terms in IMDb reviews than for Amazon reviews. This is in line with expectations since IMDb reviews contain more ambiguous sentiment terms and often include plot elements, which provide the contextualization routine with a rich selection of potential context terms.

The context terms aided in disambiguating 1339 (2369) out of 1366 (2417) sentiment terms considered ambiguous in the Amazon (IMDb) corpus. The conceptualization routine successfully grounded 1018 (74.5%) concepts of the Amazon corpus to ConceptNet nodes and 519 (38.0%) to WordNet senses. For the IMDb corpus, we were able to link 1637 (69.1%) concepts to ConceptNet and 857 (36.2%) to WordNet senses. These results show that the conceptualization routine is able to successfully leverage ConceptNet's higher expressiveness in terms of concepts and assertions.

Table 5. Enrichment statistics

	Amazon	**IMDb**
Contextualization		
- Positive context terms	793.948	2.060.333
- Negative context terms	549.120	2.608.472
ConceptNet grounding		
- Grounded concepts	1.018	1.637
- Positive	2.287 (2.141 unique)	3.649 (3.248 unique)
- Negative	2.072 (1.773 unique)	3.437 (2633 unique)
WordNet grounding		
- Senses and definitions	519	857
- Synonyms	3.015 (2.072 unique)	5.012 (3.245 unique)
- Antonyms	108 (94 unique)	159 (138 unique)

Table 5 also indicates how often the conceptualization routine was only able to ground a concept to either a positive or a negative ConceptNet node. These numbers underscore the previous conclusion that corpora with a richer selection of context terms will yield more groundings.

We used the grounded WordNet concepts to integrate WordNet senses and definitions as well as synonyms and antonyms into the knowledge base. Due to the high interconnectedness of ConceptNet, we did not include ConceptNet assertions into the refined knowledge base. It is more efficient to directly query a grounded concept on ConceptNet, rather than to replicate these data.

HC-based Evaluation

An HC-based evaluation helped assess the efficacy of the conceptualization procedure. HC gives access to a large pool of manual labourers quickly and in such a way that it is easy to set up. To extend the qualitative evaluation discussed in the previous section, we conducted a quantitative evaluation of the concept grounding by compiling a list of qualified ConceptNet groundings, i.e. those that included a part of speech tag and a short textual description of the ambiguous term's interpretation.

The evaluation corpus contained 897 positive and 725 negative concept groundings for Amazon, as well as 1273 positive and 968 negative groundings for IMDb. Each grounding was inspected by three human participants, which assigned a sentiment value between 5 (very positive) and 1 (very negative) to each grounded concept, yielding a total of 11,589 assessments. The average rating variance amounted to

0.17 (0.22) for Amazon (IMDb) review data. This evaluation provided insights into the nature of the observed ambiguities and the grounding process. The evaluators perceived 79% (72%) of the concepts from Amazon (IMDb) as neutral. The evaluation presented in Section 4.1 shows that this does not reflect their actual use in the review corpus, which emphasizes how difficult it is for human evaluators to determine a concept's polarity based on its definition alone.

For concepts considered polar in the Amazon (IMDb) dataset, the assessors agreed in 61% (68%) of cases. Restricted to positive groundings, this figure increases to 64% (71%). This positive bias can be explained by concepts such as attack, coma, debt and opposition, which can represent interesting elements of a movie plot. Similarly, addictive computer games simulating armed conflicts and warfare often receive five-star ratings.

As a result, the evaluation showed that concept grounding is a highly sensible task and that the polarity assessment of a concept, when taken out of context, turns out to be a highly difficult task even for humans.

Summary

Attempts to advance sentiment analysis tools strongly benefit from the integration of factual and affective knowledge bases. In this deliverable, we showed two promising approaches. The first approach, i.e. the creation of a *semantic knowledge base*, results in a structure combining multiple and diverse resources. Predefined rules and queries help to extract cues for trend monitoring applications that aggregate and analyse extensive amounts of textual data and aim at extracting sentiments and attitudes expressed within the text. The second approach *enriches* the *affective resources* by merging knowledge from unstructured document collections into them. A subsequent integration of features from affective knowledge bases further enhances their coverage. This strategy is fruitful in two ways. Affective knowledge helps to disambiguate polysemous sentiment terms and adds a further level of accuracy to existing systems. Vice-versa, enriching factual knowledge with affective knowledge depicts the sentiment associated with world facts and allows the creation of a clear picture of emerging trends.

We pursued an evaluation approach based on human computation (HC), relieving researchers from the cumbersome and monotonous task of manually annotating data. Concurrently, crowdsourcing also accelerates the annotation task. This is not a contradiction, as HC can access the pooled resources of large user communities. However, to ensure accurate results, the definition and setup of strict guidelines is of utmost importance.

An HC-based evaluation leveraging the CrowdFlower platform underscored the feasibility of crowdsourced evaluation. However, we found out that concepts, when taken out of context, are hard to classify even for human assessors. Accessing CrowdFlower via the uComp API will accelerate future experiments.

References

Ahn, L.v. (2006). “Games with a Purpose”, Computer, 39(6): 92-94.

Buil-Aranda, C., Hogan, A., Umbrich, J., Vandenbussche, P.Y. (2013). sparql Web-querying infrastructure: Ready for action? In: Proceedings of the 12th International Semantic Web Conference Vol.II, pp. 277-293.

Calvanese, D., Cogrel, B., Komla-Ebri, S., Lanti, D., Rezk, M., and Xiao, G. (2015). How to stay ontop of your data: Databases, ontologies and more. In The Semantic Web: ESWC 2015 Satellite Events - ESWC 2015 Satellite Events Portoroz, Slovenia, May 31 - June 4, 2015, Revised Selected Papers, pp. 20–25

Cambria, E.; Schuller, B.; Xia, Y. & Havasi, C. (2013). New Avenues in Opinion Mining and Sentiment Analysis Intelligent Systems, IEEE, 28, 15-21.

Cambria, E., Olsher, D., and Rajagopal, D (2014). SenticNet 3: A common and common-sense knowledge base for cognition-driven sentiment analysis Conference on Artificial Intelligence, pp. 1515-1521.

Cambria, E. (2016). Affective computing and sentiment analysis. IEEE Intelligent Systems 31(2), pp. 102-107.

Christiane Fellbaum (1998) WordNet: An Electronic Lexical Database. Cambridge, MA: MIT Press.

Gindl, S., Weichselbraun, A., and Scharl, A. (2010). Cross-Domain Contextualization of Sentiment Lexicons 19th European Conference on Artificial Intelligence (ECAI),Frontiers in Artificial Intelligence and Applications 215, IOS Press, pp. 771-776.

Harris, S., Seaborne, A. (2013). SPARQL 1.1 query language. W3C Recommendation, March 2013. Accessed: May 2016. Resource is available online at: https://www.w3.org/TR/sparql11-query/

Lehmann, J., Isele, R., Jakob, M., Jentzsch, A., Kontokostas, D., Mendes, P. N., Hellmann, S., Morsey, M., van Kleef, P., Auer, S., and Bizer, C. (2015). DBpedia - A Large-scale, Multilingual Knowledge Base Extracted from Wikipedia. Semantic Web Journal, IOS Press, 6(2): 167-195.

Miller G. (1995). WordNet: A Lexical Database for English. Communications of the ACM Vol. 38, No. 11: 39-41.

Musen, M.A. (2015) The Protégé project: A look back and a look forward. AI Matters. Association of Computing Machinery Specific Interest Group in Artificial Intelligence, 1(4), June 2015.

Sabou, M., Bontcheva, K., Derczynski, L. and Scharl, A. (2014). Corpus Annotation through Crowdsourcing: Towards Best Practice Guidelines. 9th Language Resources and Evaluation Conference (LREC-2014). Reykjavik, Iceland: European Language Resources Association (ELRA), pp. 859-866.

Scharl, A., Sabou, M., Gindl, S., Rafelsberger, W. and Weichselbraun, A. (2012). Leveraging the Wisdom of the Crowds for the Acquisition of Multilingual Language Resources. Proceedings of the Eight International Conference on Language Resources and Evaluation (LREC), pp. 379-383.

Scharl, A., Föls, M. and Herring, D. (2015). "Climate Challenge - Raising Collective Awareness in the Tradition of Games with a Purpose", 14th Brazilian Symposium on Human Factors in Computer Systems (IHC-2015). Salvador, Brazil. 506-509.

Sirin, E., Parsia, B., Grau, B. C., Kalyanpur, A., & Katz, Y. (2007). Pellet: A practical owl-dl reasoner. Web Semantics: science, services and agents on the World Wide Web, 5(2), 51-53.

Socher, R., Perelygin, A, Wu, J., Chuang, J., Manning, C., Ng A., and Potts C. (2013). Recursive Deep Models for Semantic Compositionality Over a Sentiment Treebank. Conference on Empirical Methods in Natural Language Processing (EMNLP).

Speer, R., and Havasi, C. (2013). ConceptNet 5: A Large Semantic Network for Relational Knowledge The People's Web Meets NLP, Springer, Berlin/Heidelberg, pp. 161-176.

Weichselbraun, A., Gindl, S., and Scharl, A. (2013). Extracting and Grounding Contextualized Sentiment Lexicons. IEEE Intelligent Systems, 28, 39-46.

Weichselbraun, A., Gindl, S., and Scharl, A., (2014). Enriching Semantic Knowledge Bases for Opinion Mining in Big Data Applications. Knowledge-Based Systems, 69, 78-85.

Weichselbraun, A., Streiff, D. and Scharl, A. (2015). “Consolidating Heterogeneous Enterprise Data for Named Entity Linking and Web Intelligence”, *International Journal on Artificial Intelligence Tools*, 24(2): 1540008 | 1-31.