Deliverable 5.2

Social Media Analysis Report

Date:	30 September 2016
Authors:	Maria Pontiki (ATHENA R.C.), Dimitris Pappas (ATHENA R.C.), Haris Papageorgiou (ATHENA R.C.)
Contributors	-
Quality Assuror	Anastasakis Leonidas (AST)
Dissemination level:	Public
	WP5
Version:	1.0
Keywords:	Social Media, Data Management, Sentiment Analysis
Description:	This deliverable provides details regarding the CAPSELLA social media analysis platform (the social media channels and the insights extracted for specific communities). The document will be updated on a regular basis reflecting the developments on the platform and the CAPSELLA pilots.

ICT-10-2015 Collective Awareness Platforms for Sustainability and Social Innovation

CAPSELLA (Collective Awareness PlatformS for Environmentally-sound Land management based on data technoLogies and Agrobiodiversity)

Project No. 688813

Project Runtime: January 2016 – June 2018

Version History

Version	Date	Description
0.1	02 Sept. 2016	ToC, First Version of report
0.2	20 Sept. 2016	An updated version of the Report based on the current develop-ments of the project, sent for internal review
0.3	28 Sept. 2016	Final version prepared
1.0	30 Sept. 2016	Final version submitted

Disclaimer

This document contains description of the CAPSELLA project findings, work and products. Certain parts of it might be under partner Intellectual Property Right (IPR) rules so, prior to using its content please contact the consortium head for approval.

In case you believe that this document harms in any way IPR held by you as a person or as a representative of an entity, please do notify us immediately.

The authors of this document have taken any available measure in order for its content to be accurate, consistent and lawful. However, neither the project consortium as a whole nor the individual partners that implicitly or explicitly participated in the creation and publication of this document hold any sort of responsibility that might occur as a result of using its content.

This publication has been produced with the assistance of the European Union. The content of this publication is the sole responsibility of CAPSELLA consortium and can in no way be taken to reflect the views of the European Union.

The European Union is established in accordance with the Treaty on European Union (Maastricht). There are currently 28 Member States of the Union. It is based on the European Communities and the member states cooperation in the fields of Common Foreign and Security Policy and Justice and Home Affairs. The five main institutions of the European Union are the European Parliament, the Council of Ministers, the European Commission, the Court of Justice and the Court of Auditors. (http://europa.eu.int/)

CAPSELLA is a project partially funded by the European Union

Executive Summary

The CAPSELLA project aims to support communities of farmers and food manufacturers in making informed decisions on their activities by offering them access to data from a variety of open data sources related to regional agro-biodiversity and food. A core activity to achieve this goal is the development of Social Data services based on the needs and interests of a community, depending on their respective roles (i.e. consumer, producer, etc.). Social media channels (e.g. Twitter, Facebook, review sites) are widely used today and thus arise as a useful way to search and mine through the available data, creating an implicit flow of information amongst all associated factors. Social Media Analysis[1] is defined as "the process of collecting data from stakeholder conversations on digital media and processing them using data mining techniques in order to extract structured insights leading to more information-driven business decisions and increased customer centrality for brands and businesses". Sentiment Analysis or Opinion Mining is a key technology in the effort to devise such strategies and to help people understand the massive amount of data available online (Pang and Lee, 2008). This report provides an overview of the CAPSELLA Social Data platform and its APIs (section 2), as also as the Social Media Analysis workflows and its services to the CAPSELLA pilots (section 3).

[1] http://www.gartner.com/it-glossary/social-analytics

Table of Contents

Table of Figures

1. Introduction

1.1 Scope

The CAPSELLA project works on making existing data more easily available to local communities of farmers and food producers/retailers through innovative technology-supported solutions that are being developed based on their requirements. An essential task towards this end is to harvest, exploit and incorporate knowledge from all available Social Media channels, in order to facilitate knowledge exchange between the existing communities and to enhance a given knowledge base of a community in general. To achieve this purpose, as a core activity of this task the software that is being developed will perform online Social Media knowledge extraction focusing on opinionated user-generated content as a key information source for understanding individual preferences and content personalization, for reputation monitoring as well as for correlation with Key Performance Indicators (KPIs), among others. The continuous extraction of this knowledge will promote candidate solutions and relevant information and ultimately result in the enhancement of the data of each of the involved communities. This software will also be shared through standard programmable interfaces (APIs) so that other developers may invoke and use it, and will be catalogued through the software services registry of the CIARD RING. To this end, the main scope of this deliverable is to present an overview of the CAPSELLA Social Data platform and its APIs, the Social Media Analysis workflows and its services to the CAPSELLA target audience.

1.2 Audience

This deliverable is a public document that serves as a roadmap for the CAPSELLA Social Media Analysis services targeting not only the CAPSELLA partners, but also other communities that may be interested in the CAPSELLA Social Data methodology, technologies and services.

2. The CAPSELLA Social Data Platform

The CAPSELLA Social Data Platform consists of five major tiers (Figure 1). The Data collection tier where datasets are being captured or ingested from a variety of sources through a set of dedicated tools. Critical considerations here are data formats, schema and metadata management issues that need to be considered and properly resolved. Transporting data from the Collection tier to the CAPSELLA platform is facilitated by a message queuing tier (MQ tier) following the Publish/subscribe interaction pattern. Publish/subscribe messaging is a pattern that is characterized by the sender (publisher) of a piece of data (message) not specifically directing it to a receiver. Instead, the publisher classifies/configures the message somehow, and that receiver (subscriber) subscribes to receive certain classes of messages. Pub/sub systems often have a broker, a central point where messages are published, to facilitate this process. In the CAPSELLA MQ tier, CAPSELLA social analytics workflows (processing tier) will play the role of the receivers enabling the efficient and effective processing of the data producing a wealth of annotations (e.g. sentiments, entities like nutrients, food ingredients, restaurants etc.). In addition, data will be distributed within the platform to provide additional protections against failures, as well as significant opportunities for scaling performance. The results (e.g. annotations) will be transmitted as new messages and further consumed by (a) a distributed indexing & visualization framework providing search and browsing functionalities (Data Exploration tier), (b) archiving to a storage system, and, (c) a persistent storage subsystem providing REST APIs to web apps, third applications and the CAPSELLA pilots (access tier). This section provides a brief description of the data collection tier (2.1), the MQ tier (2.2) and the access tier (2.3).

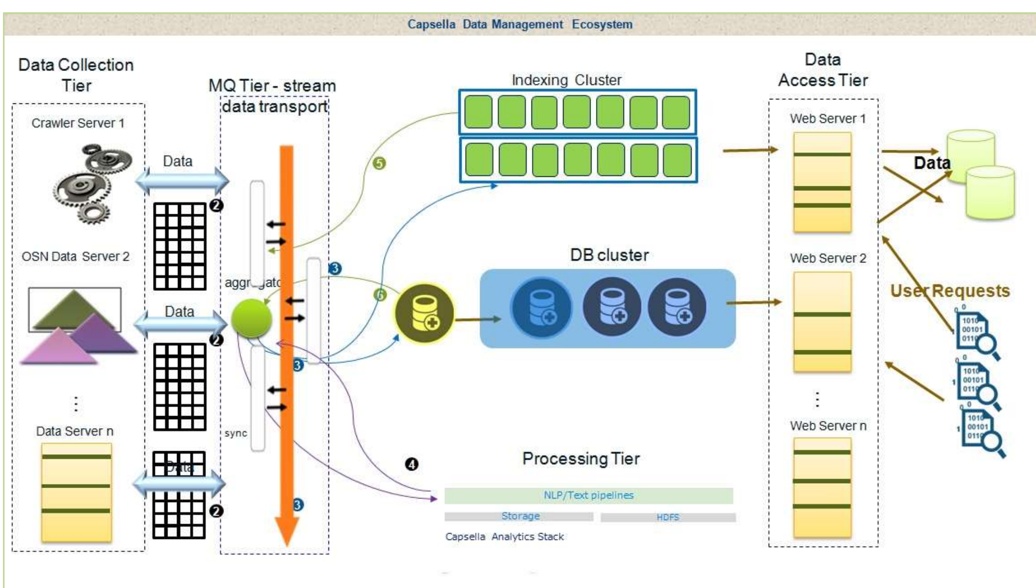

Figure 1: The CAPSELLA Social Data platform

2.1 The Data collection tier

The CAPSELLA platform will be able to ingest data streams from a variety of online social networks (OSNs) (e.g. Youtube, Facebook, Twitter, Instagram etc.) in compliance with the pilot data requirements. To this end, a series of data collectors will be deployed monitoring and collecting relevant data streams from the OSNs in real time on a regular basis. In addition, some datasets are provided as batches (very large datasets) for further processing (e.g. the Yelp Challenge Dataset is provided as a very large downloadable resource and includes information about local businesses in 10 cities across 4 countries. It consists of 2.2M reviews and 591K tips by 552K Yelp users. There are also 566K business attributes and 200,000 pictures available). In the latter case, an offline procedure will be developed enabling storing of chunks of data into the CAPSELLA platform.

The data collected will have to be cleaned and/or transformed before indexing. For example, a webpage has to be cleaned and boilerplate has to be removed (e.g. advertisements, logos, menu buttons, etc). In addition, the actual content of the page needs to be split into fields distinguishing different types of data (e.g. username, comment, location, etc.). A template-driven cleansing process will be developed and configured properly to achieve the above goals.

Since many parts of the platform use or provide data in different formats, a universal format will be adopted catering for all data exchanges. The json format is recommended as it is a broadly used open-standard format that can be easily ingested.

Monitoring the data collectors and providing CAPSELLA operators an easy way to interact with the data collections will be facilitated through proper exploratory (Figure 32) and visualization (Figure 3) tools.

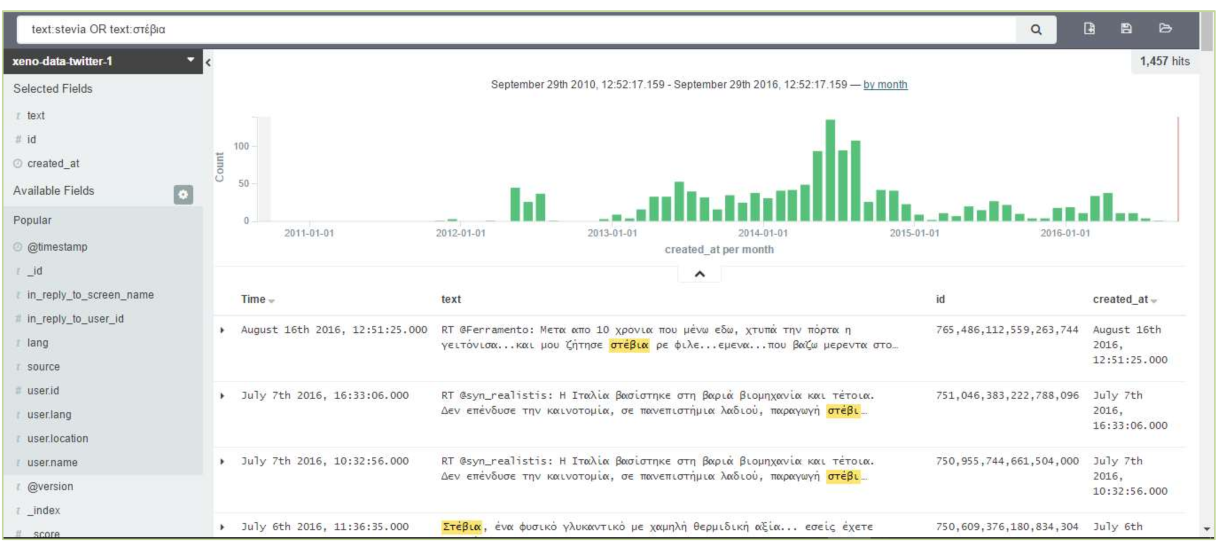

Figure 2: Exploring the twitter collection, retrieving documents using the query “stevia”

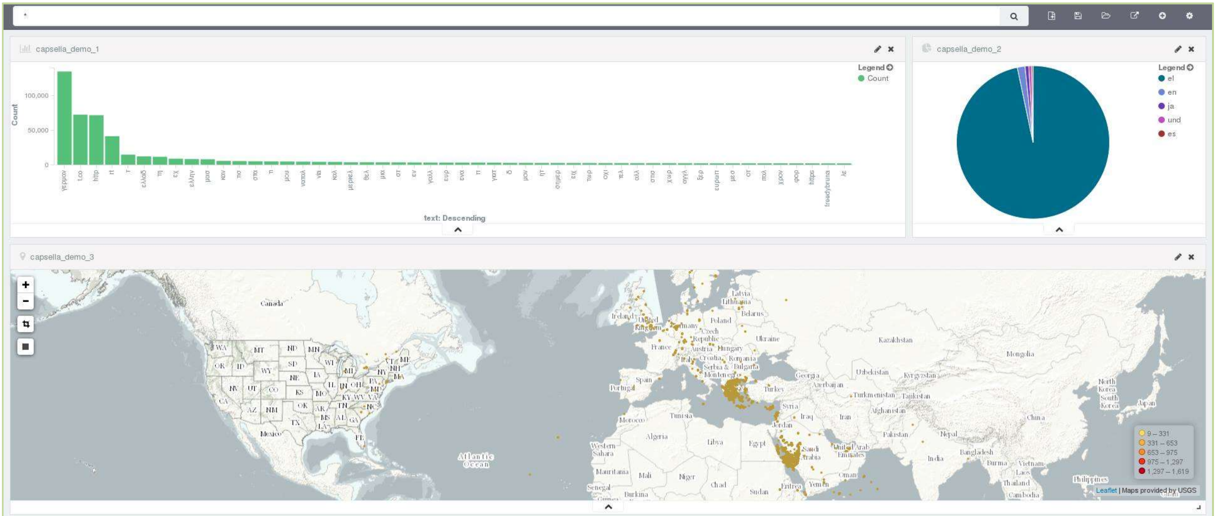

Figure 3: A group of visualizations, presenting the 50 most common terms in the dataset (upper left corner), a pie chart displaying the languages found in the dataset (upper right corner), and a map highlighting the locations presented in the dataset (down center).

2.2 The CAPSELLA message-queuing tier

The backbone of the platform is a distributed streaming tier providing “Publish and Subscribe” services to data streams, storing streams in a distributed replicated cluster and processing datasets in real time. To this end, the open source system, Apache Kafka will be exploited. Apache Kafka is a distributed messaging system and a robust queue that can handle a high volume of data. It can easily distribute messages from multiple sources to multiple destinations ensuring the data integrity as all messages are persisted on the disk and replicated within the cluster. Kafka will facilitate data exchanges, passing data as json messages from one tier of the platform to another. The basic parts of a Kafka system are the Producers, the Brokers and the Consumers. The CAPSELLA operator has to specify a *Topic* which designates a new feed (or category) of data and configure the number of partitions of this *Topic* (see Figure 4). A producer is a program which feeds a specific topic with data. The producer is responsible for choosing which message to assign to which partition within the topic. Kafka runs on clusters of one or more servers each of which is called a broker. Brokers create replicas of the data so the persistence of data is ensured. Finally, a consumer is a program which reads messages written in the partitions of a Topic. The CAPSELLA operator has to designate the actions need to be performed on the data. Such actions could be archiving, feeding the data to a new Topic, and/or feeding them to the processing chain (see section 3).

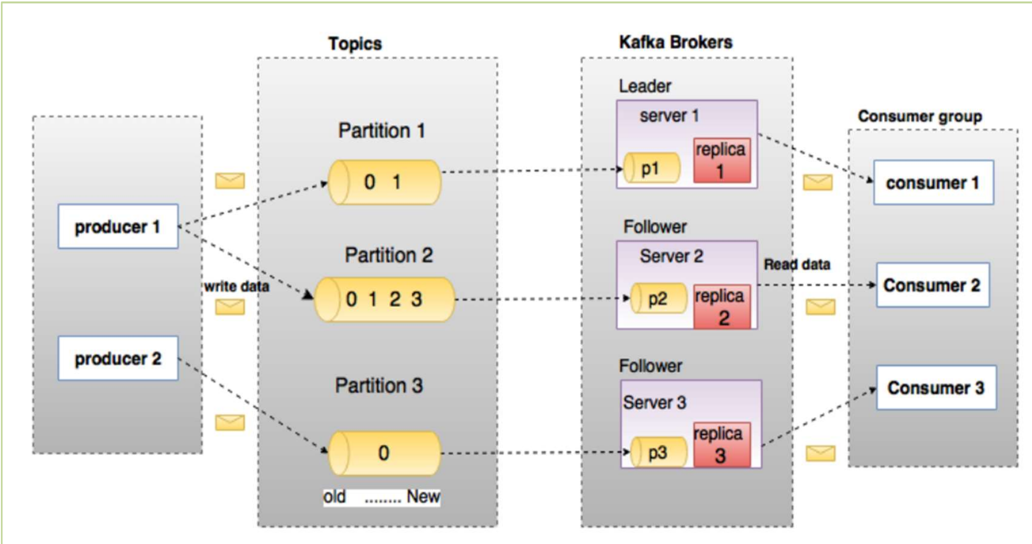

Figure 4: An indicative configuration of the MQ tier with two Producers, a Topic with three partitions, three servers that replicate the data and three Consumers

The MQ tier will be deployed in a cluster of virtual machines in a master slave configuration.

2.3 The CAPSELLA access tier

CAPSELLA provides access to its data and annotations through a managed set of REST APIs. Two task-specific APIs are foreseen. A **Data API** receiving a POST request in json form designating the configuration to be used to export a data collection as shown below.

```
{
    configuration:stevia
}
```

A configuration comprises a set of keywords (or phrases) used to query the CAPSELLA datastore and fetch relevant datasets. Such a configuration is presented below. The pilot user can specify key-words (s)he wants to include in the result set as well as keywords (s)he wants to exclude. Once a new configuration is created, a new data collection is created. This data collection will be fed into the social data processing workflows through the MQ tier and the annotations produced will update the collections.

```
[info]
name:stevia

[include]
phrase1: Sweeteners
phrase2: Stevia
phrase3: Sugar
phrase4: type A diabetes

[exclude]
phrase1: Sweetie
```

In the above example, the user creates a data collection called the “stevia” collection containing any keyword in the set “Sweeteners”, “Stevia”, “Sugar”, “Type A diabetes” but not the keyword “sweetie”. CAPSELLA will create a data collection where it will store all data returned by the specific query.

Subsequently, the Data API detects the corresponding collection and returns the data stored in it. If more than ten results have to be returned, then a method called pagination is used and the results return in batches of ten each. In that case a key is added as json filed and return with the first ten results.


```
{
    configuration:capsella,
    data:[
            {
                "text": "...",
                "contributors": null,
                "@timestamp": "2016-09-19T09:04:32.524Z",
                "entities_trends": [],
                "in_reply_to_status_id": null,
                "entities_urls": [],
                "id": 765486112559263700,
                "favorite_count": 0,
                "user_id": 582212082,
                "retweeted": false,
                "coordinates": null,
                .
                .
                .
            },
            .
            .
            .
    ],
    key: "JYHS#B@as!d"
}
```

In order to get the next ten results the end user has to make a new POST request to the same API providing only the key (s)he received:

```
{
    configuration:stevia,
    key: "JYHS#B@as!d"
}
```

An **Analytics API** returns the results of the social data analysis in the same way as the Data API. It receives a POST request in json form designating the configuration used to export one data collection as shown below. The API detects the corresponding collection and returns the data stored in it containing the annotations extracted from the Social analysis workflows as shown in **Error! Reference source not found.**. The opinions detected in the text are appended to the json data under the field "opinions" and contain information about the opinion target, the aspect category and the polarity of the aspect. Again if more than ten results have to be returned, the results return in batches of ten each in the same manner as in the Data API.

```
[
    {
        tweet_id:"1004293",
        text: "The food was lousy - too sweet or too salty and the portions tiny.",
        opinions: [
                    {
                        target:"food",
                        aspect_category:"FOOD#QUALITY",
                        polarity:"negative"
                    },
                    {
                        target:"portions",
                        aspect_category:"FOOD#STYLE_OPTIONS",
                        polarity:"negative"
                    },
                    .
                    .
                    .
                ]
    },
    .
    .
    .
]
```

3. Social Media Analytics

The first step to build a case in a Social Media Analysis initiative is to set the objectives, viz. to determine who (i.e. an individual, a business, a community) and which goals the data that is collected and analyzed will benefit. In the context of the CAPSELLA project foreseen pilots involve the use of social media data aiming to provide through the CAPSELLA platform social data services to help cooperatives' managers to better market, promote and design their products and services. The pilot cases will be supported by software services and components designed for targeted data collection (i.e. gather data related to specific entities of interest) and data mining with focus on Sentiment Analysis. The continuous extraction of this knowledge will promote candidate solutions and relevant information and ultimately result in the enhancement of the data of each of the involved communities. The analytics will also be shared through standard programmable interfaces (APIs) so that other developers may invoke and use it, and will be catalogued through the software services registry of the CIARD RING.

For example, Agroknow is developing an application for the Stevia Cooperative that is producing the La Mia Stevia product (http://www.lamiastevia.gr/) to access open data that can help them to improve their product. Focusing on Social Media, the main goal is the design and the development of a data API to query and filter the CAPSELLA platform collecting relevant Twitter data. The application comprises of the following two basic steps/services:

- **Data collection.** Collect tweets related to specific entities of interest like the LaMiaStevia product (e.g. Fig. 5) or stevia as a species in general (e.g. Fig. 6).

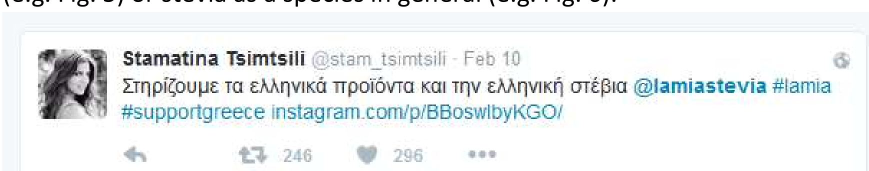

Figure 5: Example of a Tweet discussing the Greek stevia produced by LaMiaStevia

Figure 6: Example of a Tweet discussing stevia as a species in general

Another example of related entities could be competitive products and companies as in the case of imported stevia (e.g. Fig. 7).

Figure 7: Example of a Tweet discussing Paraguay stevia imported in Greece by a competitive company

- **Sentiment Analysis** will be performed on the data collections to extract the sentiment expressed towards the predefined entities of interest. For example, the Tweet in Fig.5 expresses support (i.e. a positive sentiment) for La Mia Stevia and Greek products in general, whilst the Tweet illustrated in Fig. 8 below conveys a positive sentiment about a lemonade made with stevia. The Sentiment Analysis methodology will follow - with the appropriate modifications- the general methodology described below in section 3.1.

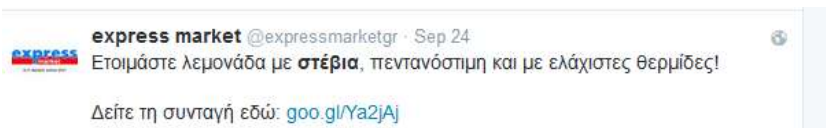

Figure 8: Example of a Tweet conveying a positive sentiment about stevia-sweetened lemonade

3.1 Sentiment Analysis: Approach and Methodology

Sentiment Analysis is defined as the computational study of opinions, sentiments, subjectivity, evaluations, attitudes, appraisal, affects, views, emotions, etc., expressed in text (Liu, 2012). Early work focused mainly on the overall positive or negative classification of a given text or text span (Pang et al., 2002; Turney, 2002). While detecting the overall sentiment of a given text or snippet has a wide range of real-world applications, analyzing unstructured text only in terms of positive and negative opinions -irrespectively of the entities mentioned in context and their aspects- is not sufficient enough to provide meaningful insights and is therefore of limited use. For example, both tweets in Fig. 5 and 8 above convey a positive opinion about stevia; however, in the first tweet the opinion is about the LaMiaStevia product, whilst in the second it is about a stevia-sweetened homemade lemonade. Similarly, restaurants customers reviews not only express the overall sentiment about a restaurant or the food, but also sentiments relating to specific aspects, such as the cleanliness and the prices of the restaurant, and the flavor or the freshness of the ingredients, respectively.

Some review sites (e.g., Amazon, TripAdvisor) provide such information in the form of multiple-aspect user ratings. However, taking into account the textual component of user reviews provides also evidence to understand the reason why behind the rating (Titov and McDonald, 2008; McAuley et al., 2012) and results in better general or personalized review score predictions than those derived from the numerical star ratings given by the users (Ganu et al., 2009). In addition, user ratings are not available in Social Media data like Twitter or Facebook. In this context, research has moved towards fine-grained approaches like aspect-based (or feature-based) sentiment analysis (ABSA) that involves identifying sentiment on different aspects of entities and entities themselves (Zhang and Liu, 2014).

An ABSA method can analyze large amounts of unstructured texts and extract information not included in the user ratings that are available in some review sites. Among several types of representations that have been suggested are tables (e.g., Fig. 9) that list the extracted aspects (e.g., price, quality) of the entity along with the respective ratings (i.e. the average sentiment perceived in reviews), visual representations such as word clouds for the extracted aspects with different colors indicating different sentiments, as well as textual summaries. Each type of representation is best suited for different applications; tables are used for direct comparison between different entities, word clouds for quick access to salient information and text summaries when fluency and coherence, subtle details, personalized and context-aware summaries are important or when visual access to the information is not available, as in the case of speech-based dialogue systems.

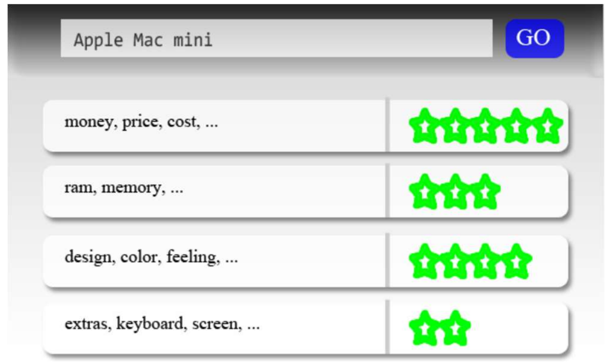

Figure 9: Table summarizing the average sentiment for each aspect of the entity "Apple Mac mini" (Pontiki et al., 2014)

Following this direction, the CAPSELLA Social Data services will integrate ABSA modules designed taking into account the specifications and the requirements of each pilot in order to provide the right type of information to the right community of people. The proposed general ABSA workflow and models are described below in sections 3.1.1 and 3.1.2, respectively. An overview of the state of the art ABSA methods and techniques is provided in section 3.2.

3.1.1 ABSA Workflow

The proposed methodology for building the CAPSELLA ABSA framework is the 4-step process illustrated in Fig. 10.

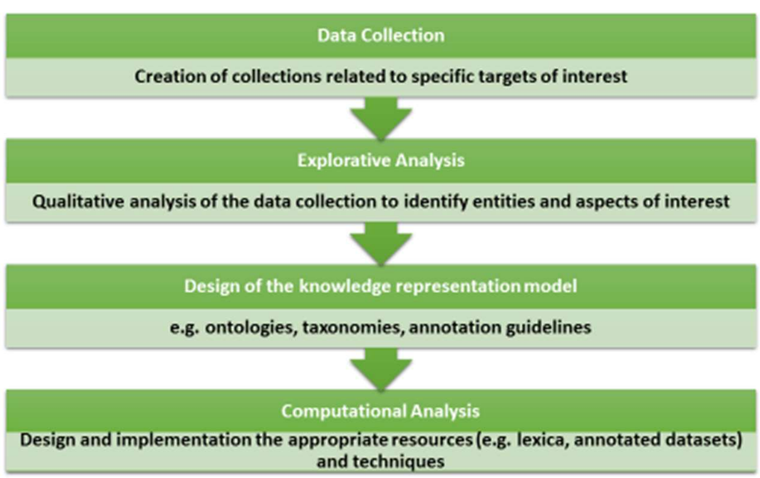

Figure 10: Proposed ABSA Workflow

The first step is to gather data related to specific entities of interest. The methodology can vary depending on the data source and the pilot case. For example, in the case of the stevia pilot we will retrieve from the Twitter data source relevant tweets using specific predefined queries/keywords (e.g. terms like "stevia" or hashtags like "#lamiastevia"). Subsequently, the collected data will be explored by experts in order to determine the different aspects of the entities of interest and their linguistic instantiations. Data exploration is an integral part of the methodology, since it helps to understand and obtain a broader view of the whole dataset and is crucial for filtering the data and clustering them into targeted collections that can be used for development and training purposes.

The third step is to translate the requirements of each pilot case into a knowledge representation schema (i.e. information types to be extracted, output format) taking also into account the findings from the data exploration. The final step is the design and the development of the resources (e.g. lexical resources, linguistic patterns, and annotated datasets) and the models/algorithms needed for the computational treatment of the ABSA representation model.

3.1.2 ABSA Representation Models

The ABSA models will be based on the representation proposed in the context of the SemEval 2015[2] and 2016[3] ABSA shared task (Pontiki et al., 2015; Pontiki et al., 2016) with the proper modifications in order to fulfill the requirements needed for each pilot. As illustrated below in Fig. 11, the ABSA representation for the restaurants domain comprises six entity types (RESTAURANT, SERVICE, AMBIENCE, LOCATION, FOOD, and DRINKS) and five attribute types (*quality, style & options, price, general,* and *miscellaneous*) that are combined to build aspect categories (i.e. entity and attribute pairs, 12 in total). In particular, the entities FOOD and DRINKS can be assigned the attributes "*quality*" (for opinions focusing on the taste, the freshness, the texture, the consistency, the temperature, the preparation, the authenticity, the cooking or general quality of the food and the drinks served in the restaurant), "*style & options*" (for opinions referring to the presentation, the serving style, the portions size, the food/menu options or variety -e.g. innovative dishes/drinks, vegetarian options- of the food and of the drinks served in the restaurant), or "*price*". The entity RESTAURANT can be assigned the attribute labels "*price*" (for opinions that refer to the restaurant prices in general), "*general*" (this attribute label is used in the case of general positive or negative sentiment about an entity type), and "miscellaneous" (for attributes that do not fall into any of the aforementioned cases). The entities types SERVICE, AMBIENCE and LOCATION, can only be assigned the attribute label "*general*" in the context of the SemEval ABSA task.

[2] http://alt.qcri.org/semeval2015/task12/

[3] http://alt.qcri.org/semeval2016/task5/

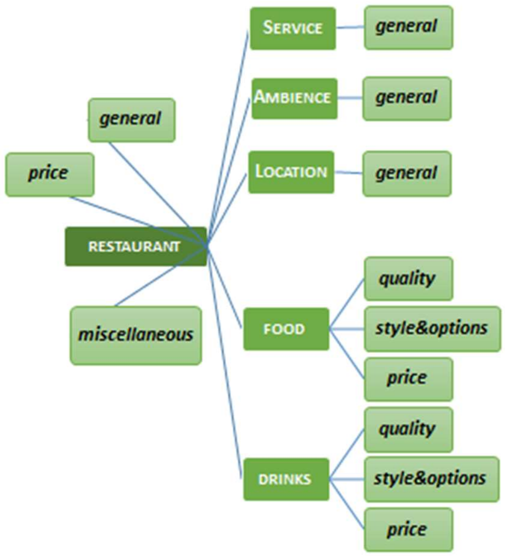

Figure 11: The SemEval 2015 and 2016 ABSA representation for the restaurants domain

Depending on the requirements of each pilot the ABSA representation schema can be simplified or enriched with more entity and attribute types. For example, the entity AMBIENCE could be assigned attributes like "*decoration*" or "*cleanliness*", and the entity SERVICE attributes like the "*staff attitude*" and the "*wait or food preparation time*". Similarly, the existing entity and attribute types can be further analyzed to finer labels. For example, the entity FOOD could be could be further analyzed to specific types of food or dishes, and the attribute "*quality*" to finer aspects like the "*freshness*" and "*taste*". In the case of the stevia pilot the ABSA representation could be transformed as illustrated in Fig. 12.

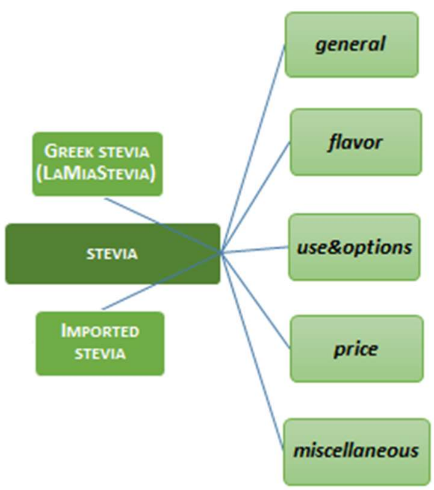

Figure 12: An indicative ABSA representation for the stevia pilot

3.2 ABSA State of the Art

Several ABSA methods have been proposed for various domains including consumer electronics (Hu and Liu {2004a, 2004b}), movies (Thet et al., 2010), restaurants (Ganu et al., 2009), and Beers, Pubs, Toys, Games and Audiobooks (McAuley et al., 2012). Depending on the approach, aspect is a synonym for both fine- and coarse grained types of information. The basic definitions are summarized below:

- Coarse predefined categories (i.e. concept names) similar to rateable aspects (e.g., Ganu et al., 2009; McAuley et al., 2012).
- Aspects are opinion targets i.e. all the targets towards which opinion can be expressed (e.g., Qiu et al., 2011).
- Aspects or features (Hu and Liu, 2014) or facets (Mei et al., 2007) denote components/ parts, subcomponents of the target entity, and attributes of the target entity or its components (Liu, 2010).
- An aspect category is a combination of an entity type E and an attribute type A (E#A pair), where E is the target entity itself, a (sub)component/part of the target entity, or other related entity, while A is a particular attribute of E (Pontiki et al., 2015; Pontiki et al., 2016). E and A are concept names (classes) from a given domain ontology and do not necessarily occur as terms in a sentence. The linguistic expression used in the given text to refer to the reviewed entity E of each E#A pair is the Opinion target expression (OTE).

For example, given the sentence "*The pizza was delicious but do not come here on an empty stomach.*" from a customer review about a particular restaurant, the output of an ABSA method would be as follows for each of the above representations respectively:

- [FOOD: positive & FOOD: negative] or [FOOD: conflict]
- "pizza": positive
- pizza [+5], size [-3] [u][4]
- [FOOD(OTE=pizza)#QUALITY: positive], [FOOD(OTE=null[5])# STYLE_OPTIONS[6]: negative]

The available ABSA methods can be divided into those that adopt domain-independent solutions (Lin and He, 2009), and those that use domain-specific knowledge to improve their results (Thet et al., 2010). Some methods treat aspect extraction and sentiment classification separately (Mei et al., 2007; Brody and Elhadad, 2010), while other approaches model the two problems jointly (Jo and Oh, 2011; Lakkaraju et al., 2014). Common techniques used for ABSA range from frequency (Zhuang et al., 2006) and syntactic relation (Qiu et al., 2011) based approaches to deep learning models (Lakkaraju et al., 2014) and neural networks (Alghunaim et al., 2015). Many approaches treat aspect extraction as a sequential labelling task (e.g. Li et al., 2010) using Conditional Random Fields (CRF)

---

[4] [u] denotes feature/aspect not appeared in the sentence (Hu and Liu, 2014).

[5] According to the SemEval 2015 ABSA schema when there is no explicit mention of the entity E, the OTE slot takes the value "NULL".

[6] According to the SemEval 2015 ABSA schema opinions evaluating the food quantity (e.g. portions size) are assigned the label "FOOD#STYLE_OPTIONS".

(Lafferty et al., 2001) or Hidden Markov Models (HMMs) (Rabiner, 1989). Topic Modelling approaches (Titov and McDonald, {2008a, 2008b}; Brody and Elhadad, 2010; Jo and Oh, 2011) usually use Latent Dirichlet Allocation (LDA) (Blei and Jordan, 2006) to learn distributions of words used to describe each aspect.

Each approach exploits a variety of features to address aspect detection and sentiment classification. The basic types of features used for ABSA are summarized below:

- Lexical features e.g. n-grams, Token shape
- Morpho-Syntactic features e.g. Lemma, Part-Of-Speech (POS), Dependency trees
- Semantic features e.g. Word clusters, Semantic Dependencies
- Lexicon based features e.g. Sentiment Lexica, WordNet
- Word Vector Representations e.g. Word2Vec (Mikolov et al., 2013), GloVe (Pennington et al., 2014)

The ABSA shared task organized for the last three years in the context of the SemEval Workshop provided for the first time the opportunity for direct comparison of ABSA methods. In particular, the task was introduced as a SemEval task in 2014 (Pontiki et al., 2014) providing training data, baselines and a common evaluation framework for both coarse- and fine-grained ABSA. It attracted 165 submissions from 32 teams that experimented with a variety of features (e.g., based on n-grams, parse trees, named entities, word clusters), techniques (e.g., rule-based, supervised and unsupervised learning), and resources (e.g., sentiment lexica, Wikipedia, WordNet). In its second run in 2015 (Pontiki et al., 2015) the task introduced the new ABSA representation attracting 93 submissions from 16 teams. The evaluation metrics used by the task organizers were F-1 measure for aspect category detection and Accuracy for sentiment polarity classification (defined as the number of correctly predicted polarity labels of aspect categories, divided by the total number of aspect categories). As illustrated in Fig. 13, the participating teams explored a variety of features, techniques and resources.

Figure 13: Overview of features, resources and techniques used for ABSA in SemEval-2015 ABSA. The font size indicates frequency.

The best F-1 score in the restaurants domain (62.68%) was achieved by the NLANGP team (Toh and Su, 2015), which modeled aspect category extraction as a multiclass classification problem with features based on n-grams, parsing, and word clusters learnt from Yelp data. A different interesting approach was the system of Sentiue (Saias, 2015) that used a separate MaxEnt classifier with bag-of-word-like features (e.g. words, lemmas) for each entity and for each attribute and was ranked 4th and 2nd in the restaurants and the laptops domains, respectively. The same team achieved also the best accuracy score (79.34%) in the restaurants domain again with a MaxEnt classifier along with features based on n-grams, POS tagging, lemmatization, negation words and publicly available sentiment lexica (MPQA, Bing Liu's lexicon, AFINN). The second best system (Zhang and Lan, 2015) achieved 78.10% in sentiment classification using features based on n-grams, PMI scores, POS tags, parse trees, negation words and scores based on 7 sentiment lexica. The lsislif team (Hamdan et al., 2015) was ranked 3rd (75.50%) and relied on a logistic regression model (Liblinear) with various features: syntactic (e.g., unigrams, negation), semantic (Brown dictionary), sentiment (e.g., MPQA, SentiWordnet).

The task was repeated in 2016 (Pontiki et al., 2016) providing participants the opportunity to improve their results. In addition, in its third run the task went multilingual attracting 245 submissions from 29 teams. The use of the same annotation guidelines for domains addressed in different languages (English, Arabic, Chinese, Dutch, French, Russian, Spanish, and Turkish) enabled also the development of cross-lingual or language-agnostic approaches. The best F-1 score in the restaurants domain was achieved again by the NLANGP team (Toh and Su, 2016) which improved its system with neural network features learnt from a Deep Convolutional Neural Network system. Deep learning models were used also by other teams NileTMRG (Khalil and El-Beltagy, 2016), bunji (Yanase et al., 2016), INSIGHT (Ruder et al., 2016), and UFAL (Tamchyna and Veselovská, 2016) achieving F-1 scores ranging from 72.88% (Khalil and El-Beltagy, 2016) to 59.3% (Tamchyna and Veselovská, 2016). The best accuracy score in the restaurants domain was achieved by the XRCE team (Brun et al., 2016) that combined Machine Learning methods (e.g. CRFs & Elastic Net regression models) with several features derived from robust syntactic/semantic parsing.

Overall, the evaluation of the ABSA systems submitted in the context of the SemEval ABSA shared task (2014, 2015, 2016) indicates that Deep Neural Networks, SVM based algorithms (Wagner et al., 2014; Kiritchenko et al., 2014; Brychcin et al., 2014; Brun et al., 2014) and CRF classifiers (Toh and Wang, 2014; De Clercq et al., 2015; Tho and Su, 2015; Hamdan et al., 2015) achieve state of the art results. Among the most effective features are those derived from sentiment polarity lexica and robust syntactic and semantic parsing. Moreover, in some cases the results indicate that the use of unlabeled data in addition to the provided annotated training data is beneficial.

4. References

Abdulaziz Alghunaim, MitraMohtarami, Scott Cyphers, and Jim Glass. 2015. A Vector Space Approach for Aspect Based Sentiment Analysis. Proceedings of the 1st Workshop on Vector SpaceModeling for Natural Language Processing, pages 116–122, Denver, Colorado, June. Association for Computational Linguistics.

David M. Blei, Andrew Y. Ng and Michael I. Jordan. 2003. Latent Dirichlet allocation. Journal of machine Learning research 3 (Jan), 993-1022.

Samuel Brody and Noemie Elhadad. 2010. An unsupervised aspect-sentiment model for online reviews. In Proceedings of NAACL, pages 804–812, Los Angeles, California.

Tomas Brychcin Michal Konkol, and Josef Steinberger. 2014. UWB: Machine Learning Approach to Aspect-Based Sentiment Analysis, Proceedings of the 8th International Workshop on Semantic Evaluation (SemEval 2014) pp. 817 - 822, Dublin, Ireland, August 23-24, 2014.

Caroline Brun, Diana Nicoleta Popa, and Claude Roux. 2014. XRCE: Hybrid Classification for Aspect-based Sentiment Analysis. Proceedings of the 8th International Workshop on Semantic Evaluation (SemEval 2014) pp. 838 - 842, Dublin, Ireland, August 23-24, 2014.

Caroline Brun, Julien Perez and Claude Roux. 2016. XRCE at SemEval-2016 Task 5: Feedbacked Ensemble Modeling on Syntactico-Semantic Knowledge for Aspect Based Sentiment Analysis. Proceedings of the 10th International Workshop on Semantic Evaluation (SemEval 2016), June 16-17, San Diego, California. ©Association For Computational Linguistics.

Orphée De Clercq, Marjan Van de Kauter, Els Lefever and Véronique Hoste. 2015. LT3: Applying Hybrid Terminology Extraction to Aspect-Based Sentiment Analysis, Proceedings of the 9th International Workshop on Semantic Evaluation (SemEval 2015) pp. 719 - 724, Denver, Colorado, June 4-5, 2015.

Gayatree Ganu, Noemie Elhadad, and Amelie Marian. 2009. Beyond the stars: Improving rating predictions using review text content. In Proceedings of WebDB, Providence, Rhode Island, USA.

Hussam Hamdan, Patrice Bellot and Frederic Bechet. 2015. Lsislif: CRF and Logistic Regression for Opinion Target Extraction and Sentiment Polarity Analysis. Proceedings of the 9th International Workshop on Semantic Evaluation (SemEval 2015), Denver, Colorado, June 4-5, 2015.

Minqing Hu and Bing Liu. 2004a. Mining and summarizing customer reviews. In Proceedings of KDD, pages 168–177, Seattle, WA, USA.

Minqing Hu and Bing Liu. 2004b. Mining opinion features in customer reviews. In Proceedings of AAAI, pages 755–760, San Jose, California.

Yohan Jo, and Alice H Oh. 2011. Aspect and sentiment unification model for online review analysis. In Proceedings of the fourth ACM international conference on Web search and data mining, pages 815-824.

Talaat Khalil and Samhaa R. El-Beltagy. 2016 NileTMRG at SemEval-2016 Task 5: Deep Convolutional Neural Networks for Aspect Category and Sentiment Extraction. Proceedings of the 10th International Workshop on Semantic Evaluation (SemEval 2016), June 16-17, San Diego, California. ©Association For Computational Linguistics.

Svetlana Kiritchenko, Xiaodan Zhu, Colin Cherry, and Saif M. Mohammad. 2014. NRC-Canada-2014: Detecting Aspects and Sentiment in Customer Reviews, Proceedings of the 8th International Workshop on Semantic Evaluation (SemEval 2014) pp. 437 - 442, Dublin, Ireland, August 23-24, 2014.

John Lafferty, Andrew McCallum, and Fernando Pereira. 2001. Conditional random fields: Probabilistic models for segmenting and labeling sequence data. In *ICML-2001*, pages 282–289.

Himabindu Lakkaraju, Richard Socher, and Chris Manning. 2014. Aspect Specific Sentiment Analysis using Hierarchical Deep Learning. NIPS Workshop on Deep Learning and Representation Learning, 2014.

Fangtao Li, Chao Han, Minlie Huang, Xiaoyan Zhu, Ying-Ju Xia, Shu Zhang,and Hao Yu. (2010). Structure-Aware Review Mining and Summarization. International Conference on Computational Linguistics (COLING).

Bing Liu. 2010. Sentiment analysis and subjectivity. Handbook of natural language processing 2, 627-666.

Bing Liu. 2012. Sentiment Analysis and Opinion Mining. Synthesis Lectures on Human Language Technologies. Morgan & Claypool Publishers.

Julian McAuley, Jure Leskovec, and Dan Jurafsky. 2012. Learning attitudes and attributes from multi-aspect reviews. In Proceedings of the 12th IEEE International Conference on Data Mining, ICDM ’12, pages 1020–1025, Brussels, Belgium.

Qiaozhu Mei , Xu Ling , Matthew Wondra , Hang Su , and ChengXiang Zhai. 2007. Topic sentiment mixture: modeling facets and opinions in weblogs, Proceedings of the 16th international conference on World Wide Web, May 08-12, 2007, Banff, Alberta, Canada

Bo Pang, Lillian Lee, and Shivakumar Vaithyanathan. 2002. Thumbs up? sentiment classification using machine learning techniques. In Proceedings of EMNLP, pages 79–86, Philadelphia, Pennsylvania, USA.

Bo Pang and Lillian Lee. 2008. Opinion mining and sentiment analysis. Foundations and Trends in Information Retrieval, 2 (1–2) (2008), pp. 1–135.

Chenghua Lin and Yulan He. 2009. Joint sentiment/topic model for sentiment analysis. In Proceedings of the 18th ACM Conference on Information and Knowledge Management, pages 375–384.

Tomas Mikolov, Ilya Sutskever, Kai Chen, Greg S Corrado, and Jeff Dean. 2013. Distributed representations of words and phrases and their compositionality. In NIPS.

Konstantina Papanikolaou, Haris Papageorgiou, Nikos Papasarantopoulos, Theoni Stathopoulou, and George Papastefanatos. 2016. “Just the Facts” with PALOMAR: Detecting Protest Events in Media Outlets and Twitter. In International AAAI Conference on Web and Social Media. North America.

Jeffery Pennington, Richard Socher, Christopher Manning. 2014. GloVe: Global Vectors for Word Representation. Proceedings of the Conference on Empirical Methods in Natural Language Processing (EMNLP 2014), pages 1532–1543, Doha, QA.

Maria Pontiki, Dimitrios Galanis, John Pavlopoulos, Harris Papageorgiou, Ion Androutsopoulos, and Suresh Manandhar. 2014. Semeval-2014 task 4: Aspect based sentiment analysis. In Proceedings of the 8th International Workshop on Semantic Evaluation (SemEval 2014), pages 27–35, Dublin, Ireland.

Maria Pontiki, Dimitrios Galanis, Harris Papageorgiou, Suresh Manandhar, and Ion Androutsopoulos. 2015. SemEval-2015 Task 12: Aspect Based Sentiment Analysis. In Proceedings of the 9th International Workshop on Semantic Evaluation, Denver, Colorado.

Maria Pontiki, Dimitrios Galanis, Haris Papageorgiou, Ion Androutsopoulos, Suresh Manandhar, Mohammad AL-Smadi, Mahmoud Al-Ayyoub, Yanyan Zhao, Bing Qin, Orphée De Clercq, Véronique Hoste, Marianna Apidianaki, Xavier Tannier, Natalia Loukachevitch, Evgeny Kotelnikov, Nuria Bel, Salud María Jiménez-Zafra, and Gülşen Eryiğit. (2016). SemEval-2016 Task 5: Aspect Based Sentiment Analysis. Proceedings of the 10th International Workshop on Semantic Evaluation (SemEval 2016), June 16-17, San Diego, California. ©Association For Computational Linguistics.

Qui Guang, Liu Bing, Bu, J., Chen, C. 2011. Opinion Word Expansion and Target Extraction through Double Propagation. Computational Linguistics 37 (1).

Lawrence R. Rabiner. 1989. A tutorial on hidden markov models and selected applications in speech recognition. In *Readings in speech recognition*, Alex Waibel and Kai-Fu Lee (Eds.). Morgan Kaufmann Publishers Inc., San Francisco, CA, USA 267-296.

Sebastian Ruder, Parsa Ghaffari and John G. Breslin. 2016. INSIGHT-1 at SemEval-2016 Task 5: Deep Learning for Multilingual Aspect-based Sentiment Analysis. Proceedings of the 10th International Workshop on Semantic Evaluation (SemEval 2016), June 16-17, San Diego, California.

José Saias. 2015. Sentiue: Target and Aspect based Sentiment Analysis in SemEval-2015 Task 12. Proceedings of the 9th International Workshop on Semantic Evaluation (SemEval 2015), Denver, Colorado, June 4-5, 2015.

Aleš Tamchyna and Katerina Veselovská. 2016. UFAL at SemEval-2016 Task 5: Recurrent Neural Networks for Sentence Classification. Proceedings of the 10th International Workshop on Semantic Evaluation (SemEval 2016), June 16-17, San Diego, California. ©Association For Computational Linguistics.

Zhiqiang Toh and Wenting Wang. 2014. DLIREC: Aspect Term Extraction and Term Polarity Classification System. Proceedings of the 8th International Workshop on Semantic Evaluation (SemEval 2014) pp. 235 - 240, Dublin, Ireland, August 23-24, 2014.

Zhiqiang Toh and Jian Su. 2015. NLANGP: Supervised Machine Learning System for Aspect Category Classification and Opinion Target Extraction. Proceedings of the 9th International Workshop on Semantic Evaluation (SemEval 2015) pp. 719 - 724, Denver, Colorado, June 4-5, 2015.

Zhiqiang Toh and Jian Su. 2016. NLANGP at SemEval-2016 Task 5: Improving Aspect Based Sentiment Analysis using Neural Network Features. Proceedings of the 10th International Workshop on Semantic Evaluation (SemEval 2016), June 16-17, San Diego, California.

Tun Thura Thet, Jin-Cheon Na, and Christopher S. G. Khoo. 2010. Aspect-based sentiment analysis of movie reviews on discussion boards. J. Information Science, 36(6):823–848.

Ivan Titov and Ryan T. McDonald. 2008a. A joint model of text and aspect ratings for sentiment summarization. In Proceedings of ACL, pages 308–316, Columbus, Ohio, USA.

Ivan Titov and Ryan T. McDonald. 2008b. Modeling online reviews with multi-grain topic models. In WWW, pages 111–120, New York, NY, USA. ACM.

Peter Turney. 2002. Thumbs up or thumbs down? Semantic orientation applied to unsupervised classification of reviews. In Proceedings of ACL, pages 417–424, Philadelphia, Pennsylvania, USA.

Joachim Wagner, Piyush Arora, Santiago Cortes, Utsab Barman. Dasha Bogdanova, Jennifer Foster and Lamia Tounsi. 2014. DCU: Aspect-based Polarity Classification for SemEval Task 4. Proceedings of the 8th International Workshop on Semantic Evaluation (SemEval 2014) pp. 223 - 229, Dublin, Ireland, August23-24, 2014.

Toshihiko Yanase, Kohsuke Yanai, Misa Sato, Toshinori Miyoshi and Yoshiki Niwa. 2016. bunji at SemEval-2016 Task 5: Neural and Syntactic Models of Entity-Attribute Relationship for Aspectbased Sentiment Analysis. Proceedings of the 10th International Workshop on Semantic Evaluation (SemEval 2016), June 16-17, San Diego, California. ©Association For Computational Linguistics.

Lei Zhang and Bing Liu. 2014. Aspect and Entity Extraction for Opinion Mining", book chapter in Data Mining and Knowledge Discovery for Big Data: Methodologies, Challenges, and Opportunities, Springer, 2014.

Zhihua Zhang and Man Lan. 2015. ECNU: Extracting Effective Features from Multiple Sequential Sentences for Target-dependent Sentiment Analysis in Reviews. Proceedings of the 9th International Workshop on Semantic Evaluation (SemEval 2015), Denver, Colorado, June 4-5, 2015.

Li Zhuang, Feng Jing, Xiao yan Zhu, and Lei Zhang. 2006. Movie review mining and summarization. Proceedings of ACM International Conference on Information and Knowledge Management (CIKM-2006), 2006.