EXPLORING THE INTEGRATION OF SOCIAL MEDIA FEEDBACK FOR USER-ORIENTED PRODUCT DEVELOPMENT

Deng, Quan (1,2); Franke, Marco (1,2); Hribernik, Karl (1,2); Thoben, Klaus-Dieter (1,2)
1: BIBA - Bremer Institut für Produktion und Logistik GmbH, Germany; 2: University of Bremen, Germany

Abstract

Product designers thrive on designing products to fulfil various expectations and needs from customers. To understand the customer expectation and needs, it is crucial to have the information on customer feedback that is generated during product usage phase. For this purpose, social media has attracted strong interest, as increasing amount of information is published daily by customers. This information is related to a wide range of products and contains product specific feedbacks. To make use of the feedbacks, different approaches were developed and described in literature. Most of them focused on the extraction of limited information to support specific tasks, which is however not flexible and general enough. Little research has provided a practicable and flexible solution to support different design tasks in various domains. This article suggests a social media wrapper approach, which can be flexibly configured to address this issue. It provides designers a holistic view of the feedbacks that widely distributed in different social media channels as well as in diversity data sources. This holistic view of feedbacks can be analyzed to earn necessary knowledge for design tasks.

Keywords: Product Lifecycle Management (PLM), User centred design, Crowdsourcing, Social media, Semantic interoperability

Contact:
Quan Deng
University of Bremen
BIBA - Bremer Institut für Produktion und Logistik GmbH, IKAP
Germany
dqu@biba.uni-bremen.de

Please cite this paper as:
Surnames, Initials: *Title of paper*. In: Proceedings of the 21st International Conference on Engineering Design (ICED17), Vol. 4: Design Methods and Tools, Vancouver, Canada, 21.-25.08.2017.

1 INTRODUCTION

Companies are constantly challenged to meet different customer expectations to achieve high acceptance of their product on the market. In order to facilitate broad acceptance of the product, product design needs to be continually improved on the basis of user feedback about what the customer really experience on the product, and how the product is performed along the lifecycle. As highlighted in various studies (e.g., Chang and Chen, 2014), customers have attracted much attention in product design and innovation process. Different amounts of input information provided by customers have high impacts on product innovation results (Chang and Chen, 2014). It appears that feedback information from customers are crucial for product design and innovation process. Integrated customer feedback information from different channels is therefore important for product design, so that these kind of information can be used and analysed to support product design tasks.

Feedback information from customers can be collected in different ways. Traditionally, it can be gathered through market survey and so on. But due to cost and time, the number of feedback is limited. Nowadays, social media has becoming an effective medium in which consumers can publish information on the Internet. It contains product specific feedbacks including usage scenarios, technological hurdles from the perspective of the daily usage, best practices or the quality etc. Thus, the data and the knowledge contained in social media has a high value for companies and could be used to improve product design. The power of social media analysis has begun to been explored by an increasing number of researchers to create business values (Fan and Gordon, 2014).

However, it lacks sufficient investigations on how to efficiently extract and integrate customer feedback information from various social media channels to support various product design tasks. Some attempts have been made to make use of social media texts to discovery specific kinds of information, such as latent product features, latent customer needs or customer opinions towards product features in order to improve product design (Tuarob and Tucker, 2015; Yagci and Das, 2015; Zhou et al., 2015). These approaches may work quite well to extract limited information to support their defined tasks. However, it is neither flexible nor general enough to extract and integrate feedback information from various social media channels for the problem solving in different tasks within various application domains according to the information needs of the designers. This paper proposed a social media wrapper approach, which makes advantage of ontology and natural language processing (NLP) to extract and integrate necessary structured feedback information to flexibly support different product design tasks in different domains based on the designers' information needs. It provides product designers with a holistic view of structured feedback information, which semantically integrates the feedback information that widely distributed different social media channels as well as feedbacks in other diversity data sources e.g. database. The integrated structured feedback information can be further analysed to earn necessary knowledge for different design tasks. The data analysis can be based on techniques such as statistic or data mining, which is however a research topic in its own right, and will not be detailed in this paper.

The rest of the paper is organized as follows. Section 2 describes the relevant literature available on the topic, while section 3 describes the proposed approach. The next section provides an example of the usage of the proposed approach. After that, a short discussion is given is the next section. Finally, the last section draws a conclusion and states current limitations and future works.

2 RELATED WORK

This section covers the previous work concerned with the usage of social media information, as well as the use of social media analysis for product design. The first part of this section summarizes the current understanding of the social media information with respect to Product Lifecycle Management (PLM) and product design. In the second part, existing technologies for social media analysis is briefly introduced, with a focus on ontology based information extraction and opinion mining. In the last part, existing works and their drawbacks are briefly presented with respect to the use of social media analysis for product design.

2.1 Social Media Information for Closed-Loop PLM

PLM is the process of managing the product related information during the entire product lifecycle, from the phase Begin of Life (BOL), through Middle of Life (MOL) to End of Life (EOL) (Kiritsis, 2011).

Kiritsis (2011) argues that there is a lack of information exchange and information feedback between different stages because the information flows are in most case interrupted shortly after product sale, and has thereby proposed the concept of closed-loop PLM that enable closed product lifecycle information loops. With respect to the improvement of product design, the information and knowledge from MOL phase, for example, the information collected by product embedded information devices (PEID) or from social media, can be handled as effective feedback.

Social media servers an effective medium in which consumers can publish their daily life information on the Internet. This increasing amount of information includes consumer feedback about products. This kind of information in the product usage phase is becoming a high-value information source for companies. The current available 'blogosphere' is more than 100 million blogs and their interconnections have become an important source of public opinion (Kietzmann et al., 2011). In the application of micro-blogging, the leading company Twitter has achieved more than 145 million users who send more 90 million 'tweets' per day, each consisting of 140 characters or less (Madway, 2010).

Consequently, a huge amount of information is available and is related to a wide range of consumer products and corresponding services. The data contains product specific feedbacks including usage scenarios, technological hurdles from the perspective of the daily usage, best practices or the quality of product-specific services like maintenance. Thus, the data and the knowledge contained has a high value for manufacturers and could be used to improve not only the product design but also the overall product service system. To achieve such an impact, the relevant knowledge must be found, extracted and aggregated from social media according to the information needs of the designers.

2.2 Technologies for Social Media Analysis

In the past few years, social media analysis has attracted strong interest from researchers, an increasing number of researchers have begun to explore the power of social media to create business values (Fan and Gordon, 2014). It involves a variety of techniques from different fields, including NLP, sentiment analysis, topic modelling and so on. Here, we highlight the technologies related to acquiring product related consumer feedback from social media. The principle idea behind acquiring product related consumer feedback from social media is to automatically extract specific information and then transform them into a higher degree of formality. The extraction and transformation can be realised through methods of NLP. In respect to the degree of information formality, the higher the degree of formality, the more downstream analysis methods are applicable. To ensure a high degree of formality and semantic interoperability, ontology is usually taken as a target representation form.

Information extraction (IE) is in principle a subfield of NLP. The general process of IE is based on automatically retrieving certain types of information from natural language text (Wimalasuriya and Dou, 2010). According to Riloff, IE is a form of NLP in which certain types of information must be recognized and extracted from texts (Riloff, 1999). While the recently emerged term ontology based information extraction (OBIE) is once again a subfield of IE, it is described as a system that processes unstructured or semi-structured text through a mechanism guided by ontologies to extract certain information and presents the output using ontologies (Wimalasuriya and Dou, 2010). According to the definitions, the main differences between OBIE and IE is the former's linkage to semantic ontology.

Researches in NLP and in particular in the subfield of IE have adopted various approaches including machine learning approaches like Latent Dirichlet Allocation (LDA) (Blei et al., 2003) and syntactic as well as rule/pattern-based approaches to build systems that extract certain information from natural language documents (Muslea, 1999; Grishman, 1997; Cowie and Lehnert, 1996). These approaches often rely on NLP techniques such as Tokenizer, Sentence Splitter and Part of Speech (POS) Tagger for understanding the linguistic structures of sentences. Corresponding NLP tools are: GATE (https://gate.ac.uk/), UIMA (https://uima.apache.org/), RapidMiner (https://rapidminer.com), etc.

While the focus of OBIE is the extraction of factual information from text, opinion mining known as sentiment analysis focuses on the extraction of subjective information. Here, methodologies are employed to meet the demands for detecting opinions and sentiments. An early extensive survey on opinion mining and sentiment analysis is performed by Pang and Lee, covering techniques that promise to directly enable opinion-oriented information-seeking systems (Pang and Lee, 2008). Liu presented another more recently survey of all important research topics in this field, in this survey a formal definition of the objective of sentiment analysis is defined (Liu, 2012). Following the definition, the necessary tasks as well as related techniques to be performed to achieve the objective are mentioned. Medhat et al. (2014), Schouten and Frasincar (2016) have given short surveys illustrating the new trends in opinion mining. For the sentiment analysis in different granularity level, various sentiment classification methodologies are preferred. In general, two main research directions, i.e. machine learning approaches and lexicon based approaches, have been focused on in recent years.

Both OBIE and opinion mining are highly active research fields. Considerable research has been conducted to take advantage of ontology for feature-based opinion mining (Zhao and Li, 2009; Peñalver-Martinez et al., 2014; Salas-Zarate et al., 2016). Ontology provides a formal structure knowledge representation for a domain. It helps the identification of product features.

2.3 Social Media Analysis for Product Design

For the improvement of product design, social media texts have been utilized and analysed from various perspectives. Popescu and Etzioni (2005) has introduced an information extraction system to extracts product features and associated opinions from product reviews. Beside the extraction of product features and associated opinions, Yagci and Das (2015) have presented the design-feature-opinion-cause relationship (DFOC) method, which can also identify the likely design cause of the customer opinion from unstructured web reviews to support product design manager. Tuarob and Tucker (2015) has mined social media networks to identify lead users and latent product features relating to specific products. Zhou et al. (2015) have extracted latent customer needs from online product reviews through semantic analysis and use case analogical reasoning. Palmer (2016) has introduced a process for identifying customer needs for product design through text analysis of online customer reviews with the analysis methods such as co-occurrence network. However, most of the researches are focusing on specific tasks, either on finding feature-opinion-cause relationships or on discovering latent customer needs etc. Few research has provided a domain- and task-independent approach, which can on one side extract flexibly ontology-defined factual and sentiment information from different social media channels according to the information needs of the individual tasks, and on the other side also enable semantic integration of diversity data sources to provide designers a holistic view of integrated structured feedback information for the purpose of tasks support.

3 SOCIAL MEDIA WRAPPER APPROACH

The research methods and approaches discussed in the above section, such as ontology based information extraction, lay the basis for our contribution. The approach proposed in this section will go further to combine semantic interoperability to enable the integration of the social media feedbacks with the information in other data sources with the aim to support various design tasks. It will meet the aforementioned requirements: domain- and task- independent, semantic interoperability.

3.1 Overview

The system architecture of the social media wrapper approach is shown in Figure 1. It follows three main steps to transform social media texts into specific domain ontology: 1) data acquisition 2) text pre-processing 3) semantic transformation. In the data acquisition phase, social media related data sources such as Facebook which are related to given design tasks are accessed. Afterwards, in the text pre-processing module, common NLP text pre-processing methods e.g. tokenizing, POS Tagging are applied on the unstructured texts to provide a basis for further text analysis. In the last step, different pieces of information including both factual information and subjective information are extracted on the basis of the given ontology. The extracted pieces of information are further associated together to populate ontology ABoxes and finally integrate into product design knowledge base which is implemented as Triple-Store. Afterwards, the product design related information extracted from social media can be additionally integrated with that from other data sources with the help of the given ontology and the Mediator for Product Design Data Federation module.

To meet the demand of domain- and task- independent, the approach is highly configurable. During the data acquisition phase, users are able to specify a list of selected social media sources to guarantee that the collected texts are highly relevant for the upcoming design tasks. Subsequently, the texts are pre-processed and semantic transformed. The semantic transformation process can be flexibly guided by specifying an ontology and a configuration file which specifies how to extract semantic information from social media and mapped them to the given ontology. In different design tasks, it is not unusual that a single text is handled from different viewpoints. This is also supported by the proposed approach, as on a single pre-processed text multiple semantic transformation processes are allowed. Each semantic

transformation process extracts then information with a defined viewpoint to meet the needs of specific design tasks.
To achieve semantic interoperability, ontology has been proven to be a successful way to combine data or information from multiple heterogeneous sources (Obrst, 2003). In our approach, ontology is used to direct the semantic transformation process to extract meaningful information from different social media sources. By doing so, information from different social media sources such as Facebook, Pinterest and Lookbook are able to be semantically integrated. The ontology is thereafter established as a mediation support for data integration of other heterogeneous data sources in order to provide product designers a holistic view of feedback information.
The detailed descriptions of the approach are presented in the following sections. The data acquisition is briefly discussed first. Then we describe shortly the text pre-processing. In the last sub-section, the techniques for text semantic transformation are presented.

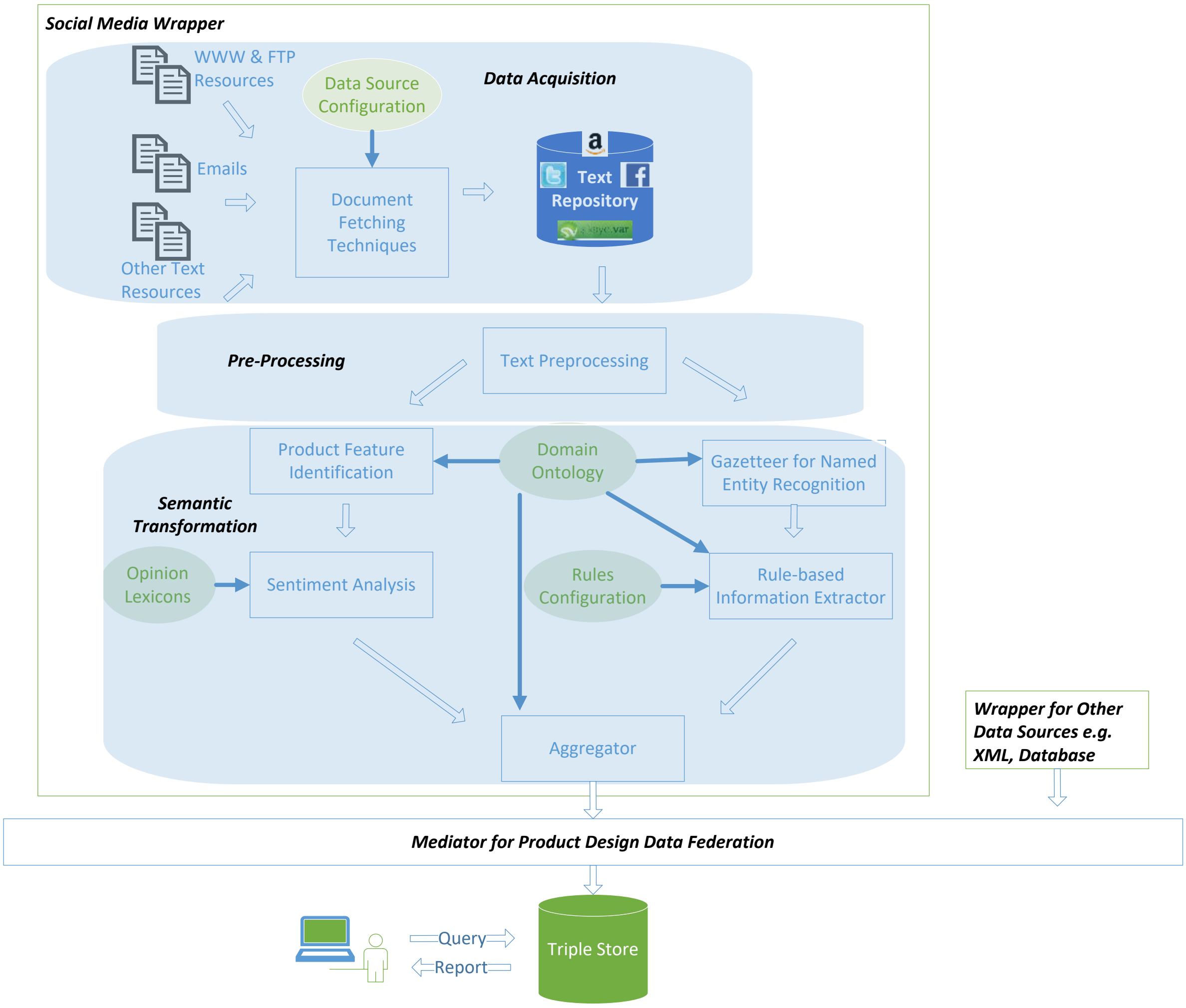

Figure 1. Social Media Wrapper Architecture

3.2 Data Acquisition

The addressed social media related data sources are categorised into three groups according to the applicable data access technologies: HTML Crawling, Web Service Invocation, File Repository Access. The first group contains websites, which offers no public API. The data acquisition is based on HTTPGET requests and downstream interpretation of the HTML content. In so doing, the data acquisition will follow included links to a specified configured depth to collect more detailed information about products. The social media website like Lookbook and Sikayetvar are for example belonged to the first group. The second group, which include popular social medias such as Facebook, offers web services for third party tools. Thus, the data acquisition is going to implement HTTPGET and HTTPPOST requests and a downstream interpretation of JSON and XML content. The remaining group provides data access via remote file repositories. The access to these repositories will be established by FTP, SMB or other necessary protocols via pre-existing clients. For that purpose, open source java libraries will be applied to support the necessary protocols. The files included in the remote repositories will be handled as text files.

The above-mentioned data acquisition methods will be capable of collecting continuous text from web resources as well as from remote files repositories. Texts from different data sources will be handled as natural language and forwarded to the pre-processing module.

3.3 Text Pre-Processing

The pre-processioning of social media text includes the common pre-processing steps in NLP, such as language detection, tokenization, stemming and lemmatization, POS tagging. As many of the pre-processing steps are similar with that in other NLP approaches, only parts of the most important steps are presented in this section. The pre-processing process is triggered by the step language detection. Depends on the language detection result, corresponding language resources would be applied for the pre-processing of the given social media text. During tokenization, the social media texts are segmented into a sequence of tokens which could be a word, punctuation, numbers, or other meaningful elements called token. In the step of POS tagging, each word form is associated with its corresponding particular part of speech to allow preliminary recognition of the phrasal units.

3.4 Semantic Transformation

The approach presented in this section can be seen as information transformation chain to transform pre-processed unstructured texts in business domains to structured semantic annotated triple entries. The task of this component is defined as following: Given an ontology TBox and a pre-processed unstructured text, Ontology ABoxes are supposed to be automatically populated using the information extracted from the pre-processed unstructured text. In this approach, both factual and sentiment information on product features will be extracted based on the aforementioned NLP tool: GATE.

3.4.1 Extracting Factual Information

Extracting factual information consists of the extraction pieces of factual information from unstructured texts and linking them with the ontology concepts and properties. In order to link specific text chunks to ontology concepts/properties, various information extraction techniques could be applied. In our approach, the approach gazetteer based named entity recognition and linguistic rules techniques are mainly investigated.

The approach gazetteer based named entity recognition involves the identifying individual entities for a particular ontology concept. The gazetteer lists are used to find occurrences of entities in text. It is clear that in this approach, the quality of the named entity recognition is highly depended on the gazetteer resource. In order to identify a concrete ontology concept, this technique requires a carefully prepared gazetteer list containing all instances of the to be identified concepts. It would be not applicable when the gazetteer resource is not available. The information for the gazetteer list can either be obtained with the help of domain ontology or from other resources.

The linguistic rules technique concerns the extraction of specific information from text using linguistic rule matching technique. The rules are in general based on regular expression combining with the information from Tokenizer, POS tags and Gazetteer etc. Obviously, there is no general rule to recognize information for all concepts from different social medias, individual rules are necessary for specific concepts. Usually, suitable extraction rules have to be manually configured based on the learning of sample documents. Extraction rules can be written in the language Java Annotation Patterns Engine (JAPE), which is a finite state transductor allowing the creation of complex extraction rules over annotations.

3.4.2 Extracting Feature-based Sentiment Information

This process is for identifying product features and respective sentiment with the help of domain ontology. It consists of two sub-modules: product feature identification and sentiment analysis.

The main idea behind product feature identification is to use formal define ontology terminology for the identifying of product features. It takes the pre-processed text and domain ontology as input to recognize the product features that are present in the text.

The module sentiment analysis handles the identification of the sentiment polarity towards target product features. In this approach, sentiment analysis will be done using generic lexicons and set of basic opinion rules. To identify the words which express opinions towards a target feature, the same strategy presented by Martinez et al. (Peñalver-Martinez et al., 2014) can be employed: 'N_GRAM After', 'N_GRAM Before', 'N_GRAM Around' and 'All_Phrase'. The sentiment polarity of the target feature is then calculated based on the sentiment values of each opinion word derived from lexical resources. For example, SentiWordNet can be used as a lexical resource for English texts to get sentiment scores (i.e. positive, negative, objectivity) of each WordNet synset. Meanwhile, basic opinion rules including the negation words such as "not" are specially handled, because this kind of words would usually change or even reverse the sentiment orientations.

4 USE CASE

In this section, we provide an example of the usage of the proposed approach in an application scenario. The application scenario regards a company focused on the design and production of garments in the textile industry. In the process of garments design, it is critical to know the fashion trend of the target groups in the target area. With correct information about trends, the designer can decide about the type of the garments, and the styles etc. with higher confidence. In order to obtain such kind of information, social medias have gained a crucial role because of the availability of the huge amount of related garment information. One example task in the design process is to decide "Which colour of specific garment category (dress, etc.) is trendy in a specific set of countries?".

To answer this question, the following information along with their connections in between need to be extracted from social media texts: Garment, Color, Garment Category, Country, Vote. Several social media sources are analysed to identify whether the texts in the social media source could supply qualified information to answer the question. Example social media sources in this application scenario are Facebook and Lookbook.

In the data acquisition phase, the texts from Facebook and Lookbook are configured to be downloaded to the local text repository. A sample snippet of the data configuration file can be found in Figure 2. The configuration items for both social media sources are not totally the same, because in the backend different approaches are applied for the text collection. For Facebook, the provided REST web service API is used to get the needed texts, while for Lookbook, a Web Crawler is applied because of the lack of necessary web service APIs. The configuration item "datawrapper_config_property_file" specifies a configuration file, which will be used to guide the further processing of the loaded social media texts.

```
<lookbook>
    <lb_dataloaders>
        <lb_dataloader pageurl="http://lookbook.nu/"/>
    </lb_dataloaders>
    <datawrapper_config_property_file>/Resources/Lookbook/config.properties</datawrapper_config_property_file>
</lookbook>

<facebook>
    <fb_dataloaders>
        <fb_dataloader pageid="FSM.QUE" postcount="10000" loadAllComments="false"/>
        <fb_dataloader pageid="asian.fashion" postcount="10000" loadAllComments="false"/>
    </fb_dataloaders>
    <datawrapper_config_property_file>/Resources/Facebook/config.properties</datawrapper_config_property_file>
</facebook>
```

Figure 2. Example configuration for data acquisition

In case of needed social media texts are collected, the texts go then through the steps of pre-processing including tokenization, POS tagging etc. and finally come to the semantic transformation process. At the semantic transformation phase, a configuration file as specified in the above configuration item "datawrapper_config_property_file" is employed to direct the transformation process. In the configuration file, a formal garment ontology and a semantic mapping file are included. The formal garment ontology has defined the important related classes such as Social Media, Garment, Color and Garment Category as well as their relationships in between. The semantic mapping file specifies the information extraction techniques for the establishment of semantic mapping from social media text to the formal garment ontology. A sample snippet of the semantic mapping file is shown in Figure 3. It gives the rules to map the following garment ontology element: class "Color", class "Garment", and object property "describesProduct" that describes the relationship between "Color" and "Garment". The sample semantic mapping has mainly applied the aforementioned techniques: gazetteer based named entity recognition and JAPE based linguistic rules. One example is the mapping of the ontology class "Garment". For the mapping, the gazetteer list "category.lst" has listed all possible garment categories, once the listed category is detected in the social media text, the respective text chunk will then be linked to ontology class "Garment" with the help of linguistic rules specified in the JAPE file "clothes.jape". In a similar way, the other necessary information such Garment Category, Country and Vote can be extracted from social media texts.

```
<ClassMapping className="Garment">
    <ID>2</ID>
    <Detection>
        <Gazetter major="Clothes" minor="">category.lst</Gazetter>
        <JAPE>Clothes.jape</JAPE>
    </Detection>
</ClassMapping>

<ClassMapping className="Color">
    <ID>5</ID>
    <Detection>
        <Gazetter major="Color" minor="">color.lst</Gazetter>
        <JAPE>Color.jape</JAPE>
    </Detection>
</ClassMapping>

<!--object property of the class Color-->
<ObjectPropertyMapping propertyName ="describesProduct">
    <SubjectID>5</SubjectID>
    <SubClass>Color</SubClass>
    <ObjectID>2</ObjectID>
    <ObClass>Garment</ObClass>
    <Detection>
        <JAPE>Mentions_Garment.jape</JAPE>
    </Detection>
</ObjectPropertyMapping>
```

Figure 3. Example mapping between unstructured text and ontology element

At the end, we can get a knowledge base with a list of feedback information extracted from social media text, as shown in Figure 4. To support designers to answer the aforementioned question "which colour of specific garment category (dress, etc.) is trendy in a specific set of countries?" with the extracted feedback information, a very simple example statistical analysis can be carried out: at the first step, the designers select all feedback records related to the garment category "Jacket" in the country "Poland"; after that, the designers group the records by color to get the color related statistic, for example, the designer may get the result that 70% of the records are with the color "black"; finally, the designers can get an answer for the question from the statistics, for example, "black" is a trendy color for the "Jacket" in "poland". When the designers want to have a more precise and unbiased answer, the following two options should be considered: 1) take more factors such as "vote" into analysis 2) have a holistic and unbiased feedback knowledge base with the help of semantic integration based on the given garment ontology and mediator for product design data federation module. To be integrated information can either from other social media channels or from other data sources e.g. historical sales data. For the support of other tasks, the designers can conduct other analyses on the knowledge base to earn the required answer when the needed information are already extracted and integrated. In other cases, the needed information for the individual tasks should be first extracted and integrated with the proposed social media wrapper approach.

Garment	Color	Garment_category	Country	Vote	SocialMediaSource
garment:1481130856560	garment:black	garment:Jacket	garment:poland	"29"	"Lookbook"

Figure 4. Example information extracted from social media

5 DISCUSSION

This paper stated that the current approaches to extract and integrate customer feedback information from various social media channels for product design tasks support is insufficient, and has taken a step forward by proposing a ontology-based social media wrapper approach. This approach allows flexible extraction of feedback information from social media channels based on the demands of the specific product design tasks in various domains. This flexibility is threefold: firstly, users can freely choose the social media channels, which are most relevant to their tasks; secondly, users can determine which information elements should be extracted for the task support through the ontology specification according to the information needs of the individual tasks; thirdly, users can interpret a piece of social media text from their own perspectives to get and correlate the needed information elements. With the help of ontology, the extracted information elements are not independent from each other but correlated together. In addition, ontology based approach can guarantee the semantic integration and interoperability of the feedback information from different channels, not only between different social media channels, but also between social media channel and other data sources, so that a relative unbiased feedback knowledge base can be constructed.

Many of recent works on the utility of social media texts to support product design have focused on the identification of product features and associated opinions (e.g. Popescu and Etzioni, 2005; Yagci and Das, 2015). Another set of works have explored the social media network with different text analysis to discover (latent) customer needs related to a product design or a specific product feature (e.g. Tuarob and Tucker, 2015; Zhou et al., 2015; Palmer, 2016). These works are quite impressive, but it is observed that they are rather task specific, which are not at all flexible enough to support different tasks in various domains. Furthermore, when the data from different social media channels as well as different data sources are not semantic integrated into a holistic view, it is not easy to avoid biased results. Meanwhile, in their works, much efforts have been investigated on the preparation of social media texts or identifying certain type of information, but the results from the preparation are difficult to be reused and integrated with other data sources to support further tasks. A limitation of the proposed approach is related to the configuration effort. To keep flexible and generic in the proposed approach, some configuration work need to be done before to use this approach.

6 CONCLUSION

In this paper, we proposed an approach to link the information and knowledge from MOL phase i.e. social media feedback to support product design. This approach uses data acquisition techniques to gather texts from social media channels and applies ontology based NLP methods on the gathered texts to extract necessary information and link them to ontologies. In this way, the social media texts are collected and directly semantic transformed into ontology Aboxes in a kind of pipeline. The ontology Aboxes, which are kept in knowledge bases, can be analysed to support various tasks in product design. The main advantage of using our approach is that it is flexible, highly configurable and general enough to be adopted in different application domains to support different product design tasks. In addition, ontology based semantic integration allows the integration of the feedback information that widely distributed in different social media channels and the information in diversity data sources to provide product designer a holistic view of feedback information.

A limitation of the proposed approach is related to the configuration effort. Indeed, if the given ontology is complex and texts from many different social media channels need to be analysed and mapped to the ontology, the current manual configuration process represent a major effort. Certain type of information is difficult to be extracted with manual lexicon based configurations. Furthermore, it is difficult to capture all needed texts from social media for design tasks, because the access to some social media channel is restricted and requires explicit permissions. Social media texts contains usually only certain types of information, and many needed information is possible not stated in the texts. Future research developments can be suggested to employee machine learning approaches to extract certain information and support (semi-)automated configuration. Further effort is also required to have a clear evaluation on the performance of the approach regarding the configuration tasks.

REFERENCES

Blei, D. M., Ng, A. Y. andJordan, M. I. (2003), Latent dirichlet allocation. the Journal of machine Learning research, 3,993-1022.

Chang, D. and Chen, C-H. (2014), 'Understanding the influence of customers on product innovation', *Int. J. Agile Systems and Management,* Vol. 7, Nos. 3/4, pp.348–364.

Cowie, Jim; Lehnert, Wendy. (1996), Information extraction. Commun. ACM 39, 1 (January 1996), 80-91. DOI=http://dx.doi.org/10.1145/234173.234209

Fan, W. and Gordon, M.D. (2014), 'The power of social media analytics', Communications of the ACM, 57(6), pp. 74–81. Available at: http://dl.acm.org/citation.cfm?id=2602574

Grishman, Ralph (1997), Information Extraction: Techniques and Challenges. In International Summer School on Information Extraction: A Multidisciplinary Approach to an Emerging Information Technology (SCIE '97), Maria Teresa Pazienza (Ed.). Springer-Verlag, London, UK, UK, 10-27.
Kietzmann, J. H., Hermkens, K., McCarthy, I. P. and Silvestre, B. S. (2011), social media? Get serious! Understanding the functional building blocks of social media. Business horizons, 54(3), 241-251.
Kiritsis, Dimitris (2011), Closed-loop PLM for intelligent products in the era of the Internet of things, *In Emerging Industry Needs for Frameworks and Technologies for Exchanging and Sharing Product Lifecycle Knowledge* 43 (5), pp. 479–501. DOI: 10.1016/j.cad.2010.03.002
Liu, B. (2012), Sentiment Analysis and Opinion Mining, Morgan and Claypool, 2012.
Madway, G. (2010, September 14), Twitter remakes website, adds new features. Retrieved December 5, 2016, from http://www.reuters.com/article/idUSN1411135520100915
Medhat, Walaa; Hassan, Ahmed; Korashy, Hoda (2014), "Sentiment analysis algorithms and applications", A survey. In: *Ain Shams Engineering Journal* 5 (4), S. 1093–1113.
Muslea, I. (1999), Extraction patterns for information extraction tasks: a survey. *In Proceedings of AAAI '99: Workshop on Machine Learning for Information Extraction*.
Obrst, L. (2003), 'Ontologies for semantically interoperable systems', *Proceedings of the twelfth international conference on Information and knowledge management, the twelfth international conference*. ACM Special Interest Group on Management Information Systems; ACM Special Interest Group on Information Retrieval, New Orleans, LA, USA, 3/11/2003 - 8/11/2003. New York, NY: ACM, p. 366. doi: 10.1145/956863.956932
Pang, B., Lee, L. (2008), Opinion mining and sentiment analysis, *Foundations and Trends in Information Retrieval* 2(1-2), pp. 1–135, 2008.
Peñalver-Martinez, Isidro; Garcia-Sanchez, Francisco; Valencia-Garcia, Rafael; Rodríguez-García, Miguel Ángel; Moreno, Valentín; Fraga, Anabel; Sánchez-Cervantes, Jose Luis (2014), Feature-based opinion mining through ontologies. In: Expert Systems with Applications 41 (13), S. 5995–6008.
Popescu, A.-M. and Etzioni, O. (2005), 'Extracting product features and opinions from reviews', *Proceedings of the conference on Human Language Technology and Empirical Methods in Natural Language Processing - HLT '05, the conference*, Vancouver, British Columbia, Canada, 6/10/2005 - 8/10/2005. Morristown, NJ, USA: Association for Computational Linguistics, pp. 339–346. doi: 10.3115/1220575.1220618
Riloff, E. (1999), Information extraction as a stepping stone toward story understanding. In: A. Ram and K. Moorman (eds), Understanding Language Understanding: Computational Models of Reading (MIT Press, Cambridge, MA, 1999).
Salas-Zarate, M. d. P.; Valencia-Garcia, R.; Ruiz-Martinez, A.; Colomo-Palacios, R. (2016), Feature-based opinion mining in financial news. An ontology-driven approach. In: Journal of Information Science.
Schouten, Kim; Frasincar, Flavius (2016), Survey on Aspect-Level Sentiment Analysis. In: IEEE Trans. Knowl. Data Eng. 28 (3), S. 813–830.
Palmer S. (2016), 'Crowdsourcing Customer Needs for Product Design using Text Analytics', Proceedings of the World Congress on Engineering, 2016(Vol I).
Tuarob, S. and Tucker, C.S. (2015), 'Automated Discovery of Lead Users and Latent Product Features by Mining Large Scale Social Media Networks', *Journal of Mechanical Design*, 137(7), p. 71402. doi: 10.1115/1.4030049
Wimalasuriya, D. C.; Dou, Dejing(2010), Ontology-based information extraction. An introduction and a survey of current approaches. In: Journal of Information Science 36 (3), S. 306–323. DOI: 10.1177/0165551509360123
Yagci, I.A. and Das, S. (2015), 'Design feature opinion cause analysis: A method for extracting design intelligence from web reviews', International Journal of Knowledge and Web Intelligence, 5(2), p. 127. doi: 10.1504/IJKWI.2015.075164
Zhao, L. and Li, C.(2009), Ontology based opinion mining for movie reviews. In 3rd International Conference Knowledge, Science, Engineering and Management,2009.
Zhou, F., Jiao, J. and Linsey, J.S. (2015), 'Latent Customer Needs Elicitation by Use Case Analogical Reasoning From Sentiment Analysis of Online Product Reviews', *Journal of Mechanical Design*, 137(7), p. 71401. doi: 10.1115/1.4030159

ACKNOWLEDGMENTS

The research presented in this paper is supported by the H2020 Framework Programme, in the context of the FALCON project (Grant no. 636868). The authors wish to acknowledge the Commission and all the FALCON project partners for a fruitful collaboration.