VISUAL TREND ANALYSIS

Ontology Evolution Support for the Trend Related Industry Sector

Jessica Huster$^{1,2}$ and Andreas Becks$^1$

$^1$Fraunhofer-Institute for Applied Information Technology FIT, Schloss Birlinghoven, 53754 Sankt Augustin, Germany
$^2$Informatik 5 (Information Systems), RWTH Aachen University Ahornstr. 55, 52056 Aachen, Germany

Keywords: Text Mining, Concept-drifts, Ontology Evolution.

Abstract: Ontologies are used as knowledge bases to exchange, extract and integrate information in information retrieval and search. They provide a shared and common understanding that reaches across people and application systems. In reality, domain specific and technical knowledge evolve over time, and so must ontologies. Creative domains, as for example the home textile industry, are representatives for quickly evolving domains. In this domain it is also important to provide methodologies for the visualisation of knowledge evolution. In this paper we report on our ontology-based trend analysis tool, which supports marketing experts and designers to identify trend drifts, and to compare the analysis results against the ontology. Furthermore means to adapt and evolve the ontology in accordance with the changing domain are provided.

1 INTRODUCTION

Ontologies have been proven to be useful in different fields such as information integration, search, and retrieval. Regarding the Semantic Web, ontologies with their definition of concepts and relationships provide the backbone for structured access and exchange of shared knowledge (Fensel, 2001b), (Flouris et al., 2008). Hence ontologies are seen as and used for “a shared and common understanding that reaches across people and application systems” (Fensel, 2001a). However, different definitions and broad interpretations exist about what can be referred to as ontology. These interpretations include simple catalogues, sets of structured text files, thesauri, taxonomies and also sets of general logical constraints, that enable automated reasoning (Welty et al., 1999). The latter models usually have a greater complexity than the former. The ontology used in this paper is a lightweight, terminological ontology that describes the domain of the home textile industry, i.e. curtains, furnishing fabrics and carpets. It mainly uses mainly sub-class-relationships.

Ontologies represent the current understanding of terms and their relations of a domain in a structured way, ensuring that a group of people agree upon the same meaning of terms. However, ontologies, once created, do not last forever because they do not only encapsulate stable knowledge. In reality, the domains which are modelled by ontologies evolve, and so must the ontologies to stay useful (Fensel, 2001b), (Stojanovic et al., 2002), (Haase and Sure, 2004). There is a need for dynamic evolution in the conceptualisation to ensure reliability and effective support through ontologies.

Meanwhile, the issue of adapting ontologies due to a certain need of change is addressed by several different, but also closely related research disciplines. This field is summarised under the term of ontology change and comprises several subfields, each focusing on another aspect of ontology change. Flouris et al. (2008) identify 10 subfields in their work, namely: ontology mapping, morphism, matching, articulation, translation, evolution, debugging, versioning, integration and merging. Regarding the Semantic Web the field of ontology evolution is of particular importance as distributed, heterogenous and dynamic domains represented by different ontologies are expected to cooperate in this field.

The term ontology evolution is used with different meanings in the literature. In this paper, we will adapt the definition by Flouris et al. (2007): ontology evolution is the “process of modifying an...
ontology in response to a certain change in the domain or its conceptualization”. Thus, our focus is on the process of adaptation of an ontology. We distinguish this process from ontology versioning which deals with maintaining several different, but related versions of the same ontology.

Creative domains are one example for quickly evolving domains in which new concepts and relationships are very frequently established. As an example for a creative domain in this paper, we will consider the domain of home textile industry. Stakeholders in this domain have to flexibly adjust to upcoming and descending trends. Therefore, the issue of modifying a domain ontology is important in this context. Furthermore, in such a domain it is important to support the modification of the domain ontology by appropriate change and update operations. Additionally, methodologies for the visualisation of the knowledge evolution are required as users need to be able to understand how the context and importance of certain concepts in the domain have changed. In this paper we present an ontology-based system that supports this industry sector to monitor trend relevant sources and identify drifts in their market. Hence, we set a focus on an intuitive change capturing, which forms the starting point for ontology evolution. Once the changes, which require a modification of the ontology have been identified, the system provides means to adapt the ontology accordingly.

The remaining paper is structured as follows: The next section presents an overview on existing approaches to the topic of evolving and maintaining ontologies. Section 3 presents our approach in more detail. Section 3.1 covers the difficulty in the creative domain of home textile sector followed by the system description (section 3.2). Section 4 presents the evaluation conducted in cooperation with carpet producing companies.

2 RELATED WORK

Many approaches focused on creating ontologies rather than on ontology evolution. This shows that these knowledge models have been treated mainly as static and that the encapsulated domain knowledge is assumed to not change (Haase and Sure, 2004).

The approaches of ontology creation and evolution can be distinguished into two main groups. One class of evolution methods follows a community based approach, whereas the other class of approaches tries to automate, at least part of the process and suggest computational methods for ontology creation.

An early work in this area to construct ontologies collaboratively is presented by Dominge (1998). His aim was to support a (possibly widespread) community in constructing an ontology, as the ontology represents a common view. He presents the two web-based systems Tadzebao and WebOnto, which complement each other. Tadzebao supports asynchronous and synchronous discussions on ontologies. WebOnto, on the contrary, provides features for collaborative browsing, creation and editing of ontologies.

The systems Ontoverse (Weller, 2006) and OntoWiki (Hepp et al., 2006) follow a wiki-based approach. This approach explicitly focuses on the collaborative construction of ontologies and supports the cooperative design process. Term meanings and problems can be discussed through the system by the user community before these terms are integrated into the system.

An important difference between these two systems is that Ontoverse has a role-based concept and restricts the usage to specific user groups. It allows parallel development of ontologies with different user groups in separated workspaces. The access to OntoWiki on the other hand is open and not limited. OntoWiki is realised using a standard Wiki, whereas Ontoverse developed a new wiki-based platform supporting the different steps of an ontology development process: a planning and conceptualisation phase; a phase of editing the structural data; and finally a maintaining phase, where updates, corrections and ontology enrichments are performed (Weller 2006). Both systems can be used also, to further evolve ontologies.

Holsapple and Joshi (2001) suggest applying a structured procedure for the collaborative working process. This procedure uses a Delphi-like method which incorporates feedback loops within the group. Step by step the terms of the ontology are enriched and better understood. In contrast to the wiki-based systems before, he does not suggest a support system.

Computational methods analyse texts and aim at detecting changes or trends in the terms. This field is known as emerging trend detection or topic evolution (Gohr et. al., 2009), (Kontostathis, 2004). The new terms serve as candidates for the ontology enrichment. An interesting approach is presented by Kruse (2009), even though not directly related to ontology evolution. He presents an interview based decision support system to identify drifts in the
meaning and usage of terms. The interview results span a semantic room of meanings, which is represented in a 3D visualisation for each person. Kruse performed an interview with 100 people in 2007 and performed the same study again in 2008. He asked the people to state their attitude regarding car brands, definition of status and moral concept as well as mobility in their daily life. Comparing the result of these two studies he identified a change in the definition of status. Status, once directly related to big, representative and expensive cars, is no longer connected to these things. It is rather linked with environmental protection, social justice and appropriate functionality. A rather small car is preferred by the people in their daily life. Already in 2007 Kruse identified these changes in the preferences, long before the market realized the change in the car branch. A year later, 2008, the press finally stated that they have problems selling big and luxury cars. This is the obvious result of the preference changes, already identified in the surveys by Prof. Kruse.

Other approaches go further and do not only extract term candidates but try to discover conceptual structures, also. Maedche and Staab (2000) present a framework that combines different methods to acquire concepts, and to identify relationships between concepts from natural language. Relationships both relevant and not relevant for the ontological taxonomy can be identified. They applied their system in the domain of telecommunications. For mining a concept taxonomy in the first step they used domain specific dictionaries. Further relations between concepts are discovered based on generalised association rules. These approaches aim at an (at least partial) automation of ontology construction and are subsumed in the field of ontology learning.

In our approach we explicitly rely on the experts and give support trough intuitive visualisation methods. The experience of Prof. Kruse with his system shows that changes are (unconsciously) realised by people before the changes reach printed media and are then realised consciously by the market or the public. Hence, the experts are an important knowledge source. The experts know their domain perfectly well and may have gut feelings and ideas how things are developing. The graphical visualisation on the other hand, supports the expert user to discover changes in domain specific knowledge in printed media and to identify missing concepts in the ontology on this way. Using the tool, they get means to monitor trend developments more efficiently by the visualisation of domain specific knowledge evolution. They can test their ideas and push them into the market to set new trends.

3 OUR APPROACH IN THE HOME TEXTILE SECTOR

3.1 Requirements

The home textile sector is a trend-related industry. Producers have to flexibly adapt to changing preferences and consumer behaviour. Companies and especially market analysts have to monitor particular fields for recent trends that may impact the company sales and success. For them, it is important and required to detect emerging topics early, and the way they evolve over time. The earlier the producers identify particular techniques as well as new colour and material combinations for a new design, the better they are able to refine the products and attract the interest of the customer. The producers are able to surprise the customer positively by new designs. On the one hand the companies have to seek for trends while on the other hand having to set those by varying colouring, colour families, materials and their combinations to create novelty and charm (Sauerwein, 2009). These products are mainly sold through innovation and design, nowadays. (Becella, 2007).

To address these requirements and provide flexible analysis of trend combinations according to the working tasks of designers and marketing experts, we developed an ontology-based system which allows analysing trend-relevant data sources using text mining technology. By accessing digitised fashion and trend magazines we enable a systematic evaluation of colour families and material groups which are mentioned in these sources. Marketing experts and designers can analyse the frequencies of colours, colour families, or material types from magazines and trend books and assess the development of colour and material statistics over time.

An initial domain ontology was created that serves as a starting point to detect trends and monitor the market relevant sources. This well known knowledge about concepts like colour, material, structure or design of surface is modelled in the ontology and forms the stable “anchors” to identify the terminological development, which is a-priori unknown. Marketing experts and designers recognise new terms, that evolve from the active use of these terms and concept combinations, and supervise their development over time. These terms may describe colours or surface structures, dominant in magazines or articles and supervise their development over time. These detected emergent terms currently not incorporated in the domain ontology are concept candidates for the ontology.
Integrated into the ontology, they can be used for a more specific search on particular trends to further ensure reliability, accuracy and effectiveness of the search (Klein and Fensel, 2001). Step by step the domain ontology evolves to a trend ontology. The search and evolution cycle is shown in Figure 1.

A system which supports designers and marketing experts in this sector should meet the following requirements:

- Systematic evaluation of colour families and material groups which are mentioned in fashion and trend magazines
- Assessment of the development of colour and material statistics over time
- Flexible access to trend information according to work tasks of designers (focus on particular areas, colours, material, etc. as well as focus on specific magazines)
- Intuitive means and features to modify the ontology for an iterative evolution

In (Becks, 2007) we reported on the overall design of the Trend Analyser system in the AsIsKnown project. The next sections describes the system features of our Term-Context-Language Analysis (TeCLA) component in more detail, followed by an evaluation in closed cooperation with marketing experts and designers of a carpet producer (cf. section 4). It combines the strengths of automatic text analysis for a visual, goal-oriented overview on trends with domain experience for iterative and efficient evolution of market specific trend ontologies.

### 3.2 System Overview

TeCLA is a semi-automatic, visual component, supporting the experts in monitoring trend relevant concept in relevant fashion magazines. It allows intuitive access to upcoming colour and material combinations. After the identification of trends the community of experts (marketing specialists, designer and trend scouts) is able to evaluate the identified drifts in the concepts, match it with known concept model in the domain ontology and possibly adapt it accordingly. The experts are thus able to form their trend specific ontology step by step. It is especially important to include the domain experts in the process of ontology evolution for several reasons (Fensel, 2001b), (Hepp et al. 2006). Firstly, the ontology aims at a consensual domain understanding and knowledge. Hence, the evolution process requires cooperation and exchange of information between different people, such as domain experts. The strength of the community cannot be replaced by a single ontology expert. Secondly, the process stays under control of the community, which enhances the acceptance of the resulting ontology. The experts are the ones who are working with the ontology. They can evolve ideas and form the ontology using their high degree of experience and implicit background knowledge. The process of trend detection is always related to experience and creativity. A fully automatic approach would not be accepted by such a user group, as the process is not under their control. Additionally, our interactive, user integrated approach meets their rather creative and weakly structured way of working.

#### 3.2.1 Visualisation and Change Detection

The first step in using TeCLA is the configuration of the analysis matrix for a particular trend detection (cf. Figure 2, showing the analysis result for the concepts “green” and “brick” in AIT). The analysis matrix defines the group of magazines and articles as well as the period and aggregation level of time for which trend relevant concepts shall be analysed in the magazines. The trend relevant concepts are selected from the ontology by the user e.g. colours, materials, architects. Based on these concepts TeCLA computes the concept term co-occurrences and represents them as *term context stars* (shown in the matrix cells in Figure 2), one for each selected concept in each cell of the analysis matrix.

The terms appearing in the context of the selected concept are determined using methods of natural language processing. Magazines are first linguistically pre-processed: The tokenised texts are automatically annotated with part-of-speech tags that indicate the grammatical categories of each word. By means of dictionaries, for each known term a matching concept from the ontology is attached (word sense tagging). See (Simov et al., 2004) for more details on the web-service based architecture used.

As soon as the user has provided the list of concepts that shall be analysed in the magazines,
The final task is then to visualise the extracted results. The graphical representation distinguishes between concepts and terms. Concepts are linked to the ontology and appear in the centre of the term context star. The terms that appear in the context of this concept surround the concept in form of bubbles. The relative size of each context term bubble corresponds to the number of times this term appears with the considered concept.

Suppose the user has analysed the context of the concepts “green” and “brick” in the magazine AIT, which is a professional journal on architecture, interior and technical solutions. In this magazine green and blue appear together as these colours are combined often. In September this colour combination has obviously grown in interest and became more popular when compared to the months before. It disappeared in the subsequent month; instead, green drew attention in the shade of “light green.”

The concept of our term context stars is related to the ideas of Heringer’s (1998) technique of visualising lexical fields and tag clouds, known...
Figure 4: Changing a term in the ontology due to a shift in the meaning from the Web 2.0 community (Hassan-Montero, 2006). This visualisation metaphor helps the users to easily recognise dominant concepts as well as the related term co-occurrences from the texts. Moreover, different term context stars of the same concept in different cells of the matrix can easily be compared regarding a change in occurrence frequency or a drift of context terms (cf. Figure 2).

The visualisation functionality was realised using a standard graphical programming library. Given a visualisation of term context stars in TeCLA, the domain expert can click on concepts or context terms to show all the articles in the considered magazines that contain the respective concept and context terms. The concept, as well as the clicked term is highlighted in the articles. This helps the users to get a quick insight in the more detailed context of the concept. Therefore this feature is especially important for the daily work: "What is precisely written in the article?" "Does it detail an interior of a restaurant or bar?" etc.

3.2 Support for Ontology Evolution

To support the evolution of the knowledge model, an appropriate feature is provided. This functionality allows the user to easily compare terms from the context star with concepts in the ontology. When the user activates the compare feature, all context terms which have a connection to the domain ontology are highlighted (cf. Figure 3). The user gets a quick overview about which terms which are already included in the ontology and which are not. The former ones can be used as search terms for further analysis, while the latter ones may be worth adding to the ontology. Additionally, the associated ontological category of the context term can be accessed by the experts.

The marketing staff and the designers in the home textile industry especially asked for a feature, which enables them to assess certain aspects more easily. "Is a colour concepts found rather in combination with other colours or materials?" "Is a specific form or surface coloured in a certain shade and hue or realised with particular materials?" All this information is important to develop ideas in their work as a designer or a marketing specialist. In July for example old bricks are used in some products (cf. Figure 3). In September old bricks are not only used to realise special designs but also used in combination with other materials to achieve particular effects (cf. Figure 3, the materials are inked in yellow). Bricks are used in combination with glass and steel, for example.
new concept the user specifies which concept of the ontology should be the super concept. He types in a name for the new concept (the URI is added from the system in the background), as well as the definition. In case of a change in the meaning of a term the user selects the option change term. Suppose the marketing and designers identified that there is a shift in the colour of brown. The tone formerly called brown is now called shift. Therefore, the expert selects the concept brown and types in the new term (cf. Figure 4). After committing the change by clicking the button, the change is directly visible in the ontology. In case of adding a new concept clicking the commit button means that a change request is send to an ontology expert. The ontology expert checks for the required update and possibly commits the change. The new version of the ontology is then made available to the system.

4 EVALUATION

To evaluate the success and usability of our system we used observational methods such as thinking aloud protocols and structured interviews to assess the functionality, usefulness and usability of the system. The users are marketing experts and designers from a carpet producer. They were given a few minutes system presentation to be able to use the features according to the given tasks during the evaluation.

In the testing scenario the users were asked to perform a whole walk-through of the system. They performed an analysis of different trend relevant concepts to identify emerging and descending trends and were asked to use the different features provided, to support their work. They were asked to write down their “trend findings”. Afterwards they were given another questionnaire to evaluate especially the usability and usefulness of the different features and the system as a whole.

The tests have shown that term context stars are a good visualisation to represent the term co-occurrences. Trend lines (increasing and decreasing) can be identified quite easily based on the colour and size of a term bubble of a context star. Common terms in different stars are easily recognised and accessed by using the comparison functionality of TeCLA. A general remark from the users was to change the colour coding of term highlighting in the different features. Instead of using an additional colour to mark the relevant terms, the irrelevant terms should be shaded in grey. As a result the important terms can be identified easier.

The overall results given by the users were positive; they would all use the tool in similar contexts again. The provided interaction possibilities and features as well as the usability were evaluated as good. The training prepared them well in using TeCLA for the given analysis tasks. Nevertheless, some result presentations can be improved to facilitate the access to the information.

5 CONCLUSIONS AND OUTLOOK

In this paper we presented an explorative approach for the stepwise evolution of a trend ontology in a creative, trend-related domain such as the home textile industry. The text mining analysis is coupled with a high degree of user interaction and community-oriented modification of the ontology. In addition to the related systems for ontology evolution as presented in section 2 our system provides and focuses on an intuitive visualisation that supports the trend observation over a period of time. Popular concepts and co-occurring terms in the texts can be recognised and analysed over time. The identified drifts give the experts an overview on the market changes and help to evolve the ontology. The system provides means to modify the ontology accordingly. Thereby the ontology can evolve with the domain thus better supporting the designer in searching the magazines efficiently for further trends and trend changes.

The community in this case is formed by marketing experts, designers, trend scouts and product managers. This is a group of domain experts with sound knowledge working closely together. They evaluate recognised upcoming and descending terms and decide what should be modified in the ontology accordingly. We do not provide collaboration support in the system, explicitly, as this group of experts is small and usually closely working together.

During the whole development cycle informal feedback was collected from our project partners from the domain of carpet production several times. In these evaluation phases it turned out early that an explorative, community based approach with a high degree of user involvement is necessary to establish trust in the analysis results. The structured evaluation proved this approach again. The user interaction helps to derive ideas on possible trend lines, taking into account the experience and background knowledge of the experts. In this
practical approach, we bring together the two aspects of a trend experience and intuition of experts as well as observation and analysis of relevant market sources.

As mentioned in the beginning the aspect of ontology versioning is closely related to the topic of evolution. In the trend sector, it is also helpful to have access to older versions of the ontology and to be able to follow the changes made in the ontology. Hence our work will focus on a history feature in the next step. This feature will list all the changes which were applied to get the current ontology. All the changes can then be retraced by the experts and provide additional information on the evolution of trend concepts.

ACKNOWLEDGEMENTS

The research presented in this paper has been funded by the AsIsKnown project (http://www.asisknown.org) within the Information Society Technologies (IST) Priority of the 6th Framework Program (FP6) of the EU, and by the Research School within the Bonn-Aachen Int. Centre for Inform. Techn. (B-IT).

REFERENCES


Fensel, D. 2001a. OIL: An Ontology Infrastructure for the Semantic Web IEEE Intelligent Systems, 16(2)


Kruse, P. 2009. Ein Kultobjekt wird abgewrackt, GDI Impuls, Wissensmagazin für Wirtschaft, Gesellschaft, Handel, Nr. 1

Maedcke, A., Staab, S., 2000. Mining Ontologies from Text, Proc. of the 12th Int. Conf., EKAW, 189-202, Juan-les-Pins, France


http://www.asisknown.org