SEMANTIC DRIFT IN ONTOLOGIES

Jon Atle Gulla, Geir Solskinnsbakk
Norwegian University of Science and Technology, Norway

Per Myrseth, Veronika Haderlein, Olga Cerrato
Det Norske Veritas, Norway

Keywords: Semantic Web, Ontology, Evolution, Text mining.

Abstract: Ontology evolution is the process of incrementally and consistently adapting an existing ontology to changes in the relevant domain. Even though ontology management and versioning tools are now available, they are of limited use for ontology evolution unless the desired changes are known beforehand. Ontology learning toolsets are often employed, but they require large document sets and do not take the existing structures into account. Semantic drift refers to how concepts' intentions gradually change as the domain evolves. When a semantic drift is detected, it means that a concept is gradually understood in a different way or its relationships with other concepts are undergoing some changes. A semantic drift captures small domain changes that are hard to detect with traditional ontology engineering approaches. This paper discusses a new approach to detecting and assessing semantic drift in ontologies. The method makes use of concept signatures that are constructed on the basis of how concepts are used and described. Comparing how signatures change over time, we see how concepts' semantic content evolves and how their relationships to other concepts gradually reflect these changes. An experiment with the DNV's business sector ontology from 2004 and 2008 demonstrates the value of this approach to ontology evolution.

1 INTRODUCTION

Ontologies are becoming increasingly important in enterprises' pursuit of more efficient IT architectures. The ontologies define standardized vocabularies that support application integration and more integrated operations inside and across enterprises. Also, new ontology-supported applications now range from intelligent information retrieval solutions to service composition and intelligent agents.

Ontology evolution is the timely adaptation of ontology structures to changes in the domain. The underlying requirement to all ontologies is that their content is consistent with the way phenomena are understood and referred to in the domain. When the perception of the domain changes, this has to be reflected in the ontology as well.

Unfortunately, developing and maintaining ontologies is still a tedious and expensive undertaking. As opposed to data models in traditional transaction systems, ontologies' large scope necessitates the involvement of domain experts of different backgrounds and different roles. As models of real world phenomena they are also intrinsically complex and hard to validate. On top of this the formal notation of many ontologies makes it difficult to maintain the models unless ontology experts are available.

Since ontologies need to be updated and evaluated at regular intervals, the maintenance costs tend to grow unacceptably high if appropriate tool support is not available.

Most ontologies today are maintained manually by dedicated teams of domain experts and ontology modelers. Traditional modeling techniques are applied, which requires long face-to-face sessions with modeling, discussion, and evaluation. For smaller updates, though, it should be possible to employ more cost-effective approaches with less human involvement. Most of the concepts and structures are already there, and the task is to verify whether anything has to be changed, added or deleted. Ontology evolution, thus, should lend itself better to tool support than full-fledged ontology engineering projects.

Gulla J., Solskinnsbakk G., Myrseth P., Haderlein V. and Cerrato O.
SEMANTIC DRIFT IN ONTOLOGIES.
DOI: 10.5220/0002788800130020
In *Proceedings of the 6th International Conference on Web Information Systems and Technology (WEBIST 2010)*, page
ISBN: 978-989-674-025-2

In this paper we present a new approach to ontology evolution that makes use of concept representations – signatures – that capture small semantic changes to concepts over time. Since these signatures are constructed automatically from textual descriptions of existing concepts, they are geared towards updating existing structures rather than developing new ontologies.

In section 2 we discuss the problem of semantic drift in ontologies. Section 3 is devoted to concept signatures, whereas Section 4 demonstrates how these signatures are generated to analyze evolutionary changes to a real industrial ontology. A discussion of results is given in Section 5, followed by related work in Section 6 and conclusions in Section 7.

2 SEMANTIC DRIFT

An ontology is formally defined as an *"explicit specification of a conceptualization"* (Gruber 1993). It provides an abstract simplified view of the world that is shared by a community and prepared for a particular purpose.

An ontology language like OWL represents this conceptualization in terms of classes, individuals, properties and various constraints and operators. Even though other languages choose other primitives, they tend to categorize phenomena along the same line to accommodate a sound logical foundation. For this paper, though, it suffices to assume that ontologies consist of concepts that are related – taxonomically and non-taxonomically – to each other.

2.1 Evolutionary Changes

Stojanovic et al (2003) define ontology evolution as a cyclic process consisting of *change capturing, change representation, semantics of change, change implementation, change propagation* and *change validation.* Whereas ontology management and versioning systems deal with the representation, implementation and propagation of changes, the more difficult part of change capturing has been left to manual effort and some limited ontology learning support.

The captured ontology changes fall into two distinct categories:

- *Existential Changes.* Existing ontology concepts may be deemed irrelevant, and new concepts may need to be added to the ontology. An ontology of computers, for example, may not need to include floppy disks any more, as these are not used by modern computers. Similarly, GPS receivers are now a natural part of a mobile phone ontology, even though it had nothing to do with phones 10 years ago.
- *Relational Changes.* Both taxonomic and non-taxonomic relationships between concepts may change over time. In the example above, GPS receivers may now be modeled as a part of a smart phone, and computers now are more closely related to games and entertainment than a few years ago.

In principle, changes may be imposed to the ontology from three kinds of analyses: *Structure-driven* changes are motivated from structural properties of the existing ontology itself. *Usage-driven* changes reflect changes in users' behavior over time, while *data-driven* changes stem from a modification of the underlying knowledge such as text documents (Stojanovic 2004).

Our approach combines the usage-driven and the data-driven approach to ontology evolution. The object of our analysis is a collection of text documents, though the documents are assumed to be allocated to the correct ontology concepts by the users.

2.2 Semantic Drift in Ontologies

A concept's semantic value – i.e. our understanding of the concept – may change over time in response to general changes to the domain or our own insight. Our perception of computers, for example, is very different from what people associated with computers when the first PCs were introduced. We say that the meaning of computers has drifted as the technology developed and computers got ever more powerful.

We may define the notion of *semantic drift* as *the gradual change of a concept's semantic value as understood by the relevant community* example, may not need to include floppy

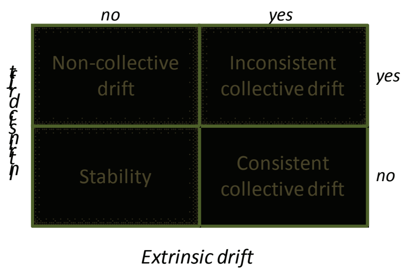

Figure 1: Types of semantic drift.

Moreover, we distinguish between *intrinsic* and *extrinsic* semantic drift.

Intrinsic drift means that a concept's semantic value is changed with respect to other concepts in the ontology. This will typically be reflected in changes to the relationships in the ontology. Extrinsic drift is when a concept's semantic value is changed with respect to the phenomena it describes in the real world. In the ontology an extrinsic drift may cause all kinds of changes.

Figure 1 sums up the nature of semantic changes associated with intrinsic and extrinsic drift. If a concept is exposed to extrinsic, but no intrinsic drift, it means that the whole ontology is undergoing a collective consistent drift that may not necessitate any changes to the ontology. On the other hand, no extrinsic drift and substantial intrinsic drift means that a concept's relationships to other concepts in the ontology may no longer be correct, even though the concept itself has not changed its meaning. In cases of both extrinsic and intrinsic drift we are dealing with inconsistent collective drift of concepts in an ontology that is no longer valid.

3 CONCEPT SIGNATURES

An ontology consists of inter-related concepts and normally has a sound logical foundation that allows some reasoning and verification checks. The meaning of an individual concept is however not entirely clear. Providing a taxonomic structure and adding associations between concepts give us some semantic clues, though it is not sufficient to recognize the concept in the real world. Logically, we assume the existence of an interpretation that maps for example the concept *Computer* to the set of all computers in the world, though for all practical purposes these interpretations are not available and machine-processable to us.

Most ontologies, thus, provide informal textual descriptions that try to help us understand how the concept is to be interpreted. In the petroleum ontology for ISO15926 there is a concept *Christmas tree* that is modeled as an artefact and decomposed into a number of specialized Christmas trees (Gulla 2009). These structures do not help us recognize Christmas trees in the petroleum business, though a simple natural language comment linked to the concept may give us an impression of what it is: *"An artefact that is an assembly of pipes and piping parts, with valves and associated control equipment that is connected to the top of a wellhead and is intended for control of fluid from a well."*

3.1 Definition

For our purposes it is more useful to link concepts to our linguistic world than to an imaginary interpretation function that points to real world phenomena. The textual description of Christmas tree above is not accurate, but is available and can be analyzed linguistically and statistically. As long as languages are used fairly consistently, the analysis of linguistic expressions can tell us how a community deal with a concept at particular points in time.

We define a concept signature as follows:

A concept signature $S_{c,t}$ is a materialization of the concept C through linguistic forms at some time t.

The signature is not a semantic representation of the concept. It merely shows how words and linguistic expressions are used to refer to and discuss the concept. The signature thus can be used to relate concepts at a linguistic level without being forced to formalize a mapping to real-world phenomena.

A concept signature is represented as a vector

$S_{c,t} = (u_1,.., u_n)$,

where u_i is the weight of linguistic unit i. Linguistic units may be individual words, phrases, argument structures, or any other linguistic structure that can be systematically extracted from text.

Examples of concept signatures from our DNV study are given in Figure 3. The linguistic units in this case are individual nouns and noun phrases, and their weights indicate their relative importance in understanding the concept. For *Consulting* in 2004, the top-ranked phrases *process industry* and *advanced cross-disciplinary competence* tell us that consulting was considered a cross-disciplinary activity with a primary focus on the process industry. The bottom-ranked phrase environmental performance reveals that DNV only rarely thought of consulting as related to environmental issues.

4 CONSTRUCTING SIGNATURES FOR DNV CASE

Det Norske Veritas (DNV) is an international company specializing in risk management and certification. As an industrial conglomerate DNV is involved in a number of business segments that each constitute a subdomain within risk management and

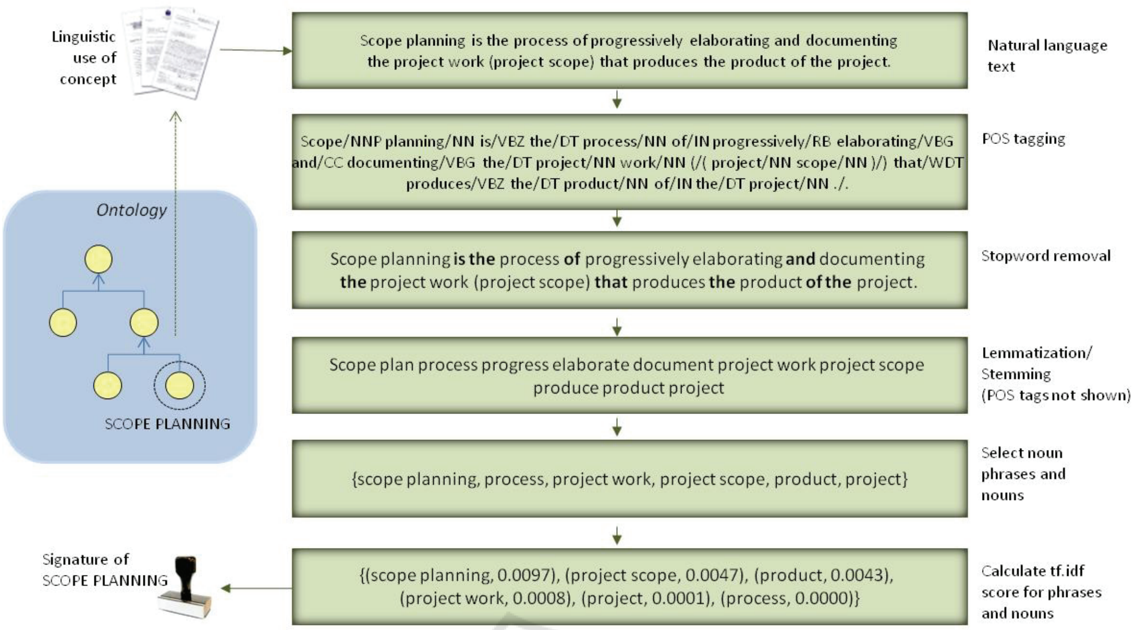

Figure 2: Generating concept signature for SCOPE PLANNING.

certification. Their web site mirrors their business activities and forms a taxonomy of DNV's business activities. Each web page at their site represents a concept in this taxonomy, and the text of the web page is our source for understanding this concept.

In 2004 this taxonomy counted 227 concepts (web pages) that on the average were described by texts of a few hundred words each. As their business domain evolved, their taxonomy was expanded into 369 concepts in 2008.

Constructing concept signatures for all their concepts in 2004 and 2008, we followed the procedure below for each concept:

- *Preprocessing Stages*: After collecting the text describing the concept, the text was tagged using the Penn treebank tag set. Irrelevant stop words were removed, and the resulting text was stemmed.
- *Selection of Linguistic Units*: Two lists were generated from the stemmed text above: (1) List of noun phrases, and (2) list of individual nouns only.
- *Signature Construction:* For every element of the two lists, the tf.idf score was computed. The tf.idf score of term t for concept C is given as

$tf_{i,c} * idf_{i,c}$,

$tf_{i,c} = f_{i,c}/max_j(f_{j,c})$ *and* $idf_i = log(N/n_i)$. *The variable* $f_{i,c}$ is the frequency of term i in concept C's text, $f_{j,c}$ is the maximum frequency of any term in this text, N is the number of concepts, and n_i is the number of concepts, whose text descriptions contain term i.

The two lists of elements with tf.idf scores are then merged into a vector representing the signature of that concept.

The whole procedure is illustrated in Figure 2, and examples of signatures generated are found in Figure 3. *Consulting* in 2004, as illustrated by the signature in Figure 3(a), was best understood as part of the process industry and international affairs. In 2008 the consulting concept had more to do with EFTA, performance issues and risk management.

5 USING CONCEPT SIGNATURES TO DETECT DRIFT

The concept signatures tell us how concepts are referred to in the linguistic communities. Our understanding of the totality of these terms is our implicit understanding of the concept. Since the concept signatures are formally represented as vectors, they can also be compared using standard information retrieval calculations like cosine similarity and euclidian distance. This enables us to run some automatic tests on possible semantic drift in ontologies.

Phrasal terms		*Single terms*	
process industry	*4.63*	*firm*	*1.95*
advanced cross-disciplinary competence	*4.63*	*compet*	*1.72*
international clients	*2.66*	*cross*	*1.69*
effective risk handling	*2.66*	*matur*	*1.35*
fast-moving world	*2.66*	*strong*	*1.30*
strong business orientation	*2.66*	*advanc*	*1.13*
international experience	*2.66*	*enhanc*	*0.92*
improved health	*2.66*	*dividend*	*0.92*
firm base	*2.66*	*differ*	*0.84*
genuine industry knowledge	*2.66*	*foundat*	*0.84*
worldwide network	*2.66*	*experienc*	*0.78*
strong technological competencies	*2.66*	*usa*	*0.78*
enhanced public confidence	*2.66*	*save*	*0.78*
direct savings	*2.66*	*manag*	*0.75*
unique independence	*2.66*	*technolog*	*0.75*
technology competencies	*2.66*	*perform*	*0.74*
better safety management	*2.66*	*base*	*0.74*
full access	*2.66*	*genuin*	*0.74*
experienced consultants	*2.31*	*provinc*	*0.74*
environmental performance	*2.11*	*fast*	*0.74*

(a)

Phrasal terms		*Single terms*	
efta inspection	5.91	efta	1.23
real performance	5.91	risk	0.57
industry best practices	5.21	softwar	0.55
risk management services	4.81	consult	0.55
right questions	4.81	knowledg	0.55
business functions	4.52	smart	0.51
operational excellence	4.30	inspect	0.50
knowledge management	3.71	busi	0.48
improvement opportunities	3.20	function	0.48
friday last week	2.95	manag	0.42
ict systems	2.95	abil	0.42
new premises	2.95	object	0.40
norwegian competition authorities	2.95	real	0.38
høvik	2.95	uncertainti	0.35
efta surveillance authority	2.95	question	0.34
efta team	2.95	technolog	0.33
other asset	2.95	complex	0.31
onboard dnv navigator	2.95	â	0.31
management control	2.95	km	0.31
smart ways	2.95	columbia	0.31
telecoms contract	2.95	copyright	0.31
columbia shipmanagement	2.95	improv	0.29
clients she threats	2.95	surveil	0.28
systems functionality	2.95	privaci	0.28
significant risk factor	2.95		
environment risk management	2.95		
in-depth industry insight	2.95		
smart organizations	2.95		

(b)

Figure 3: (a)Signature of 'consulting' from 2004. (b) Signature of consulting from 2008.

5.1 Individual Concepts

A concept exposed to extrinsic change will have significantly different signatures at different points of time. This means that the cosine similarity of signatures at times t_1 and t_2 will be below a certain threshold α:

$$Sim(S_{C,t1}, S_{C,t2}) < \alpha$$

where

$$Sim(X,Y) = \sum_{i=1}^{n}(x_i * y_i)/(\sqrt{\sum_{i=1}^{n} x_i^2} * \sqrt{\sum_{i=1}^{n} y_i^2})$$

The constant α depends on a number of factors and defines what counts as significant in this context. For our analysis of DNV, *Consulting* in 2004 and 2008 had a cosine similarity of 0.27. Other tests with consulting indicate that this is a fairly small similarity that reflects a genuine change of meaning over the years. The concept *Seaskill*, on the other hand, had a larger cosine similarity of 0.45 and seems not to be drifting significantly.

A low similarity score is an indication that the concept has undergone substantial extrinsic changes. To what extent that should be reflected in changes to the ontology depends on the possible changes to related concepts.

5.2 Non-taxonomic Relationships

Non-taxonomic relationships constitute semantic associations between concepts. Important permanent relationships tend to be modeled explicitly in the ontology, whereas less obvious or fluctuating ones are often left out of the model. If the importance or stability of a relationship changes over time, a reconsideration of which relationships to include will be needed.

Let us define the Concept Relation vector for concept *C* at time *t* as follows:

$$R_{C,t} = (r_{C,L1},\ldots, r_{C,Lm})$$

where $r_{C,Li} = Sim(S_C,S_{Li}) \geq \beta$

The concept relation vector for concept C provides a ranked list of concepts that are semantically related to C. The relation score, which is between 0 and 1, reveals the relative strength of the relationships compared to all other concepts related to C. The constant β gives a lower bound for when two concepts are to be regarded as related.

Normally, you would like to concentrate on high-level concept relationships first to make sure that ontology relationships are defined and kept at the highest possible level. This keeps the ontology more general and prevents unnecessary duplications from being introduced at lower levels in the ontology. Figure 4 shows the top-level concept relation vector for *Consulting* in 2004. Only top-level

Concepts most similar to Consulting in 2004	
process_industry	*0,313683*
asset_operation	*0,233114*
energy	*0,225704*
qualification_verification	*0,122025*
transportation	*0,102305*
classification	*0,086659*
organisation	*0,082843*
technologyservices	*0,075651*
careers	*0,072296*
certification	*0,067242*
publications	*0,066085*
press	*0,04665*
maritime	*0,045062*
location	*0,044662*

Figure 4: Concept relation vector for Consulting.

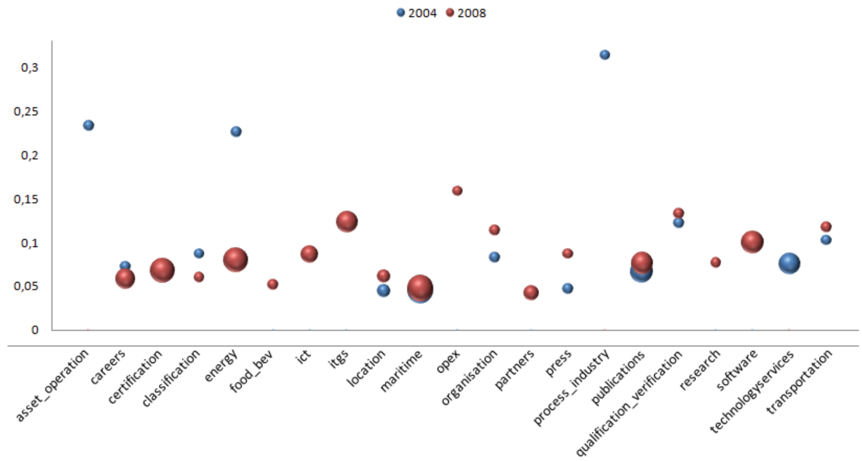

Figure 5: Consulting's non-taxonomic relationships to other major concepts in 2004 and 2008.

concepts are included, and all related subconcepts are incorporated into the top-level concept's relationship to consulting. That is, the relation score for each top-level concept like *Process_industry* and *Asset_operation* are average scores of all their subconcepts related to consulting.

Figure 5 shows how a high-level temporal analysis is conducted by means of concept relation vectors. In addition to using top-level concept relation vectors for 2004 and 2008, we have also recorded the number of subclasses supporting each top-level concept's relation score.

For every top level concept related to *Consulting*, there is one bullet for 2004 and one for 2008 in the figure. The strength of these relationships – the relation score – is indicated along the vertical axis, whereas the number of subclasses underlying every top-level concept is reflected by the size of the bullets. For example, consulting's relationship to careers has not changed much over the years with a relation score of about 0.06-0.07. However, in 2004 the relationship was limited to only one subclass of careers, while in 2008 there were relationships between consulting and 13 subclasses of careers. In the diagram this is shown by the much larger size of the bullet for 2008.

As seen from the results, the nature of consulting in DNV has shifted from maritime and process-oriented industries to ICT, software and risk management. This suggests significant intrinsic changes to the consulting concept that should impose changes to the ontology.

More generally, a substantial change of relation score to another concept necessitates an evaluation of whether this relationship should exist in the ontology or not. A small bullet means that the relationship is only relevant for a few subclasses and may therefore not be represented as a relationship to the top-level concept in the ontology. A large bullet, like for maritime in Figure 5, implies that many subclasses are related to the concept, indicating that the relationship in the ontology should be linked to the top-level concept rather than directly to its subclasses.

5.3 Taxonomic Relationships

Concept signatures may also be used to analyze the hierarchical structures of the ontology. In Figure 6 we have calculated the similarity between *Consulting* and all its specializations and parts for 2004 and 2008, filtered out those below a certain threshold β and ranked them according to similarity scores. A high similarity score means that the specialization is central to the core understanding of the superclass.

It is however not obvious how such a ranked list of specializations should be interpreted. Other experiments with concept signatures reveal that we should not expect a very high similarity between super and subordinate concepts, though there should always be some minimum similarity for the properties they share (Solskinnsbakk 2009).

As seen from the figure, the composition of *Consulting* has been fairly stable over these years. Specializations like *Process, General industries, Safety health environment* and *Enterprise Management* are equally central in 2008 as in 2004. A few interesting changes should be noted, though. *Asset operations* and *Project management* (PM) were seen as core activities of consulting in 2008, but were rather distant in 2004. We also see that DNV terminated its software consulting activities between 2004 and 2008.

6 DISCUSSION

We have in this paper shown how the notion of concept signatures helps us analyze the evolutionary aspects of ontologies. The method uncovers semantic drift among concepts in the ontology, both with respect to real-world phenomena and the concepts' relationships to other concepts in the ontology.

Our technique relies on good text fragments that describe or define the existing concepts in the ontology. Because it makes use of the existing ontology, it does not suffer from the noise that has hampered traditional ontology learning approaches. Only real concepts are subjected to the analysis, and we can concentrate fully on verifying the quality of these concepts as they are currently modeled. Since the analysis is geared towards the temporal development of the concepts, we need not to worry about the exact relationship between text and concepts, as long as we can assume that this relationship is unchanged over time. Unfortunately,

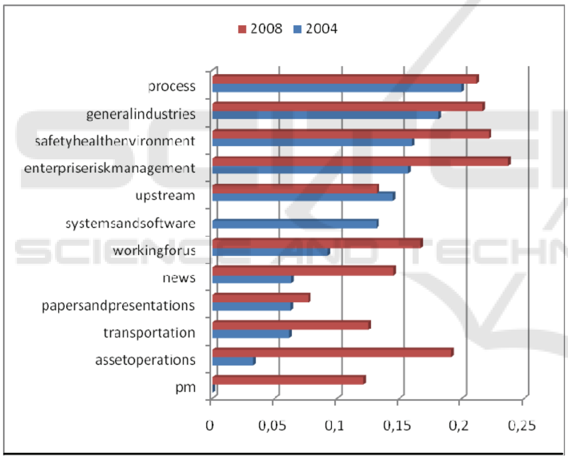

Figure 6: Specializations of *Consulting*.

this also means that the method will not detect any missing concepts in the ontology.

In a temporal perspective there will always be some semantic drift. Our understanding of concepts change as the domain change, and many concepts reflect more technological level or state of the art than fixed and permanent terminologies. This does not mean, though, that ontologies should be updated whenever a noticeable semantic drift is detected. Before updating the ontology, we need to understand both the nature of semantic drift and the extent of semantic drift among all the ontology concepts.

A fundamental problem of our current approach is the generation of concept signatures. Since we depend on texts attached to every single concept, these texts tend to be rather short and shallow. Our statistical approach would benefit from longer texts, from which more reliable statistical data can be extracted.

7 RELATED WORK

Our approach to detecting semantic drift draws on research on ontology learning and evolution (Haase & Sure 2004, Stojanovic 2004). However, standard data-driven ontology learning methods tend to use uncategorized text both to extract concepts and describe their properties (e.g. Gulla & Sugumaran 2008). This makes it difficult to take into account the existing ontology and any manual additions to it.

Some recent work on belief change theory (Flouris et al. 2006, Lee et al. 2004) and collaborative environments (Noy et al. 2006) provide alternative approaches to ontology evolution, though neither addresses the way concepts are materialized through language.

Enkhsaikhan et al. (2007) describe a method for building term clusters that describe existing top-level concepts like *Politics* and *Economy*. This enables an analysis of temporal concept development similar to ours, though their approach does not use vectors or linguistic characterizations of concepts.

Our focus on individual concepts' evolution rather than the ontology as a whole is similar to work done in logic and conceptual structures (Foo 1995, Wassermann 1998).

The idea of concept signatures is inspired by the concept vectors used in Su's ontology mapping approach (Su & Gulla 2006), though her vectors did not try to capture any temporal development of concepts. The vectors contained both definitional and non-definitional terms and were merely used to recognize product similarities across product catalogs.

8 CONCLUSIONS

This paper has presented a new approach to detecting semantic drift in ontologies over time. The notion of concept signatures is introduced and used to capture deeper linguistic characterizations of concepts.

The approach has been applied to an informal ontology maintained by a large enterprise in Norway. Data about the ontology from 2004 and 2008 were used to generate concept signatures and analyze the way the terminology has developed. The analysis shows that the method is able to capture small semantic changes to concepts that are hard to detect manually or by means of traditional ontology learning techniques. Primarily, these are changes to the concepts' relation to reality, but the method also uncovers secondary changes to the relationships among concepts in the ontology. The detected semantic changes shed light on why and how the ontology had been updated between 2004 and 2008.

Our current approach makes use of standard statistical methods for constructing concept signatures. If the textual descriptions of concepts are short, the statistical data is too limited to produce signatures of the necessary quality. Our future research, thus, will look into the use of more sophisticated linguistic techniques in the signature generation process. This includes both deeper grammatical analysis of sentences and utilization of semantic lexica.

ACKNOWLEDGEMENTS

We would like to thank the LongRec project, funded by the Research Council of Norway under project number 176818/I40, for supporting this research.

REFERENCES

Enkhsaikhan, M., W. Wong, W. Liu, and M. Reynolds. Measuring Data-Driven Ontology Changes using Text Mining. In *Proceedings of the Sixth Australiasian Conference on Data Mining and Analytics*, pp. 39-46. 2007.

Flouris, G., D. Plexousakis, and G. Antoniou. A Classifiation of Ontology Change. In Proceedings of the 3rd Italian Semantic Web Workshop, Semantic Web Applications and Perspectives (SWAP). 2006.

Foo, N. Ontology Revision. In Proceedings of the 3rd International Conference on Conceptual Structures (ICCS'95), Lecture Notes in Artificial Intelligence, Vol. 954, pp. 16-31, 1995.

Gruber, T. (1993). A Translation Approach to Portable Ontology Specifications. *Knowledge Acquisition*, **5**(2), 199-220.

Gulla, J. A. Experiences with Industrial Ontology Engineering. In J. Filipe & J. Cordeiro (Eds.), *Enterprise Information Systems. Lecture Notes in Business Information Processing 19*, Springer, 2009.

Gulla, J. A. and V. Sugumaran. An Interactive Ontology Learning Workbench for Non-Experts. In *Proceedings of the 2nd International Workshop on Ontologies and Information Systems for the Semantic Web (ONISW 2008)*, pp. 9-16, Napa Valley, October 2008. ACM Press. ISBN 978-1-60558-255-9.

Haase, P. and Y. Sure. *D3.1.1.b State of the Art on Ontology Evolution.* SEKT Deliverable, 2004.

Lee, K. and T. Meyer. A Classification of Ontology Modification. In *Proceedings of the 17th Australian Joint Conference on Artificial Intelligence*. 2004.

Noy, N., A. Chugh, W. Liu, and M. Musen. A framework for ontology evolution in collaborative environments. In *Proceedings of the 5th International Semantic Web Conference (ISWC'2006).* Lecture Notes in Computer Science. 2006.

Solskinnsbakk, G., J. A. Gulla, V. Haderlein, P. Myrseth, and O. Cerrato: Quality of Subsumption Hierarchies in Ontologies *Accepted for publication at NLDB 2009: 14th International Conference on Applications of Natural Language to Information Systems,* June 24-26 2009, Saarland University, Saarbrücken, Germany.

Stojanovic, L. *Methods and Tools for Ontology Evolution.* PhD thesis. University of Karlsruhe. 2004.

Su, X. and J. A. Gulla. An Information Retrieval Approach to Ontology Mapping. *Data & Knowledge Engineering,* Volume 58, No. 1, pp. 47-69, 2006.

Wassermann, R. Revising Concepts. In Proceedings of the 5th Workshop on Logic, Language, Information and Communication (WoLLIC-98), 1998.