Evolution in the Ontology Based Knowledge Management Systems

Nenand Stojanovic  
*University of Karlsruhe*, nst@aifb.uni-karlsruhe.de

Ljiljana Stojanovic  
*University of Karlsruhe*, ljiljana.stojanovic@fzi.de

Siegfried Handschuh  
*University of Karlsruhe*, sha@aifb.uni-karlsruhe.de

Follow this and additional works at: [http://aisel.aisnet.org/ecis2002](http://aisel.aisnet.org/ecis2002)

**Recommended Citation**

Stojanovic, Nenand; Stojanovic, Ljiljana; and Handschuh, Siegfried, "Evolution in the Ontology Based Knowledge Management Systems" (2002). ECIS 2002 Proceedings. 123.  
[http://aisel.aisnet.org/ecis2002/123](http://aisel.aisnet.org/ecis2002/123)
EVOLUTION IN THE ONTOLOGY-BASED KNOWLEDGE MANAGEMENT SYSTEMS

Nenad Stojanovic
Institute AIFB - University of Karlsruhe, Germany
+49 721 608 7363
nst@aifb.uni-karlsruhe.de

Ljiljana Stojanovic
FZI Research Center for Information Technologies at the University of Karlsruhe, Germany
+49 721 9654 804
Ljiljana.Stojanovic@fzi.de

Siegfried Handschuh
Institute AIFB - University of Karlsruhe, Germany
+49 721 608 6554
sha@aifb.uni-karlsruhe.de

ABSTRACT

An ontology-based knowledge management system uses an ontology to represent explicit specification of a business domain and to serve as a backbone for providing and searching for knowledge sources. But, dynamically changing business environment implies changes in the conceptualisation of a business domain that are reflected on the underlying domain ontologies. Consequently, these changes have effects on the performance and validity of the KM system. In this paper we make an analysis of the problems induced by using not-evolved ontologies and present an approach for enabling consistency of the description of knowledge sources in an ontology-based KM system in the case of changes in the domain ontology. This approach is based on our research on ontology evolution and ontology-based annotation of documents. The proposed method is implemented in our semantic annotation framework so that efficient acquiring and maintaining of ontology-based metadata is supported.

1. INTRODUCTION

Ontology-based knowledge management systems represent a class of KM systems that are centred on an explicit conceptualisation of a domain model, so called domain ontology [Gr93]. This ontology provides semantic means for structuring knowledge repository, which enables more effective providing and searching for knowledge in a knowledge management system. For example, a content of a knowledge sources can be described using some controlled domain vocabulary constrained by the set of assumptions about intended meaning of used terms. That is inherently provided by using an ontology and the underlying process is called semantic annotation [Ha01]. In that way, when one
searches for a knowledge source that is about the animal “jaguar” it is avoided to retrieve a source that describes the “jaguar” car-maker.

The advantages of using ontologies in structuring information in a knowledge management system are already shown [St01], but the main bottleneck in the development of such systems (as in any knowledge based system) weather the development of the domain ontology itself. We mention only several challenges an ontology engineer has to cope with:

- an ontology is a model of a real domain and a domain entity can be conceptualised in different manners in the ontology (for example, is the domain concept Computer Science Course, a concept or an instance in the underlying ontology);

- an ontology is the common-shared agreement about the meaning of the used domain entities and therefore the ontology development should be a collaborative process, in which the different views and arguments about the same domain entity have to be compromised;

- some drawbacks or gaps in the structure of the domain ontology can be obtained only during the using of that ontology in the real applications (for example, the missing relation between two concepts);

- modelled domain is changed over time and consequently, the ontology is the matter of continual changes.

From the knowledge management point of view this last point is the crucial one. In knowledge management system the domain ontology reflects, indirectly, the characteristics of the business model of the enterprise and should be changed according to the changes in the business policy. In the high changed business environment these changes occur very often. An inappropriate ontology evolution system can lead to a not domain-relevant knowledge management system, i.e. the system that posses knowledge which is (partially) not relevant for the given domain. Although this change management problem has great impact on the business value of a knowledge management system, this problem is not well addressed in the KM literature [Ha00]. Our ongoing research is oriented toward a framework for dealing with changes in the domain in a knowledge management system.

In this paper we present an approach that enables consistency in the annotations of knowledge sources in the case of changes in the domain ontology. The approach is based on our research in the area of ontology evolution and ontology-based annotation of documents [Ha01]. The proposed method is included in our semantic annotation framework CREAM, so that the efficient acquisition and maintenance of ontology-based metadata is supported.

The benefits of the proposed approach are manifold:

- In the case of changes in the domain ontology, annotations of knowledge sources can be automatically updated;

- An ontology-evolving model enables the categorisation of required/derived changes. In such a way an incorrectness, which leads to the more critical decreasing of the system’s performances, can be managed firstly;

- A special ontology for the maintenance of the annotation is introduced – herein after called evolution ontology. It offers new search possibilities for knowledge sources, not only according to the content, but also according to the author, date, format, relevance and their combinations.

From the knowledge management system point of view the proposed approach will enable us to develop a robust knowledge management solution that copes with the high-changeable business conditions.

The paper is organised as follows: Section 2 describes the anomalies in an ontology from the point of view of an ontology-based knowledge management system. Section 3 explores the problem of changes in the ontology and analyses the effects of the change on the ontology itself and on the underlying objects. In section 4 we describe a method to analyse and propagation changes made in the ontology.
Further we present an integration tool for implementing this method in a KM scenario. Before we conclude, we give a survey of related work in the categories knowledge management, ontology evolution and annotation environments.

2. THE MAINTENANCE PROBLEM IN AN ONTOLOGY-BASED KNOWLEDGE MANAGEMENT SYSTEM

There are two main causes for making changes in domain ontology that serves as the backbone in a knowledge management system:

- a (top) manager can analyse the business environment and propose some changes in the business model and consequently in the underlying ontology and

- a (decision making) system can analyse the past data and discover some changes in the business policy, which are consequently reflected in the underlying ontology.

These reasons can be seen as top-down deduction and bottom-up induction of changes, respectively. The second one is related to the learning changes in the ontology from the real-data, obtained from the use of the ontology-based knowledge management system and our research is focused on this approach - change induction. Only to note, that although the causes of changes in the mentioned approaches are different, the both of them are resolved in the same way, which is described in the section 3.

In this section we briefly analyse types of anomalies in an ontology and the consequences of the using such incorrect domain ontology in the knowledge management applications, motivating in that way research in the ontology evolution.

From the point of view of the (mall-) functionality of the ontology in the underlying knowledge management system, we distinguish two types of anomalies:

- structural, which disables the application of (part of) the ontology (for example, the definition of some ontological entities are not correct and the using of such entities is not possible). This kind of anomalies is on the syntax level and can be discovered automatically.

- semantic, which disables the application of the ontology in the right manner (for example some structures in the ontology are not optimally designed). This kind of anomalies can result in the incorrect domain model, which causes some mistakes in the underlying knowledge management system, for example some knowledge elements are missing in the knowledge repository. One example for such a not correct domain model is represented on the Figure 1. Moreover, this kind

**Figure 1. Example of an incorrect domain model:**

Concept “Customer” is divided into three subconcepts: “Software”, “Consulting” and “Hardware”, according to the provided service. All three subconcepts have the property “bonus”, among others. It means that “bonus” is the substantial characteristic of the concept “Customer” and that all the instances of the “Customer” (including instances of subconcepts) have to have that property. Therefore the property “bonus” should be defined on the level of the concept “Customer”. In the opposite case, it is possible to define an instance of the concept “Customer” which will have only “propertyX”.

- semantic, which disables the application of the ontology in the right manner (for example some structures in the ontology are not optimally designed). This kind of anomalies can result in the incorrect domain model, which causes some mistakes in the underlying knowledge management system, for example some knowledge elements are missing in the knowledge repository. One example for such a not correct domain model is represented on the Figure 1. Moreover, this kind
Evolution in the Ontology-based Knowledge Management Systems

of anomalies can lead to the sub optimal domain model, which makes the use of the domain model less efficient in the particular problem-solving. One of the benefits of resolving this anomaly can be the optimisation of the querying the knowledge repository, as well as the discovery of new knowledge, highly relevant for the particular problem.

In the next sections we describe a method (section 3, 4) to analyse and propagate changes made in the ontology. Furthermore, we present an integration tool (section 4) for implementing this method in a KM scenario.

3. ONTOLOGY EVOLUTION

One critical point in applying ontologies to real-world problems is that domains are changing fast (new concepts evolve, concepts change their meaning, new business rules are defined, etc.) and user needs are changing, too. Thus, the corresponding ontologies have to evolve as well. Ontology evolution is the timely adaptation of the ontology to the changed business requirements, to the trends in the ontological instances and to the way of using of the ontology-based applications, as well as the consistent management/propagation of these changes because a modification in one part of the ontology may generate subtle inconsistencies in other parts of the same ontology, in the ontology-based instances, depending ontologies and applications. This variety of causes and consequences of the ontology changes makes ontology evolution a very complex process (figure 2) that is described in the following.

3.1 Semantics of change

An ontology has to be consistent according to its structure (concepts, inheritance graph, relations, axioms). This is “semantics of change” phase (cf. figure 2 (A)) that refers to the effect of the change on the ontology itself. In order to retain consistency of the ontology, set of required changes is expanded with the additional (derived) changes in ontology. For example, the deletion of relation
domain can provoke the deletion of the relation as well in the case that there are no other concepts defined as domain of this relation.

The additional changes in the ontology are derived automatically. The approach is based on the sound and complete set of axioms (provided with an inference mechanism) that formalises the dynamic of the ontology evolution. The compliance of the available ontology changes with the axioms automatically ensures ontology consistency, without need for explicit checking as incorrect ontology version cannot actually be generated [Fr00]. While the focus of the paper is on the knowledge management, we will omit here the description of our approach used for “semantics of change” and concentrate on the “change propagation” problem, which has great impact on the knowledge searching process.

Inputs of this phase (A) are required changes (1) and source ontology (2a) and outputs are list of required and derived changes (3) and modified source ontology (2b).

3.2 Change propagation

Potentially, an ontology change might corrupt the instances, dependent ontologies as well as application programs running against the ontology and/or the knowledge base. The task of the change propagation phase is automatically bringing all dependent elements to a consistent state after an ontology update has been performed. Block (B) in the figure 2 depicts this phase. Output is the list of changes (7), which have to be done. In the rest of the section we will analyse the effect of the change propagation on the corresponding inputs.

Effect of changes on the dependent ontologies

An ontology update might corrupt ontologies that depend on the modified ontology. They are built from the modified ontology or they import it. This problem could be solved by recursive applying ontology change procedure on these ontologies in order to preserve their conceptual, structural and behavioural consistency [Fr00].

Effect of changes on the ontological instances

When the ontology is modified, the instances need to be changed in such a way that the ontology and instances remain consistent with each other. Basically, if the ontology is modified instances must be transformed to confirm to the modified ontology. It means that continuous adaptation of the annotated information to the new semantic terminology and relationships is necessarily.

Effect of changes on the applications

Changes in the ontology might invalidate applications that are already running on top of the ontology and the knowledge bases, especially if they rely on certain schema characteristics, which are lost after the ontology update. In the ideal case, the conceptual knowledge that is necessary for an application should be merely specified in the ontology. However, practice applications also use an internal model that may become incompatible with the ontology [KF01]. Moreover, although the application’s programs are written to be as generic as possible, there are a certain number of “hard-coded” elements that should be treated special in some way. In most of the web applications, where some queries are “hard-coded” into the service that is invoked as a response on the specific action, the query rewriting process is needed [FL96].

4. EVOLUTION OF THE METADATA

This section introduces the backbone of our approach - evolution ontology that supports, alleviates and automates the evolution process. Thereupon, we present our method for solving the change propagation problem and our annotation framework, which integrates the ontology evolution process.
4.1 Evolution ontology

Since ontology evolution requires additional meta-level reasoning capabilities that allow inspecting changes and their logical dependencies, we define a special ontology, so-called evolution ontology. We distinguish between domain ontology that is changed and the evolution ontology that enables better management of these changes.

Ontological changes are represented using the top level concept "Change", its subconcepts (AddConcept, AddRelation, etc.) and its relations [Ol99]. For every change, it is also useful to know who is author of the change and when it is happened (time). The cause of the change is used to represent the source of the change (business requirements or the learning process) and the relevance of the change describes whether and how it can fulfil the requirements. Also, ontology evolution is managerial process and it needs some properties to support decision-making like cost, priority, etc. Order of the changes is also very important while it enables recovery of implemented changes. Moreover, change propagation cannot be done after every change in the ontology (it requires too much time) even thought the change causes instance inconsistency. Consequently, only the order of the changes can guaranty that the instances “picture” the real status of the ontology structure. To solve semantics of change problem, the evolution ontology contains axioms that derive additional changes. Similarly to the ordering of the change: this type of the dependency between changes is represented as a relation parentChange.

The second part of the evolution ontology represents semantic information about the domain ontology explicitly (relations prototypical, primary_key, etc.), because the conceptual structure of the evolution ontology aims to provide enough mechanisms to deal with problems of syntax as well as semantic inconsistencies that arise when the domain ontology is changed [TB01]. The third part of the evolution ontology aims to support data-driven self-improvement of the domain ontology. For example, the fact that these are no instances of some concept is a sign that this concept should be deleted. We enforce formal discovering of changes by representing these heuristics as axioms in the evolution ontology.

The benefits of using the evolution ontology are manifold: First, changes are formally represented. Second, a history of changes is stored. Third, based on the formal representation and the history of changes the change-propagation problem may be approached. Using the same representation model for the ontology and analysis of changes simplifies storage and allows reuse of system components like searching.

4.2 Evolution of the metadata

In this section we present our method for the change propagation problem based on consistency analysis of already existing metadata and the performed change in the domain ontology. It is divided into three steps described in the following.

Metadata capturing

When an ontology is modified, instances need to be changed in such a way that the ontology and instances remain consistent with each other. If the instances are on the Web, they are collected in the knowledge base using tools like focused crawler¹ (process “capture” in the figure 3). In order to speed up the whole change propagation process, only the instances that depend on the change are gathered. This dependency information is obtained from the instance of the evolution ontology that represents the performed change. Moreover, the output of this step is one list that makes references between located instances and Web documents.

¹ Kaon.semanticweb.org
The main problem is how to find an application that uses the ontology that is changed. An application can be semi-automatic maintained only if exists metadata describing which ontology and/or ontological entities that application uses. Thus, annotation of applications is necessary.

**Metadata analysis**

In the second step, automatic translation of the instances is performed according to the changes in the ontology [SSV02]. In order to avoid overhead of the system, which may heavily increase if the changes are performed every time the ontology has to be modified, the categorisation of the changes is embedded in the evolution ontology. We distinguish between:

- ontology-extending changes that do never have an impact on the existing instances (e.g. creating a new relation);
- changes that provoke syntax inconsistencies in the ontological instances (e.g. deleting a concept that already has instances);
- changes that provoke semantic inconsistencies in the ontological instances (e.g. creating a new sibling concept does not lead to the invalidity in the set of instances but an analysis of the meaning of the instances is needed).

The axiomatic part of the evolution ontology enables the verification of the formal characteristic of the instances. The analysis of the semantic consistency is based on the meta information (e.g. primary_key) defined in the evolution ontology.

This step provides an output in the form of list of modified instances with the reference to the corresponding resource (knowledge source). Only this step is performed in the case that instances are already gathered in the knowledge base.

**Generation of a proposal for modifications**

In the last step “out of date” instances on the Web are replaced with the corresponding “up-to-date” instances. As depicted in the figure 3, some modifications of the instances can be done automatically (process “update”), but for the instances that are “write-protected” the notification has to be sent (process “notification”) to the author of the annotation in order to inform her/him about the changes and to suggest how to correct the instance. Information about author is saved in the property “Author” in the evolution ontology.

Using the method for metadata evolution does not solve all problems. However, we provide guidelines, which suggest which resources’ metadata have to be checked, and eventually changes to run again the changed ontology.

**4.3 Framework**

In order to support the proposed approach for ontology evolution based on the maintenance of the instances we have adapted our CREAM framework [Ha01] presented in the figure 3. The Evolution Ontology, Evolution Component and related links are the new elements and they are described in the previous section.
Figure 3. Architecture of CREAM

**Document Editor/Viewer:** The document editor/viewer visualizes the document content and the annotations.

**SOEP² - Ontology and Fact Editor:** The user can browse and edit the ontology and retrieve for one concept all instances or for one instance all properties.

**Crawler:** The creation of relational metadata must take place within the Semantic Web. During metadata creation subjects must be aware of which entities exist already in their part of the Semantic Web. This is only possible if a crawler makes relevant entities immediately available.

**Annotation Inference Server:** Relational metadata, proper reference and avoidance of redundant annotation require querying for instances, i.e. querying whether and which instances exist. For this purpose as well as for checking of consistency, we provide an annotation inference server in our framework. The annotation inference server reasons on crawled and newly created instances and on the ontology. It also serves the ontological guidance and fact browser, because it allows querying for existing concepts, instances properties.

**Document Management:** In order to avoid redundancy of metadata creation efforts, it is not sufficient to ask whether instances exist at the annotation inference server. When a metadata creator decides to capture knowledge from a Web page, he does not want to query for all single instances that he considers relevant on this page, but he wants information, whether and how this Web page has been annotated before. Considering the dynamics of HTML pages on the web, it is desirable to store annotated web pages together with their annotations. When the web page changes, the old annotations may still be valid or they may become invalid. The metadata creator must decide based on the old annotations and based on the changes of the web page.

5. RELATED WORK

**Knowledge management and annotation/ontology evolution**

As known to authors the problem of maintaining description (annotations) of knowledge sources in an ontology-based KM system in the case of changes in the domain ontology is not treated in the

---

² kaon.semanticweb.org
literature and therefore we here present an analysis of the annotation systems for the knowledge management purposes. The last presented system gives the best view on the maintenance problem in the knowledge management community. Annotate [Gi99] is a system that uses information retrieval methods to support KM in an organisation. It enables document annotations on the web and captures global usage history. Annotate is not ontology-based and therefore does not treat the problem of managing validity such knowledge item descriptions. In [DPP00] paper author presents several issues with the design and implementation of organisation memories in distributed companies. They have designed a tool, based on the domain model in the form of ontology, capable to capture the content of the documents and the context, in which they were created. A sophisticated retrieval engine can retrieve the annotated documents based on their context. The presented system seems very similar to ours; it has very suitable user interface which support process of creating document annotations, it is integrated in the general ontology engineering environment, but it is not adapted to new web infrastructure (Semantic Web) and does not consider ontology evolution problem.

A very interesting, field research study of managing changes in a knowledge management system is given in [Ha00]. The authors consider two types of changes: (i) functional changes that are about new KM-systems in the organization, new versions of a KM-system and new features in one KM-system and (ii) structural changes that deal with new business models, new subsidiaries and new competencies in the organisation. The results of the study show that managing the evolution of KM-systems on an ad hoc basis can lead to unnecessary complexity and KM-systems failures and that KM research has paid little attention to the evolution of KM-systems.

Ontology evolution

There are very few approaches investigating the problems of changing in the ontologies. The most similar approach to our approach is described in the paper [KF01]. As the authors also mentioned the most important flaw is the lack of a detailed analysis of the effect of specific changes on the interpretation of data.

The problem of schema evolution and schema versioning support has been extensively studied in relational and database papers. [Ro96] provides an excellent survey on the main issues concerned. [Fr00] introduces an approach to schema versioning, which considers a (conceptual) schema change as a (logical) schema augmentation. In contrast to our approach, this semantic approach does not address the change propagation problem, which concerns the effects of schema changes on the underlying instances. For the change propagation problem, several solutions have been proposed and implemented in real systems. In all cases, simple default mechanisms can be used or user-supplied conversion functions must be defined for non-trivial extant object updates. However, there are no approaches that treat data on the web.

Annotation

We know of three major systems that intensively use knowledge markup in the Semantic Web, viz. SHOE [HH00], Ontobroker [De99] and WebKB [PP99]. All three of them rely on knowledge in HTML pages. They all started with providing manual mark-up by editors. However, our experiences [Er00] have shown that text-editing knowledge mark-up yields extremely poor results, viz. syntactic mistakes, improper references, and all the problems sketched in the scenario section. The approaches from this line of research that are closest to CREAM are the SHOE Knowledge Annotator and the WebKB annotation tool. The SHOE Knowledge Annotator is a Java program that allows users to mark-up webpages with the SHOE ontology. The SHOE system [Lu97] defines additional tags that can be embedded in the body of HTML pages. The SHOE Knowledge Annotator is rather a little helper (like our earlier OntoPad [Fe99], [De99]) than a full-fledged annotation environment. WebKB uses conceptual graphs for representing the semantic content of Web documents. It embeds conceptual graph statements into HTML pages. Essentially they offer a Web-based template like interface like knowledge acquisition frameworks described next.
6. CONCLUSION

Ontology used in an ontology-based KM system is related to the business strategy and indirectly to the business environment. In the highly changing environment it is obvious that an ontology as a domain backbone is also a matter of change. The changes in the ontology have to be propagated to all ontology-based descriptions of the knowledge sources in order to enable consistency of the searching process.

In this paper we have presented an approach for enabling consistency of the descriptions of the knowledge sources in the case of the changes in the domain ontology. The approach is based on our research in the ontology evolution and the ontology-based annotation of the Web documents. The proposed method is implemented in our semantic annotation framework so that efficient acquisition and maintenance of the ontology-based metadata are supported.

The proposed approach has many benefits: automatic updating of the ontology-based description of the content of the knowledge sources, new possibilities for searching for the knowledge sources according to the author, date, format of the knowledge sources, to name but a few. From the knowledge management system point of view the proposed approach enables us to develop robust knowledge management solution, which copes with the high-changing business conditions.

Combining this approach with the ontology learning methods, which enables learning ontologies from the knowledge sources, leads us to the some kind of self-organising knowledge management systems.

Acknowledgements: We thank our colleagues and students at the Institute AIFB, University of Karlsruhe, and FZI Research Center for Information Technologies at the University Karlsruhe, without whom the research would not have been possible – especially Steffen Staab and Alexander Mädche. Research for this paper was partially financed by EU in the IST-2000-28293 project “OntoLogging” and by US Air Force in the DARPA-DAML project “OntoAgent”.

REFERENCES


