BUSINESS ONTOLOGIES COOPERATION

Mina Ziani, Danielle Boulanger and Guilaine Talens
University of Lyon – Jean Moulin, Modeme Team, 6 Cours Albert Thomas, Lyon, France

Keywords: Ontology alignment, Ontology integration, Hybrid ontology, Mappings, Semantic interoperability.

Abstract: In the context of the cooperation between heterogeneous and distributed information sources, ontologies are a main issue. To represent shared knowledge, an hybrid domain ontology is designed and to respect each point of view of different experts, local ontologies are created. Since experts are willing to cooperate, similarities must be identified to build mappings between the concepts of the different ontologies. We propose a computer-aided system to allow the experts to choice similarity measures on demand. We apply this work to the geotechnic domain which involves various businesses.

1 INTRODUCTION

The knowledge management in the context of heterogeneous and distributed information sources is a great challenge. The difficulty is to represent all the domain knowledge specificities and to allow the cooperation between experts with different points of view. Ontologies are a promised approach for the knowledge representation in a formal way.

We propose in this paper an approach to establish interoperability between knowledge contained in different local ontologies. An hybrid ontology (Visser, 2002) is designed. It consists to describe each information source in a local ontology and to represent the vocabulary shared by all the experts.

Each expert community built her own business ontology. A global ontology is automatically designed and contains only the common concepts and properties of all the local ontologies.

To reach cooperation between the several ontologies and to allow semantic interoperability, different techniques are used. In particular, ontology alignment, or matching, is the process of determining relationships or correspondences (subsumption, inclusion …) between entities of different ontologies. The correspondences are also called alignments. Very often, these relations are "equivalence relations" discovered through similarity measures between ontologies' entities. In the existing alignment system, these measures are used, alone or aggregated according to a particular strategy. This one depends to the domain, the type, and the use of the ontology. So, we have developed a generic computer-aided system which guides the experts to choice the similarity methods and measures to be used in the alignment process according to the ontology characteristics.

We applied this work to the geotechnics. It is a complex domain involving different businesses: Project management, geologists and chemists…To design ontology for this domain we use information contained in the geotechnic referentials and documents.

We present some tools and frameworks for merging or aligning ontologies. After, we describe the hybrid ontology approach to represent geotechnical knowledge. In order to allow cooperation between experts, we propose a framework enable to create mappings. Finally, we conclude with some perspectives.

2 RELATED WORKS

In **OntoDB** project, each information source builds its local ontology from the concepts and relations contained in a global domain ontology. So, semantic integration is automatic. This research work utilizes a strong hypothesis: A database administrator defines a relevant ontology and he adds the subsumption relations existing between his local ontology and the global ontology (Nguyen Xuan, 2006).

OWSICS architecture involves two ontology levels:

Ziani M., Boulanger D. and Talens G..
BUSINESS ONTOLOGIES COOPERATION.
DOI: 10.5220/0003651802140219
In *Proceedings of the International Conference on Knowledge Engineering and Ontology Development* (KEOD-2011), pages 214-219
ISBN: 978-989-8425-80-5

The local information sources and the cooperation (global). Information sources are semantically explained with corresponding local ontologies. At the global level reference ontology describes the domain semantics. A semi-automatic method has been developed to allow the creation of mappings, but only between a local ontology and the reference ontology (Abrouk, 2008).

Except OntoDB, cooperation systems need ontology alignment. Then, we focus on several tools to align ontologies. Most of them use terminological, conceptual or extensional similarity measures and combine them according to an aggregation strategy. They differ in their functioning and the interactions they offer to the users.

PROMPT is a computer-aided system for comparing, merging, aligning and managing ontologies. Its alignment and merge module, called Anchor PROMPT, allows the expert, to find mapping in the following way: (i) the system calculates terminological measures to determine an initial set of similar concepts, (ii) from this list, an algorithm analyzes paths in the sub-graphs bounded by these concepts and indicates which classes frequently appear in the same positions on similar paths (Noy, 2001).

OLA (OWL Lite Alignment) is a system implementing an algorithm of ontology alignment written in OWL. It measures the similarity between two entities from the similarity between their characteristics (classes, properties, relations with the other entities…). The final similarity is the weighted sum of these similarities. The weights are associated relatively to the type of entities to be compared. The algorithm uses a fixed point method with iterations to improve the similarity of two entities. When there are no possible improvements, alignments between two ontologies are proposed (Euzenat, 2004).

AROMA (Rule Ontology Matching Approach association) is an approach of alignment of ontologies represented in OWL. It allows discovering semantic links (subsumption or equivalence between two entities: Classes or properties). There are three steps in the process of alignment: The first one consists in the acquisition of “relevant terms” for each concept and its ancestors. These terms, contained in the descriptions and instances of ontology entities, are extracted with tools of Natural Language Processing. The second step allows creating relations of subsumption between the ontology entities from rules of association. In the final step, the system analyzes the relations previously obtained in order to: (i) deduce the relations of equivalence; (ii) find inconsistences and eliminate them (iii) delete the redundant relations; (iv) select the best alignment for every entity (David, 2009).

More recently, frameworks have appeared in the ontology alignment systems. They allow multiple combinations of strategies to calculate the similarity. For example:

COMA++ (*COmbining MAtching*) is a generic system of matching schemas (Do, 2002; Massmannn 2006). This framework allows the importation, storage, and edition of schemas and produces alignment algorithms in order to transform or merge those schemas. It provides an extensible library of alignments, a module for the combination of the results and a platform to estimate the various measures. The user can interact during the matching process by selecting the measures aggregation strategy (the average, the weighted sum,…).

MAFRA (*MApping FRAmework for distributed ontologies*) is an interactive framework, dynamic and progressive, for the alignment of ontologies distributed through the semantic web (Mädche, 2002). The steps of the MAFRA alignment process are: (i) Importation and standardization of ontologies to align, (ii) similarities to compute between elements of different ontologies from a combination of similarity measures, (iii) formalization of mappings by establishing "semantic bridges" between the entities of different ontologies, (iv) execution of mappings to transform the instances of source ontology to the instances of a target ontology based on semantic bridges, and finally the results' verification.

FOAM (*Framework for Ontology Alignment and Mapping*) is a framework used in several systems for data integration, ontology merging, and ontology evolutions... The tool implements several measures and strategies for similarities research and allows to create mappings between ontologies described in OWL. The general process of alignment is as follows: One selects the pairs of entities to be compared and the characteristics on which to perform the comparison. The system calculates a similarity for each pair and for each characteristic. These results are combined to obtain the final similarity between each pair of entities. From these results, FOAM forwards a set of suggestions for alignment, to be validated by the users (Ehring, 2007).

RiMOM (*Risk Minimisation based Ontology Mapping*) is an interactive framework which implements several strategies to align ontologies

(Tang, 2006); (Li, 2009). The RiMOM's alignment process is as follows: The first step consists in selecting the used measures depending on the assumed similarity between the ontologies (terminological or structural). In the second step, several measures are applied. Then, the results are combined using linear interpolation function. The third step is the propagation of the similarities (from concept to concept, from property to property, from concept to property). The final step consists in generating mappings from the results previously obtained. The process is iterative, with a validation of results at each iteration.

Tools and frameworks we presented differ in the measures used and the aggregation strategy. These systems do not guide user, according to the context, to use a particular methods to discover similarity. In addition, these systems do not permit to reuse previously calculated result.

In order to improve the interactions with the users, we design a computer-aided system allowed selected similarity methods according to the characteristics of the concepts to align. The expert is guided in the selection of the methods to be aggregated. Our contribution is to propose a computer-aided system which calculates similarities and discovers semantic relations. Similarities and mappings are stored in databases and can be reused in order to avoid new computation.

3 PRELIMINARY WORKS

Geotechnics is the science studying the grounds according to diverse aspects: Mechanics of grounds, geology, techniques of building…The complexity of geotechnical domain and the knowledge heterogeneity imply the sharing and the management of knowledge to be difficult (Faure, 2007).

To represent knowledge of geotechnical domain, ontologies are a main issue. The existing concepts in the domain are too numerous to be represented in a single ontology (approximately 5000 concepts), so we propose to design an hybrid ontology.

Each group of experts builds an ontology representing concepts, properties and instances of his business: A local ontology. A global ontology is automatically created by the system to represent a shared vocabulary.

Initially, we classify the ontologies by level. The first level corresponds to the ontology sharing the most common concepts with the others. It represents the target ontology. The source ontologies are the other local ontologies.

The concepts and properties of the source ontology involved in the target ontology are integrated to the global ontology in the ascending order of ontology levels.

The result is an ontology represented by a conceptual graph. This one is verified by the means of an algorithm of integration (Ziani, 2008) which allows the deletion of the relations providing cycles in the global ontology. These last ones are stored in the local ontologies. Finally, we obtain a consistent and consensual ontology including the common concepts and properties of local ontologies and not the conflicting relations which connect them.

Let a target ontology "Tunnel" containing the concepts and instances used in tunnel engineering and a source ontology "Project management" containing the concepts and instances used by the experts in project management (cf. figure 1).

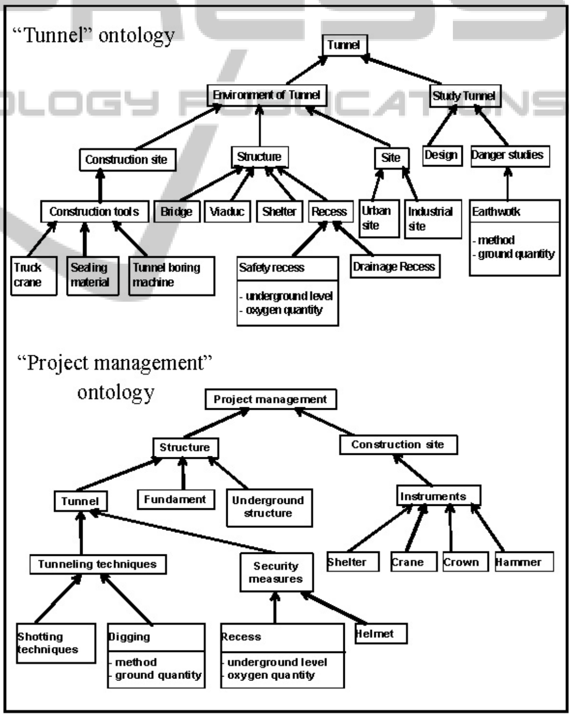

Figure 1: Description of the ontologies "Tunnel" and "Project management".

The integration program allows deleting the relation directed from "structure" to "tunnel" in the global ontology, because there is a cycle between these two concepts and the relation between them is not in the target ontology "Tunnel". This relation is preserved in the source ontology "Project management" (cf. figure 2).

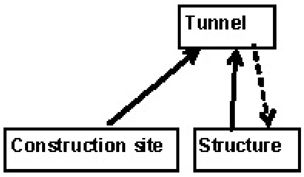

Figure 2: Global ontology.

This representation is possible relatively to the plurality of the geotechnic sub-domains and the existence of a vocabulary shared by all the experts.

The hybrid ontology contains ten business ontologies (currently, we have pointed ten differents sub-domains) and a global one, written in the OWL language.

Each local ontology is represented as a tree, simple to implement and destinated to geotechnical experts. Its size can vary according to the business (between 100 and 1000 concepts).

4 A FRAMEWORK FOR ONTOLOGY ALIGNMENT

To permit cooperation business ontologies, we propose a computer aided system (MOON). It guides the expert in the process of mapping creation between concepts of local ontologies (Ziani, 2011). The architecture of this system is presented in figure 3.

Most of the alignment systems use various measures of similarity to deduce the similarity between two entities. The difficulty is to choose the right measure or the combination of measures to find the similarity. Our system helps the geotechnical expert in the process of similarity research between concepts and generates mappings between them.

When an expert wants to cooperate with another, he sends to the system a request including two ontologies: Departure (corresponding to the business of the expert) and research and, the concept to align (departure concept) (1). The system loads the two local and the global ontologies (2). The objective is to discover the concepts of the research ontology to align with the concept of the departure ontology. Then, the system verifies in the similarity database if there are synonyms for the departure concept (3). Several methods of similarity measures are implemented in the framework and can be proposed to the expert (4). The interest of the framework is to reuse different implemented measures (terminological, conceptual...) and to allow the combination of several similarity measures. The goal of the system is to give to the geotechnical experts several measures and to help them choosing the best ones.

The expert selects the methods and measures (5). The result is a set of similarities between the departure concept and the concepts found in the research ontology. These can be of different types (equivalence, subsumption) and are stored in the similarity database (6). Then, they are proposed to the domain expert (7). The expert can validate or not these similarities. The semantic links proposed and the expert names are stored in the mapping database (8). If the expert knows the concepts to align, he directly stores the relation in the mapping database. This involved the update of the generated mappings (9).

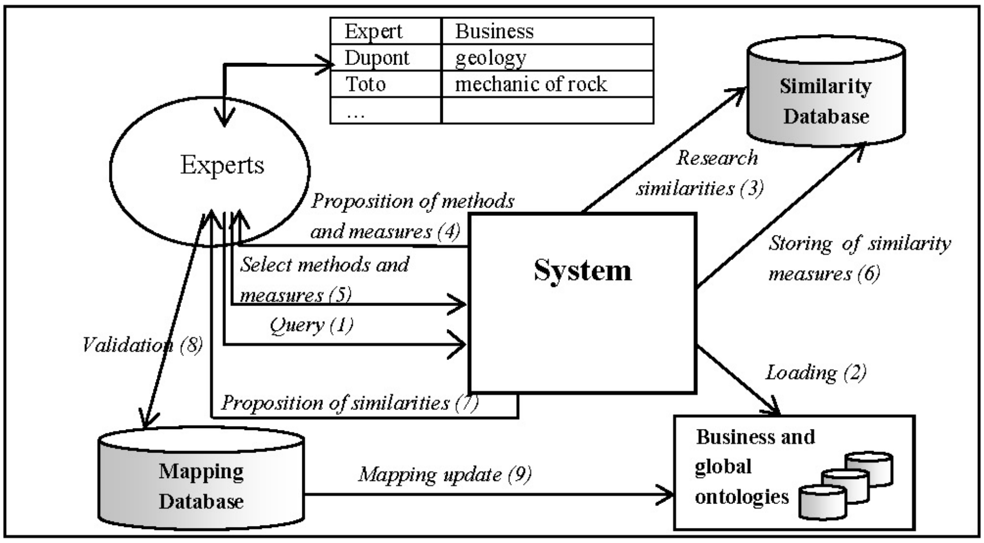

Figure 3: Architecture of the MOON system.

When an alignment, between two concepts, is proposed, the system researches in the mapping database a relation between the same concepts. If there is no relation between the concepts, this one is automatically generated in the ontologies. Otherwise, there are two possibilities: Either the same alignment exists, in this case there is no modification to be brought to the ontologies, or there is a contradictory alignment: In this case, we cannot create this latter. The alignment previously created is deleted. It can be recreated only by a third expert who confirms one of the existing solutions or by an expert who modifies the alignment which he has previously proposed in the mapping database.

The process to guide the experts to choose similarity measures is explained in the following paragraphs:

At first, the system calculates the terminological similarity measures between the departure concept and all the concepts of the research ontology. When the terminological similarity measure gives a result, it means the compared concepts are lexically similar.

After, the system verifies if the concepts are semantically similar from the global ontology. It researches and compares the smallest subsuming concept of the departure concept and the found concept which exists in the global ontology. If they are identical, they are considered as potentially similar. On the contrary, if a link of subsumption in the global ontology exists, the system deduces a similarity between them. But if they are no relation between them in the global ontology, the system deduces that the compared concepts are not semantically similar.

Then, if the departure concept contains at least two attributes, the system proposes to calculate the similarity measure based on the concept properties. This measure gives the concepts in ontology research which have common attributes with the departure concept.

Then, the system proposes a method based on the hierarchical structure of ontology: Counting the edges. It consists in (i) researching the smallest subsuming concepts of the departure which exists in the global ontology, (ii) calculating the number of edges between this common concept and the departure concept, (iii) selecting all the concepts in the research ontology from the level N-2 of the common concept until the level N+2, (iv) suggesting to the expert the found concepts if their number is not very important (subordinate or equal to 10).

In addition, to improve these results, the system proposes extensional methods: Two classes are similar if they share a subset of instances for attributes chosen by the expert.

Suppose that the “Project management” ontology has to cooperate with “Tunnel” ontology. The discovery of mappings through the system for the concepts “digging” and “crane” gives the following results:

For the concept "digging", the terminological similarity measures between concepts do not give result. The system calculates the similarity measure based on the attributes. It finds the same attributes in the concept “digging” and the concept “earthwork” containing in the “Tunnel” ontology: “method” and “ground quantity”. Therefore, it suggests to the expert to create the relations between them.

For the concept “crane”, the terminological similarity measures find the concept “truck crane”. Similarity based on the attributes is not proposed because the concept “crane” contains less than 2 attributes. Consequently, the system proposes the method of the edge counting: It researches the smallest concept subsuming this concept and existing in the global ontology. It finds “construction site”. There are 2 edges separating these two concepts, so it selects all the concepts in the “Tunnel” ontology from the level 0 (level of the concept “construction site”) to 4 (4^{th} level from the concept “construction site”). They find the concepts: “construction site”, “construction tools”, “truck crane”, “sealing material” and “tunnel boring machine”. The number of these concepts is less than 10, the system proposes them as similar concepts with priority to the concept “truck crane” finding earlier.

This work offers a guide to an expert in the process of creating mappings. All the result found must be stored in the similarity and mapping databases.

5 CONCLUSIONS

In this article we have presented an approach for the representation of heterogeneous, distributed knowledge and the cooperation between experts of various businesses. We applied this work to the geotechnical domain.

An hybrid ontology is designed: In the first time, the business ontologies are conceptualized by a consensus between several experts in the same sub-domain. At the end, the global ontology is automatically built from these local ontologies by an integration approach. To allow the cooperation between the experts, we have implemented a prototype with a module to guide experts to align ontologies. In current systems the measure scheduling is fixed; on the contrary, in our proposition, the system selects the measures relatively to the ontology characteristics.

A similarity database stores all the calculated similarities and a mapping database containing all the relations validated by the experts. All the stored measures can be reused to avoid new computations and so not to perform the process.

Currently, the system of the geotechnical knowledge management is partially implemented.

The local ontologies and the created mappings between concepts evolve. They imply modifications in the global ontology and the generated mappings. So, the first perspective of this work is to analyze the consequences of the hybrid ontology evolution and to propose some solutions to maintain the consistence of all the ontologies (local and global). There are diverse systems which manage the ontology evolution (Stojanovic, 2004; Jaziri, 2010; Djedidi, 2010). Our future contribution will manage an hybrid ontology evolution.

The second perspective is to estimate all the mappings stored in the similarity database. The interest is to deduce other semantic relations.

Finally, the third perspective is to study the scalability of the hybrid ontology and the alignments between concepts.

REFERENCES

Abrouk L., Cullot N., Ghawi R., Gomez Carpio G-V., Poulain T., 2008. Cooperation of information sources in OWSCIS System, In *CGCT'2008.*

David J., 2009. AROMA results for OAEI 2009, In *OM'2009, 4th ISWC workshop on ontology matching*, p.147-152, 2009.

Djedidi R., Aufaure M-A. *ONTO-EVO^{A}L* an Ontology Evolution Approach Guided by Pattern Modeling and Quality Evaluation, In *FoIKS'2010,* p. 286-305.

Do H., Rahm E, 2002. COMA – a system for flexible combination of schema matching approaches, In *VLDB'02, 28th International Conference on Very Large Data Bases*, p. 610-621.

Ehring M., 2007. Ontology Alignment: Bridging the Semantic Gap: Semantic Web and Beyond, *Springer.*

Euzenat J., Valtchev P., 2004. Similarity-based ontology alignment in OWL-lite, In *ECAI'04, 15th European Conference on Artificial Intelligence*, p. 333-337.

Faure N., 2007. Un système d'aide à la modélisation des connaissances en géotechnique, *Thèse de doctorat en informatique, Université Jean Moulin – Lyon3.*

Jaziri W., Sassi N., Gargouri F., 2010. Approach and tool to evolve ontology and maintain its coherence, *Metadata, Semantic and Ontologies*, vol. 5, n°2, p. 151-166.

Li J., Tang J., Luo Q., 2009. RiMOM: A Dynamic Multistrategy Ontology Alignment Framework, *IEEE Transactions on Knowledge and Data Engineering*, vol. 21, n°8, p. 1218-1232.

Mädche A., Motik B., Silva N., Volz R., 2002. MAFRA-a MApping FRAmework for distributed ontologies, In *EKAW'02, International Conference on Knowledge Engineering and Knowledge Management*, Springer, p. 235-250.

Massmann S., Engmann D., Rahm E., 2006. COMA++: Results for the Ontology Alignment Contest OAEI, In *OM'2006, 4th ISWC workshop on ontology matching.*

Nguyen Xuan D., Bellatreche L., Pierra G., 2006. A Versioning Management Model for Ontology-Based Data Warehouses», *DaWaK* 2006, vol. 4081, 2006, p. 195-206.

Noy N., Musen M., 2001. Anchor PROMPT: Using non-local context for semantic matching, In *IJCAI'01, Workshop on Ontology and Information Sharing,* p. 63-70.

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

Tang J., Li J., Liang B., Huang X., Li Y., Wang K., 2006. Using bayesian decision for ontology mapping, *Journal of Web Semantics*, vol. 4, n° 4, p.243-262.

Visser U., Stuckenschmidt H., Schlieder C., Wache H., Timm I., 2002. Terminology Integration for the management of distributed information resources, *Kuenstliche Intelligenz Journa*l, vol. 16, p. 31-34.

Ziani M., Boulanger D., Talens G., 2008. Designing an Hybrid ontology From Domain Ontologies, In *JFO'08, 2^{e} Journées Francophones sur les Ontologies*, Lyon, p. 41-47.

Ziani M., Boulanger D., Talens G., 2011. Système d'aide à l'alignement d'ontologies métier–Application au domaine de la géotechnique, *revue Ingénierie des systèmes d'informations*, vol. 16, n°1, p. 89-112.