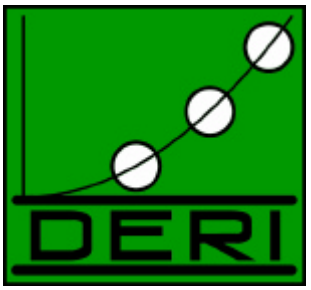

Deliverable:
D18

Version:
0.2

Research Topic Ontology

Semantic Web Portal Project

Lars Ludwig

DERI Technical Report 2004-09-27

September 2004

DERI Galway
University Road
Galway
IRELAND
www.deri.ie

DERI Innsbruck
Technikerstrasse 13
A-6020 Innsbruck
AUSTRIA
www.deri.at

INTRODUCTION

The Semantic Web Portal working group has developed and continues to develop a **research portal ontology**. This research portal ontology shall enable DERI to publish and document its research activities in a Semantic Web Portal. While the work on the portal ontology has seen considerable progress during the last months, an important part of it is still missing: the research topics.

So far **research topics** are seen as mere string value attributes of concepts such as 'publication', thereby reducing the possibilities of describing complex and networked research contents.

Value attributes would soon lead to a de-normalization of research topics described, because many documents and activities relate to the same research topics, which would then be stored again and again as more or less similar string value properties of instances. In the end, however, research topics might become one of the important access keys to information stored in DERI's Semantic Web Portal.

Research topics would best be separated into a special **domain ontology** to increase reusability in other.

A **radical position of knowledge management** would even maintain that a research topic ontology could hold by itself all the information of DERI's research activities. Instead of using such an ontology to annotate research results stored in documents annotated by the portal ontology, it could itself be the store for knowledge about the research domain.

However, for practical reasons, we will focus on **a solution in line with the current development path** of the Semantic Web Portal working group.

The domain ontology proposed is different from the portal ontology in that the former might easily become **subject of more heated debates on how concepts and instances representing theoretical ideas are to be defined and related to each other**, while the latter seems to be more or less a mere reflection of the visible world expressible in general concrete concepts - such as 'person' or 'organization' - and their specific sub-concepts.

The processes of the portal ontology creation have followed a rather **classical ontological engineering method**. In short, given suitable ontologies were integrated, new ontology concepts were proposed by chosen experts, and discussions were stimulated to reach sufficient consensus. Even though new concepts and proposals to improve the portal ontology will arise in future, the ontology engineering can be centralized and be taken care of by only few people.

The research topic ontology would be different in this respect as most members of DERI are active in a quite **fluent research process**. Moreover, one often finds that many people at DERI hold **different views on research topics**. To express it in another way: all research fellows of DERI will become ontology engineers in the creation of a research domain ontology.

Of course, it would be quite easy to just decide to leave the development of the research topic ontology to a few selected people, or to restrict its scope considerably as done in the current portal ontology. One might argue that even the portal ontology will be engineered by all

members of DERI, because instances belonging to the ontology are to be created via the portal, for example in annotations; and research topics can be defined in instances instead of concepts as well, thus **enabling a core ontological engineering process and a following controlled extension and development of the research topic knowledge base on instance level**.

In Knowledge Management, ontology updates cannot be delegated to a specialised work force. This idea is an ivory-tower view comparable to delegating data-updates to database administrators. We here come across a general problem of ontology usage in knowledge management. The 'minimal ontological commitment' and 'consensual conceptualization' demands do not suffice the representation of an area of fluent and heterogeneous individual knowledge creation. However, most of the application fields of ontologies will be knowledge management of one kind or another.

METHODOLOGY

An appropriate way of implementing a research domain ontology is here proposed by a **two-step process**. First, a **traditional ontological engineering** methodology is to be used to construct a basis ontology schema for research topics. Second, the concepts from the ontology schema developed are being applied in **communities of interest** in form of instances.

In the first step an **advocate for each community of interest** should contribute to the development of the ontology schema in order to later on more easily intermediate between centralised engineering of the ontology schema and de-centralised evolving user requirements.

Communities of interest might comprise for example:

1. Semantic Web Portal (Technology)
2. Semantic Web Services
3. Semantic Knowledge and Innovation Management
4. Ontology Management
5. Ontology Languages

This list is not to be understood as a complete suggestion, but rather as a rough first draft to be revised and completed.

For the first step, it is recommended to follow the **Methondology**[1] ontological engineering methodology. This methodology describes in its **conceptualization process**, which is of central importance to us, 8 action steps relevant here:

1. Build glossary of terms
2. Build concept taxonomy
3. Build ad hoc binary relation diagrams
4. Build concept dictionary
5. Describe ad hoc binary relations
6. Describe instance attributes (class attributes that can be instantiated)
7. Describe class attributes

[1] Fernández, M., Gomez-Perez A. and Jurista N. Methondology: From ontological art towards ontological Engeneering. In Spring Symposium Series, Standford University, 1997

8. Describe constants (constant table)

The conceptualization is **followed by processes of formalization, implementation, and maintenance**. The conceptualization should be **preceded, if applicable, by an evaluation of related ontologies** at hand.

Each conceptualization activity would deliver a **working result** which can be made subject to a **working group refinement session**. Such a session would at the same time serve to **coordinate the ontological engineering efforts** of the advocates to ensure consistency and provide necessary interdependencies.

To build **a pre-formalized representation of the concept taxonomy,** conceptualization activity 2, the author suggests using **Artificial Memory** as a Knowledge Management System for shared development and visualization. By the **Web-Interface** of Artificial Memory in combination with its **flexible browsing features** we would be able to have telephone conferences between members located in Innsbruck and those located in Galway instead of personal meetings when refining and coordinating the individual efforts of the subject matter advocates. A (yet to be implemented) RDF-export of the research topic ontology schema could aid the formalization efforts to be made after conceptualization later on.

As it is might well be found that the Semantic Web Portal working group is unable to cover all the subject areas of relevance in a single engineering cycle, a **cyclic approach** might be needed. The explicit decision for a cyclic approach could mean that, to make things easier and to soon **benefit from a learning cycle**, the working group could restrict itself in the beginning to a **limited number of subject areas** and communities of interest respectively.

The author recommends dealing soon with a technical issue that arose in the context of the development of the Portal Ontology: **In OWL light research terms cannot be implemented as concepts** as would usually be done. This is because in OWL light and OWL DL instances cannot directly refer to concepts. Instead, concepts would have to be implemented on the instance level mimicking a concept hierarchy, accompanied or not accompanied by a corresponding concept hierarchy. The issue arising now is that **due to the common level of appearance the difference between abstract classes and classes serving to create instances cannot be easily implemented and thereby enforced**. Thus, if the Portal user had the means to create instances, which is in itself beneficial for the scenario of a research portal, the process could not be controlled to the degree used to in traditional ontology management.

PROPOSAL OF ACTIVITIES

Following find an exemplary **proposal of activities** related to the first 2 activities of the Methondology conceptualization phase. The activities proposed cumulate into a feedback session of all working group members and a final draft of a term glossary and concept taxonomy regarding the first ontological engineering cycle.

Further activities will be defined as soon as there is an agreement on the general path of progress in order to avoid work that might easily have to be abolished later on in the process.

0. Destine advocates for each community of interest or a choice of communities of interest

1. Build glossary of terms
    a. Separately research for relevant ontologies at hand
    b. Separately collect potential cross-community or top-level concepts (such as 'Semantic Web' or 'methodology')
    c. Consolidation of these concepts by central ontology engineer (avoid duplicates)
    d. Separately collect research area / community-relevant concepts
    e. Consolidation of these concepts by central ontology engineer (avoid duplicates)
    f. Feedback to advocates regarding consolidation / intersections and joint glossary
    g. Possibly feedback from advocates between each other regarding *missing* terms
    h. Integration of *missing* terms into joint glossary by central ontology engineer
2. Further development of activity plan by Lars Ludwig
3. Build concept taxonomy (identify disjoint concept sets, strive for exhaustive decomposition)
    a. Preparation of ontology schema to model taxonomy in Artificial Memory
    b. [Session] Jointly build common concept taxonomy for top-level concepts
        i. generate common understanding on taxonomy-modelling by paradigmatic modelling of parts of one research area
    c. Separate building of research area-related concept taxonomies using a single Artificial Memory database by all advocates
    d. Joint session of advocates representing, discussing, and refining individually developed concept taxonomy-parts
    e. Presentation and explanation of the concept taxonomy to the other Semantic Portal working group members
    f. Integration of Feedback of Semantic Portal working group members regarding concept glossary and concept taxonomy
    g. Publishing of results