ONTHOLOGY - An Ontology Metadata Repository

Jens Hartmann
Institute AIFB, University of Karlsruhe, Germany
hartmann@aifb.uni-karlsruhe.de

Abstract

Most ontologies today exist in pure form without any additional information, e.g. authorship or domain specific information. The proposed Ontology Metadata Vocabulary (OMV) aims to establish a standard which enables users from academia and industry to identify, find and apply – basically meaning to reuse – ontologies effectively and efficiently. Our demo/poster contribution consists of the presentation of the up-and-running metadata portal ONTHOLOGY ("anthology of ontologies") which implements the proposed OMV to support users in accessing and reusing of ontologies. OMV is available for download at http://omv.ontoware.org/, ONTHOLOGY is available at http://www.onthology.org/.

1. Introduction

Ontologies are commonly used for a shared means of communication between computers and between humans and computers. To reach this aim, ontologies should be represented, described, exchanged, shared and accessed based on open standards such as the W3C standardized web ontology language OWL. However most ontologies today exist in a pure form without any additional information about authorship, domain of interest and other meta data about ontologies. Ssearching and identifying existing ontologies which are potentially reusable because they e.g. are applied in similar domains, used within similar applications or who have similar properties is a rather hard and tedious task.

We argue that metadata in the sense of machine processable information for the Web[1] helps to improve accessibility and reuse ontologies. Further, it can provide other useful resource information to support maintenance. Thus we claim that metadata not only help when applied (or, attached) to documents, but also to ontologies themselves.

As a consequence, ontologies which are annotated by metadata require an appropriate technology infrastructure as well. This includes tools and metadata repositories which comply to the ontology metadata standard and which provide the required functionalities to support reuse of ontologies. Such tools and repositories typically should support the engineering process, maintenance and distribution of ontologies.

In this paper we present the up-and-running portal infrastructure ONTHOLOGY as reference implementation which shows the benefit of applying such standard in a centralized scenario. The main functionality of the portal is to store, manage and making accessible ontology meta data for large user communities.

2. Ontology Metadata Vocabulary (OMV)

The presented metadata portal stores information according the metadata vocabulary OMV which has been proposed as metadata standard in the EU IST thematic network of excellence Knowledge Web[2].

OMV distinguishes between an **ontology conceptualisation** and an **ontology implementation**. This separation is based on following observation: any existing ontology (implementation) has some kind of *core idea* (conceptualisation) behind. From an ontology engineering perspective, a person initially develops such a *core idea* of what should be modeled in his mind. Further, this initial conceptualisation might be discussed with other persons and after all, an ontology will be *realised* using an ontology editor and stored in a specific format. Over time there might be created several *realisations* of this initial *cconceptualisation* in many different formats, e.g. in RDF(S)[3] or OWL[4].

The distinction between an ontology conceptualisation and ontology implementation leads to an efficient mechanism, e.g. for tracking several versions and evolvements of ontologies as well as for different representations of one knowledge model (conceptualisation) in different ontology languages. Such an *ontology conceptualisation* can be seen as representation of the conceptual model behind an ontology.

Besides these two main classes, additional classes are re-

[1] http://www.w3.org/Metadata/

[2] http://knowledgeweb.semanticweb.org

[3] http://www.w3.org/RDF/

[4] http://www.w3.org/TR/owl-features/

quired to represent useful information about ontologies by such vocabulary. Therefore OMV provides further classes and properties representing *environmental information* and *relations*, e.g. such as *persons*, *engineering tools* or even *license models*. The complete metadata ontology is illustrated in [1].

3. ONTHOLOGY — Ontology Metadata Repository

As the importance of metadata increases with the number of existing ontologies, the demand for a supporting technologies like storage and access techniques becomes important as well. We present the conceptual design of a centralised ontology metadata repository and its implementation, so-called ONTHOLOGY standing for "anthology of ontologies".

Centralised systems allow to reflect long-term community processes in which some ontologies become well accepted for a domain or community and others become less important. Such well accepted ontologies and in particular their metadata need to be stored in a central metadata portal which can be accessed easily by a large number of users whereby the management procedures are well defined. A main goal of a centralised metadata portal is to act as large evidence storage of metadata resp. their related ontologies to facilitate access, reuse and sharing as required for the Semantic Web.

We identified several different user roles for ONTHOLOGY: The *visitor* is an anonymous user, he is allowed to browse the public content of the portal. A Visitor can become a *user* by completing an application form on the website. In order to avoid unnecessary administrative work, a user is added automatically to the membership database. Users can customize their portal, e.g. the content of their start-page or their bookmarks. If a user wants to submit metadata to the portal, this submission has to be reviewed before it is published. ONTHOLOGY establishs a *review process* in order to ensure a certain level of quality. *Reviewers* check the new submissions before it is published. The *technical administrator* is responsible for any other task mainly the maintenance of the portal.

Functionalities of ONTHOLOGY can be separated into two groups based on the usage. Indeed, *basic functionalities* which are provided to every user who accesses the portal and *sophisticated functionalities* for reviewers and administrators. The main operations a user can perform on the repository are (i) *Search*, (ii) *Submit* and (iii) *Export*.

The search and export can be performed by any visitor without being registered to the repository. Since providing new metadata is based on a certain community confidence, a visitor has to register at the portal to be become a registered user.

A metadata portal mainly consists of *a large data repository* in which metadata can be stored. Exemplary, Sesame[5] or KAON[6] can be used as back-end metadata repository. *Access* and in particular the *managament* of the repository must be guaranteed, too. Therefore, ONTHOLOGY is based on SEAL, the AIFB conceptual architecture for building SEmantic portALs. In SEAL ontologies are key elements for managing community web sites and web portals. They support queries to multiple sources, but beyond that also intensive use of the schema information itself to allow for automatic generation of navigational views such as navigation hierarchies that appear as `has- part`-trees or `has- subtopic` trees in the ontology. In addition to that mixed ontology and content-based presentation is supported. Further information can be found at [2].

In addition to the central storage and maintenance, ONTHOLOGY cooperates with the decentralised system Oyster[7] which stores and retrieves metadata in a P2P manner. The benefit of connecting both systems lies mainly in the simple reuse of existing ontology metadata information from such networks of users who are willing to share them. Whereas the portal is expected to contain data which matures according to quality insurance procedures over time, the ad-hoc P2P network enables quick and easy distribution of data without much control. In combination, both systems ensure efficient and effective ontology metadata management for various use cases.

4. Conclusion

To conclude, reusing existing ontologies is a key issue for sharing knowledge on the Semantic Web. Our contribution aims at facilitating access and reuse of ontologies which are previously unknown for ontology developers and users through the ONTHOLOGY metadata portal. As metadata standard we use the Ontology Metadata Vocabulary (OMV). Next steps include the standardization of OMV on a wider scope and the development of further extensions to ONTHOLOGY, in particular the linking of ONTHOLOGY with Oyster requires additional efforts.

5. References

[1] J. Hartmann and R. Palma. OMV - Ontology Metadata Vocabulary for the Semantic Web, 2005. v. 1.0, available at `http://ontoware.org/projects/omv/`.

[2] J. Hartmann and Y. Sure. An infrastructure for scalable, reliable semantic portals. *IEEE Intelligent Systems*, 19(3):58–65, May/June 2004.

[5] `http://www.openrdf.org/`
[6] `http://kaon.semanticweb.org/`
[7] `http://ontoware.org/projects/oyster`