Oyster – Sharing and Re-using Ontologies in a Peer-to-Peer Community

Raul Palma
Facultad de Informática. Universidad Politécnica de Madrid
Campus Montegancedo, s/n 28860
Boadilla del Monte. Madrid. Spain
34-913367439
rpalma@fi.upm.es

Peter Haase
Institute AIFB
Universität Karlsruhe (TH)
D-76128 Karlsruhe, Germany
49-7216083705
haase@aifb.uni-karlsruhe.de

Asunción Gómez-Pérez
Facultad de Informática. Universidad Politécnica de Madrid
Campus Montegancedo, s/n 28860
Boadilla del Monte. Madrid. Spain
34-913367439
asun@fi.upm.es

ABSTRACT
In this poster, we present Oyster, a Peer-to-Peer system for exchanging ontology metadata among communities in the Semantic Web. Oyster exploits semantic web techniques in data representation, query formulation and query result presentation to provide an online solution for sharing ontologies, thus assisting researchers in re-using existing ontologies.

Categories and Subject Descriptors
H.3.4 [**Information Storage and Retrieval**]: Systems and Software—*Information networks*
K.6.4 [**System Management**]: Centralization/decentralization

General Terms
Management, documentation, design, reliability

Keywords
Ontology, Peer-to-Peer, Repository, Metadata

1. INTRODUCTION
Currently efficient knowledge sharing and reuse is rather difficult, as it is hard to find and share ontologies available among the community due to the lack of standards for documenting and annotating ontologies with metadata information. This raises the problem of having many isolated ontologies created by many different parties. Besides the costs of the duplicate efforts, this also hampers interoperability between ontology-based applications. Oyster[1] is a Peer-to-Peer application that exploits semantic web techniques in order to provide a solution for exchanging and re-using ontologies. To achieve this, Oyster implements a proposal for a metadata standard, called Ontology Metadata Vocabulary (OMV)[2] [2] which is based on discussions and agreements carried out in the EU IST thematic network of excellence Knowledge Web[3] as a way to describe ontologies. The decentralized approach provides an ideal solution for users that require a repository to which they have full access and can perform any operation without any consequences to other users. For example, users from academia or industry might use a personal repository for a task dependent investigation, or ontology engineers, might use it during their ontology development process to capture information about different ontology versions. We argue that a decentralized system is the technique of choice, since it allows the maximum of individuality while it still ensures exchange with other users. A centralized approach, on the other hand, allows reflecting long-term community processes in which some ontologies become well accepted for a domain or community and others become less important. However, both approaches could be combined to cover a variety of use cases.

Demos and Posters of the 3rd European Semantic Web Conference (ESWC 2006), Budva Montenegro, 11th - 14th June, 2006

2. OYSTER
Oyster provides an innovative solution for sharing and re-using knowledge (i.e. ontologies), which is a crucial step to enable Semantic Web. The Oyster system has been implemented as an instance of the Swapster system architecture[4]. In Oyster, ontologies are used extensively in order to provide its main functions (importing data, formulating queries, routing queries, and processing answers).

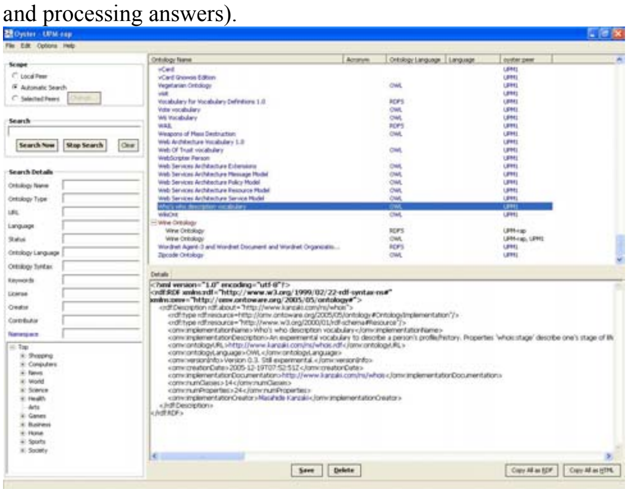

Figure 1. Oyster Screenshot

Creating and Importing Metadata: Oyster enables users to create metadata about ontologies manually and also to import ontology files in order to automatically extract the ontology metadata available and let the user to fill in the missing values. For the automatic extraction, Oyster supports the OWL[5], DAML+OIL[6], and RDF-S[7] ontology languages. The ontology metadata entries are aligned and formally represented according to two ontologies: (1) the proposal for a metadata standard OMV

[1] Available at http://oyster.ontoware.org/
[2] More information at http://omv.ontoware.org/
[3] http://knowledgeweb.semanticweb.org/
[4] http://swap.semanticweb.org/
[5] http://www.w3.org/TR/owl-guide/
[6] http://www.w3.org/TR/daml+oil-reference
[7] http://www.w3.org/TR/rdf-schema

that describes the properties of the ontology, and (2) a topic hierarchy (i.e. DMOZ[8]) that describes specific categories of subjects to define the domain of the ontology.

Formulating Queries: Users can search the repository for ontologies by means of simple keyword searches, or more advanced, semantic searches (c.f. the left pane of figure 1). Queries are formulated in terms of these two ontologies. This means that queries can refer to fields like name, acronym, ontology language, etc. or they may refer to topic terms.

Routing Queries: As shown in the upper left pane of figure 1, users may query a single specific peer (e.g. their own computer, or a certain peer because this peer is known as a big provider of information), or they may query a specific set of peers (e.g. all the members of a specific organization), or they may query the entire network of peers (e.g. when the user has no idea where to search). In the latter case, queries are routed automatically through the network depending on the expertise of the peers, describing which topic of the topic hierarchy a peer is knowledgeable about. In order to achieve this expertise based routing, a matching function determines how closely the semantic content of a query matches the expertise of a peer [1].

Processing results: The results matching a query are presented in a result list (c.f. upper right pane in figure 1). The answer of a query might be very large and may contain many duplicates due to the distributed nature and potentially large size of the Peer-to-Peer network. Such duplicates might not be exact copies because of the semi structured nature of the metadata, so the ontologies are used again to measure the semantic similarity between different answers and to remove apparent duplicates. Then a merged representation that combines the knowledge from the individual and potentially incomplete items is presented to the user. Details of the particular results are shown in the lower right side of Figure 1. Users can save the results of a query into their local repository for future use.

3. OMV IN OYSTER

Oyster provides an infrastructure for storing, sharing and finding ontologies making use of the proposal for a metadata standard OMV. OMV comprises the OMV Core, which captures information relevant to most of the ontology reuse settings and various OMV Extensions that allow ontology developers/users to specify task/application-specific ontology-related information (e.g. ontology merging, alignment or versioning, evaluation of ontologies or ontological engineering methodologies). These extensions should be compatible to the OMV core, but at the same time fulfill the requirements of a domain, task or community-driven setting. The OMV elements are classified according to the type and purpose of the contained information such as availability (e.g. URI, URL), provenance (e.g. creator, contributor), applicability (e.g. domain, ontology type), relationship (e.g. import, backward compatibility), format (e.g. ontology language and syntax), statistics (e.g. number of classes or properties) and general information (e.g. name, description). Furthermore, OMV classifies elements according to their impact on the prospected reusability of the described ontology content as required, optional, and extensional. OMV also models additional classes and properties required to support the reuse of ontologies, especially in the context of the Semantic Web, such as Party, Organisation, Person, LicenseModel, OntologyLanguage, OntologySyntax and OntologyTask,. For a complete description of OMV please refer to [2].

4. RELATED WORK

A closely related application is the **Onthology**[9] central repository, which also exploits the OMV. Onthology offers a complementary application to Oyster as both applications have a different usage perspective: Oyster as a decentralized system is the technique of choice for users who needs the maximum of individuality while still ensuring exchange with other users with up-to-date information. Ontology as a centralized system allows reflecting long-term community processes in which some ontologies become well accepted for a domain or community and others become less important. There exists similar approaches to our proposed solution, but in general their scope is quite limited. E.g. the **DAML ontology library**[10] provides a catalog of DAML ontologies that can be browsed by different properties. The **FIPA ontology service**[11] defines an agent wrapper of open knowledge base connectivity. The Semantic Web search engine **SWOOGLE**[12] makes use of particularly metadata which can be extracted automatically. Finally the **SchemaWeb Directory**[13] is a repository for RDF schemas expressed in RDFS, OWL and DAML+OIL.

5. CONCLUSIONS AND FUTURE WORK

To conclude, the reuse of existing ontologies within communities is a key issue for sharing knowledge on the Semantic Web. This task, however, is rather difficult because of the heterogeneity, distribution and diverse ownership of the ontologies as well as the lack of sufficient metadata. As we summarized in this paper, our contribution, Oyster, addresses exactly these challenges by implementing a proposed standard for metadata for describing ontologies. Oyster is already being applied in the KnowledgeWeb project which has partners across the European Union. Oyster is ranked as the number one in the list of top downloaded projects of Ontoware[14] (650 downloads, including all versions and releases). Currently, there are around 250 ontologies shared in Oyster network. We are in the process of collecting usage statistics. Finally, our future work includes addressing many challenges like the integration of Oyster with central repository, handle change propagation, evaluation of expertise ranking, using trust information and evaluation of performance.

6. ACKNOWLEDGMENTS

Our thanks to our partners from the EU projects Knowledge Web (FP6-507482) and NeOn (FP6-27595) for their present and future collaboration.

7. REFERENCES

[1] Haase P. et al. Peer selection in peer-to-peer networks with semantic topologies. In proc of ICNSW'04, Paris, June 2004.

[2] Hartmann, J., Palma, R.OMV-Ontology Metadata Vocabulary for the semantic web.2005. http://omv.ontoware.org

---

8 http://dmoz.org/

9 http://www.onthology.org/

10 http://www.daml.org/ontologies

11 http://www.fipa.org/specs/fipa00086/XC00086C.html

12 http://swoogle.umbc.edu/

13 http://www.schemaweb.info/

14 http://ontoware.org/