Querying Heterogeneous and Distributed Learning Object Repositories via Ontology-Based Mediation[1]

Philipp Kärger*, Erica Melis+, Carsten Ullrich+
*kaphi@activemath.org, +{melis|cullrich}@dfki.de
*Saarland University, +German Research Center for Artificial Intelligence

1. Introduction

We propose a mediator architecture that allows a learning system to retrieve learning objects (LOs) from heterogeneous repositories of LOs and not exclusively from its proprietary repository. A mediating component accepts queries formulated in a uniform query language, translates them into repository specific queries and passes them to each connected repository. For the translation of queries, an ontology-based query-rewriting method has been developed which uses the specified knowledge representation of the repository and an ontology mapping to compute the rewriting steps for translating. For distributing queries we realized a mediating architecture, which passes the translated queries to the repositories. This architecture has been integrated in the Web-based, user-adaptive and interactive e-learning environment ActiveMath. Currently, it enables the ActiveMath's course planner to access four heterogeneous LO repositories. The course planning component can create courses adapted to the learner's goals, knowledge, learning behaviour and a specific learning scenario [1]. To select adequate LOs, the course planner initially queried its proprietary repository MBase for LOs. The challenge our architecture faces is to integrate more than one repository into ActiveMath's course generation. Currently, with the help of the architecture here-proposed, the course planner retrieves LOs additionally from the repository of the DaMiT-system, of the MathsThesaurus, and the LeActiveMath ExerciseRepository.

2. Mediator Approach

The mediator provides a single interface accepting queries which contain a metadata specification of LOs and returns the Uniform Resource Identifiers (URIs) of the LOs meeting this specification. The query language we developed is based on the *Ontology of Instructional Objects* (OIO) introduced in [2]. A query asking for all LOs which are easy exercises training the concept `asymptote` looks as follows:

```
(relation isFor asymptote²)
(class Exercise)(property hasDifficulty easy).
```

The OIO-terms in the queries (underlined in the example) have to be replaced with the corresponding terms a repository uses to describe its LOs. To guarantee the correct substitution of terms one has to define an ontology mapping between the OIO (as the source ontology) and each of the pedagogical ontologies describing the metadata structure of the target repositories. Therefore, to define an accurate mapping, one has to make explicit and specify an ontology representing the knowledge structure and metadata semantics of each target repository. We developed an XML-based ontology mapping language where an ontology mapping comprises a set of *mapping pattern*s (see Fig. 1 as an example). Each mapping pattern consists of a *matching pattern* and a set of *replacement patterns* expressing the semantical containment between the concept specified in the matching pattern and the ones in the corresponding replacement patterns. The *query expansion* guarantees that the mediator, if asked for category C, returns not only objects belonging to C but objects belonging to subcategories of C, too.

Our architecture is a mediation information system architecture as introduced by Wiederhold in [3]. The advantage of a mediating architecture is that the querying component does not have to know the specification of the data sources and their query languages [3].

For each repository a wrapper is integrated comprising the specification of the ontology of the repositories knowledge (as an OWL-ontology-definition) and the mapping between the terms of the OIO and the terms the repository uses (see Fig. 2). The mediator utilizes the OWL file for query expansion and the mapping specification for query rewriting.

[1]This research was carried out in the LeActiveMath project, funded under the 6th Framework Program of the European Community. The authors are solely responsible for its content.

[2]Typically, identifiers of learning objects are URIs. For readability we use simple terms throughout this paper instead

```
<OIOMapping>
  …
  <MappingPattern>
    <MatchPattern>
      <ClassRestriction name="Example"/>
      <RelationRestriction name="isFor"/>
    </MatchPattern>
    <ReplacementPattern>
      <RelationRestriction
               name="example_for"/>
    </ReplacementPattern>
  </MappingPattern>
  …
</OIOMapping>
```

Figure 1. *Extract of an ontology mapping*

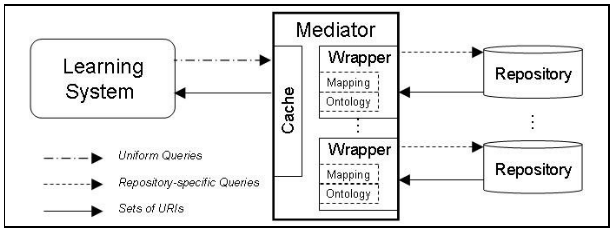

Figure 2. *Mediator Architecture*

A repository is not always equipped with a powerful caching mechanism. Hence we integrated a caching mechanism into the mediating component.

Four repositories have been integrated into the mediator architecture: 1. **ActiveMathMBase** (`www.mathweb.org/mbase`), a knowledge base for mathematical LOs enriched with metadata. 2. **DaMiT** (`damit.dfki.de`), an adaptive tutoring system imparting knowledge about Data Mining. 3. **MathsThesaurus** (`thesaurus.maths.org`), an online multilingual Mathematics Thesaurus. 4. The **LeActiveMath ExerciseRepository**, a database of interactive exercises (`mathdox.org/repository`). Four repository-specific ontology mappings have been specified to rewrite the queries from ActiveMath's course planner. Using the mediator's cache, the architecture made the course planning more than ten times faster and therefore more usable. The course planner can now generate personalized courses from each of the repositories, and also from several repositories in case the repositories use the same URI for the same concepts.

3. Related Work

Several approaches for data integration by ontology mapping as well as for federation (i.e., the reuse and exchange) of LOs exist. Edutella [4] is a Peer-To-Peer approach for sharing information of the semantic web. Regarding ontology mapping, expressive mapping languages have been developed (e.g., [5]) but not yet been implemented. Therefore, we decided to develop a query rewriting specific approach which is less powerful but expressive enough for our translation purposes.

4. Conclusion and Further Work

We propose an ontology-based mediation approach which enables an e-learning system to query several heterogeneous LO repositories. After analysing and representing the knowledge representation and the metadata structure of a repository it is easy to use the repository as a service by specifying a wrapper which calls the retrieval facilities of the new repository. This technique was successfully implemented and tested for the e-learning environment ActiveMath and four repositories.

Further improvements are planned for the mediator architecture. To ease the creation of ontology mappings an ontology mapping editor would be useful. Another planned improvement is to increase the expressiveness of the query language to enable more complex search facilities on repositories (such as text search, provision of disjunctions, etc.).

Our approach focuses on mapping of concepts and not on mapping of instances. To provide an architecture which allows the generation of mixed courses comprising LOs from different repositories, we are planning to integrate an instance mapping technology basing on domain ontologies and mappings between them. Note that our architecture does not solve the problem of LO *presentation*: we are managing heterogeneous metadata standards, heterogeneous knowledge representation and heterogeneous storing technologies.

5. References

[1] Ullrich, C.: Course generation based on HTN Planning. *Proceedings of 13th Annual Workshop of the SIG Adaptivity and User Modeling in Interactive Systems.* (2005) 74–79

[2] Ullrich, C.: The learning-resource-type is dead, long live the learning-resource-type! *Learning Objects and Learning Designs* 1(1) (2005) 7–15

[3] Wiederhold, G.: Mediators in the architeture of future information systems. *The IEEE Computer Magazine* (1992)

[4] Nejdl, W., Wolf, B., Qu, C., Decker, S., Sintek, M., Naeve, A., Nilsson, M., Palmer, M., Risch, T.: Edutella: A P2P Networking Infrastructure Based on RDF. In: *Proceedings of WWW2002.* (2002)

[5] de Bruijn, J., Foxvog, D., Zimmermann, K.: Ontology Mediation Patterns Library V1. D4.3.1, SEKT-project (2005)