SEMANTIC WEB AND PEER-TO-PEER TECHNOLOGIES FOR DISTRIBUTED LEARNING REPOSITORIES

Wolfgang Nejdl

*Computer Science Department, Stanford University, nejdl@db.stanford.edu**

Abstract

Metadata for the World Wide Web is important, but metadata for Peer-to-Peer (P2P) networks is absolutely crucial. In this paper we discuss the open source project Edutella and related projects, which combine semantic web and peer-to-peer technologies in order to make distributed learning repositories possible and useful. We describe the main services of the Edutella network infrastructure and its architecture based on the exchange of RDF metadata, starting with the query service as one of the core services of Edutella. We discuss the Edutella Common Data Model (ECDM) as basis for the Edutella query exchange language (RDF-QEL-i) implementing distributed queries over the Edutella network, and discuss as one example an O-Telos-Peer with native Datalog query and inference capabilities. We then sketch a new P2P routing topology which minimizes broadcast traffic and distance between the peers in such a network, as well as a modification language needed for supporting distributed update and annotation. A short section discusses the changing role of adaptation in these distributed repositories (as open corpus hypermedia instead of the closed ones investigated in most research projects so far). Finally, we discuss an application of the Edutella network for digital libraries (OAI-P2P) and the use of Edutella in the ELENA project which aims to create smart spaces for learning.

1. Introduction

While in the server/client-based environment of the World Wide Web metadata are useful and important, for *Peer-to-Peer (P2P)* environments metadata are absolutely crucial. Information Resources in P2P networks are no longer organized in hypertext like structures, which can be navigated, but are stored on numerous peers waiting to be queried for these resources if we know what we want to retrieve and which peer is able to provide that information. Query-

---

*on leave from the University of Hannover, Germany

ing peers requires metadata describing the resources managed by these peers, which is easy to provide for specialized cases, but non-trivial for general applications.

P2P applications have been successful for special cases like exchanging music files. However, retrieving a song like "Material Girl from Madonna" does not need complex query languages nor complex metadata, so special purpose formats for these P2P applications have been sufficient. In other scenarios, like exchanging educational resources, queries are more complex, and have to build upon standards like IEEE-LOM/IMS [IEEE-LTSC, 2001, IMS, 2001] metadata with up to 100 metadata entries, which might even be complemented by domain specific extensions.

Furthermore, by concentrating on domain specific formats, current P2P implementations appear to be fragmenting into niche markets instead of developing unifying mechanisms for future P2P applications. There is indeed a great danger (as already discussed in [Dornfest and Brickley, 2001]), that unifying interfaces and protocols introduced by the World Wide Web get lost in the forthcoming P2P arena.

Edutella Infrastructure.   The Edutella project [Edutella, 2002, Nejdl et al., 2002c, Nejdl et al., 2002b] addresses these shortcomings of current P2P applications by building on the W3C metadata standard RDF [Lassila and Swick, 1999, Brickley and Guha, 2000]. The project is a multi-staged effort to scope, specify, architect and implement an RDF-based metadata infrastructure for P2P-networks based on the recently announced JXTA framework [Gong, 2001]. The initial Edutella services will be *Query Service* (standardized query and retrieval of RDF metadata), *Replication Service* (providing data persistence / availability and workload balancing while maintaining data integrity and consistency), *Mapping Service* (translating between different metadata vocabularies to enable interoperability between different peers), *Mediation Service* (define views that join data from different metadata sources and reconcile conflicting and overlapping information) and *Annotation Service* (annotate materials stored anywhere within the Edutella Network).

The Edutella infrastructure aims to provide the metadata services needed to enable interoperability between heterogeneous JXTA applications. The main application area we discuss within this paper are P2P networks for the exchange of educational resources (using schemas like IEEE LOM, IMS, and ADL SCORM [ADL, 2001] to describe course materials), other application areas are possible as well.

As the query service is one of the core services in Edutella, upon which other services are built, we specify in Section 2 the Edutella Common Data Model (ECDM) as basis for the Edutella query exchange language and format (see [Nejdl et al., 2002c, Nejdl et al., 2002b]). Section 3 discusses as one example

an O-Telos-Peer [Wolpers et al., 2002], built upon the ConceptBase system [Jarke et al., 1995], which shares with RDF(S) a very similar data model, and provides native Datalog query and inference capabilities (see also [Nejdl et al., 2001]). In section 4 we sketch a new P2P routing topology, called HyperCuP, which is described in [Schlosser et al., 2002]. The HyperCuP topology minimizes broadcast traffic and distance between the peers in such a network, allowing it to scale to large numbers of peers. Section 5 discusses annotation and modification in the Edutella network, described in [Nejdl et al., 2002c, Nejdl et al., 2002b] and [Nejdl et al., 2002a]. Section 6 discusses the changing role of adaptation in these distributed repositories (as open corpus hypermedia instead of the closed hypermedia systems investigated in most research projects so far), see also [Henze and Nejdl, 2002]. Finally, we discuss an application of the Edutella network for digital libraries (OAI-P2P) [Ahlborn, 2002] and the use of Edutella in the EU/IST project ELENA which aims to create smart spaces for learning.

2.   Edutella Query Service

The Edutella Query Service is a standardized query exchange mechanism for RDF metadata stored in distributed RDF repositories and serves both as query interface for individual RDF repositories located at single Edutella peers as well as query interface for distributed queries spanning multiple RDF repositories (storing RDF statements based on arbitrary RDFS schemata).

One of the main purposes is to abstract from various possible RDF storage layer query languages (e.g., SQL) and from different user level query languages (e.g., RQL, TRIPLE): The Edutella Query Exchange Language and the Edutella Common Data Model provide the syntax and semantics for an overall standard query interface across heterogeneous peer repositories for any kind of RDF metadata. The Edutella network uses the query exchange language family RDF-QEL-i (based on Datalog semantics and subsets thereof) as standardized query exchange language format which is transmitted in an RDF/XML-format.

We will start with a simple RDF knowledge base and a simple query with the following RDF XML Serialization:

```
<lib:Book about="http://www.xyz.com/sw.html">
  <dc:title>Software Engineering</dc:title>
</lib:Book>

<lib:Book about="http://www.xyz.com/ai.html">
  <dc:title>Artificial Intelligence</dc:title>
</lib:Book>

<lib:AI-Book about="http://www.xyz.com/pl.html">
  <dc:title>Prolog</dc:title>
</lib:AI-Book>
```

Evaluating the following query (plain English)

"Return all resources that are a book having the title 'Artificial Intelligence' or that are an AI book."

we get the query results shown in Figure 1, depicted as RDF-graph.
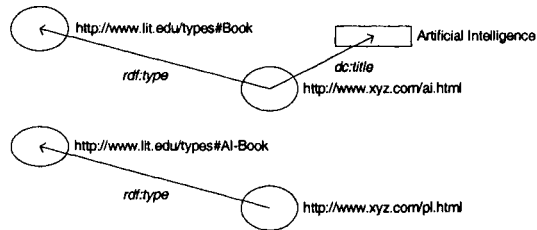

Figure 1.   Query Results as RDF Graph

Edutella peers can be highly heterogeneous in terms of the functionality (i.e., services) they offer. A simple peer has RDF storage capability only. The peer has some kind of local storage for RDF triples (e.g., a relational database) as well as some kind of local query language (e.g., SQL). In addition the peer might offer more complex services such as annotation, mediation or mapping.

To enable the peer to participate in the Edutella network, Edutella wrappers are used to translate queries and results from the Edutella query and result exchange format to the local format of the peer and vice versa, and to connect the peer to the Edutella network by a JXTA-based P2P library. To handle queries, the wrapper uses the common Edutella query exchange format and data model for query and result representation. For communication with the Edutella network the wrapper translates the local data model into the Edutella Common Data Model ECDM and vice versa, and connects to the Edutella Network using the JXTA P2P primitives, transmitting the queries based on ECDM in RDF/XML form.

In order to handle different query capabilities, Edutella defines several RDF-QEL-i exchange language levels, describing which kind of queries a peer can handle (conjunctive queries, relational algebra, transitive closure, etc.) The same internal data model is used for all levels.

Edutella Common Data Model (ECDM).   The ECDM is based on Datalog, which is a well-known non-procedural query language based on Horn clauses without function symbols. A Datalog program can be expressed as a set of rules/implications (where each rule consists of one positive literal in the consequent of the rule (the head), and one or more negative literals in the antecedent of the rule (the body)), a set of facts (single positive literals) and the actual query literals (a rule without head, i.e., one or more negative literals).

Literals are predicates expressions describing relations between any combination of variables and constants such as title(http://www.xyz.com/book.html, 'Artificial Intelligence'). Disjunction is expressed as a set of rules with identical head. Additionally, we can use negation as failure in the antecedent of a rule, with the semantics that such a literal cannot be proved from the knowledge base. A Datalog query then is a conjunction of query literals plus a possibly empty set of rules [Silberschatz et al., 2001].

Datalog queries easily map to relations and relational query languages like relational algebra or SQL. In terms of relational algebra Datalog is capable of expressing selection, union, join and projection and hence is a relationally complete query language. Additional features include transitive closure and other recursive definitions.

The example knowledge base in Datalog reads

```
title(http://www.xyz.com/ai.html,'Artificial
                              Intelligence').
type(http://www.xyz.com/ai.html,Book).
title(http://www.xyz.com/sw.html,'Software
                              Engineering').
type(http://www.xyz.com/sw.html,Book).
title(http://www.xyz.com/pl.html,'Prolog').
type(http://www.xyz.com/pl.html,AI-Book).
```

Each RDF repository can be viewed as a set of ground assertions either using binary predicates as shown above, or as ternary statements "s(S,P,O)", if we include the predicate as an additional argument. In the following examples, we use the binary surface representation.

Example Query in (binary) Datalog notation.

```
aibook(X)  :- title(X, 'Artificial Intelligence'),
             type(X, Book).
aibook(X)  :- type(X, AI-Book).
?- aibook(X).
```

Since our query is a disjunction of two (purely conjunctive) subqueries, its Datalog representation is composed of two rules with identical heads. The literals in the rules' bodies directly reflect RDF statements with their subjects being the variable X and their objects being bound to constant values such as 'Artificial Intelligence'. Literals used in the head of rules denote derived predicates (not necessarily binary ones). The query expression "aibook(X)" asks for all bindings of X, which conform to the given Datalog rules and our knowledge base, with the results:

```
aibook(http://www.xyz.com/ai.html)
aibook(http://www.xyz.com/pl.html)
```

Internally Edutella Peers use a Datalog based model to represent queries and their results. Figure 2 visualizes this data model as UML class diagram.
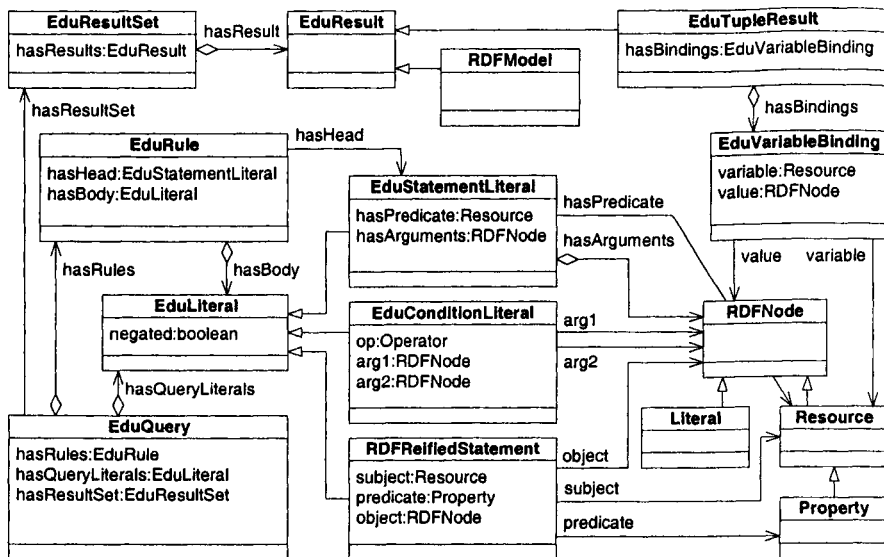
Figure 2.   Edutella Common Data Model (ECDM)

The Edutella Wrapper API includes the ECDM as well as wrappers for different query languages, and is available as source code from the Edutella Project Page[1]. We are currently extending the ECDM with two important features from the TRIPLE language [Sintek and Decker, 2002], namely model IDs (to explicitly distinguish between different models/peers) and function symbols (useful to construct new IDs for models, resources and for complex values expressed as terms instead of strings).

Our current prototype environment features a set of different peers to demonstrate various aspects of the translation from ECDM to local query languages. It contains the QEL query exchange mechanism, a simple mediator and the wrapping of different repository peer types, including an OLR (Open Learning Repository) based peer [Dhraief et al., 2001] using a subset of IMS/LOM RDF metadata stored in a relational database, a DbXML-based peer [Qu and Nejdl, 2001] as a prototype for an XML repository using a simple mapping service to translate from RDF-QEL-1 queries (conjunctive queries) to Xpath queries over the appropriate XML-LOM schema, AMOS-II-based peers [Risch and Josifovski, 2001] with local repositories, KAON-based peers [Maedche et al., 2002] allowing remote annotation [Handschuh et al., 2001] using an RDF-based ontology format, and an O-Telos-Peer [Jarke et al., 1995] and [Wolpers et al., 2002], which we describe in section 3.

The environment also supports the design and integration of other tools which make use of RDF metadata. The Ont-O-Mat editor and annotation tool

is described shortly in section 5 and more detailed in [Nejdl et al., 2002c], the concept browser Conzilla [Nilsson and Palmér, 1999] allows us to use a graphical query representation for QEL-1 and QEL-2.

3. Advanced Datalog Capabilities in an Edutella-O-Telos-Peer

O-Telos, ConceptBase and RDF. Once we have introduced Datalog as a query language, we can also use it for general inference purposes or to express integrity constraints on RDF data. Basic RDF is very similar to a binary relational model (if we view properties as predicates with two arguments), which is extended by RDF schema into an object oriented model (using classes and class hierarchies, instantiation of classes, as well as domains and ranges for properties). A very similar scheme has been used for the object-oriented meta modeling language O-Telos [Kramer et al., 1991] [Mylopoulos et al., 1990], as we have discussed in a previous paper [Nejdl et al., 2001] in more detail, where we have described an RDF variant called O-Telos-RDF, which extends RDF using statement IDs and explicit instantiation not only of classes but also of properties, and provides better reification capabilities and more powerful meta-modeling functionality compared to RDF(S).

The O-Telos language has been implemented in a system called Concept-Base [Jarke et al., 1995], which is a deductive object-oriented database, very useful as a repository for modeling and storing metadata (e.g. [Jeusfeld et al., 1998]). ConceptBase implements a powerful query and reasoning mechanism (rules and constraints) based on (stratified) Datalog and uses the Lloyd-Topor transformation [Lloyd and Topor, 1984] to allow arbitrary first order logic formulas in the body of rules.

Our O-Telos-Peer provides advanced query capabilities (up to RDF-QEL-5) for RDF data stored in or imported into the ConceptBase subsystem of the O-Telos-Peer. In order to store the RDF metadata in ConceptBase the peer has to translate RDF to O-Telos. We have developed a translation based on the axioms from O-Telos-RDF [Nejdl et al., 2001], which is quite straightforward and stresses the similarity of the two formalisms RDF(S) and O-Telos/O-Telos-RDF, with differences in notation and serialization (triples vs. quadruples and XML vs. frame syntax). Using this translation we can import arbitrary RDF(S) data into ConceptBase and then use the reasoning capabilities of ConceptBase for querying, inferencing, and integrity checking on RDF(S) data.

Query Classes. RDF queries are posed to the O-Telos-Peer using RDF-QEL queries as described in section 2. RDF-QEL queries are translated into O-Telos queries using the interesting feature of O-Telos Query Classes as specified in [Jarke et al., 1995]. O-Telos results are then translated to either

a RDF graph result in XML serialization or the RDF triple format. In our book example we stored the RDF description of the resource with the URL http://www.xyz.com/jv.html and the title "Just Java". Now suppose we want to know from the database which books have the title "Just Java", which is depicted in RDF-QEL-1 and XML as follows:

```
<?xml version='1.0' encoding='ISO-8859-1'?>
<!DOCTYPE rdf:RDF
 [ <!ENTITY a 'http://purl.org/dc/elements/1.1/'>
   <!ENTITY rdf 'http://www.w3.org/1999/02/22-rdf-syntax-ns#'>
   <!ENTITY b 'http://www.edutella.org/edutella#'>
   <!ENTITY c 'http://www.w3.org/2000/01/rdf-schema#'>
 ]>
<rdf:RDF xmlns:a="&a;" xmlns:rdf="&rdf;" xmlns:b="&b;"
         xmlns:c="&c;" >
  <b:QEL1Query rdf:about="#genQuery">
    <b:hasVariable rdf:resource="#X"/>
    <b:hasVariable rdf:resource="#Y"/>
    <b:hasResultType rdf:resource="&b;GraphResult"/>
  </b:QEL1Query>
  <b:Variable rdf:about="X" a:title="Just Java" c:label="X">
    <a:title rdf:resource="#Y"/>
  </b:Variable>
  <b:Variable rdf:about="Y" c:label="Y"/>
</rdf:RDF>
```

The query asks for at least one resource in the database which has a property called dc:title with a value "Just Java". The resulting resource(s) are returned as RDF graph in XML notation. The peer constructs from the RDF query representation the respective O-Telos Query Class, taking the ECDM representation of the query as input. In our example the corresponding query class defines the attributes title and namespace, the constraint varX declares that all instances of the answer set must have an attribute named title with the value "Just Java".

```
QueryClass EduQuery isA Individual with
    retrieved_attribute
        title : String;
        namespace : String
    constraint
        varX : $ exists Y/Individual A(this,title,Y)
                      and A(this,title,"Just Java") $
end
```

The answer consists of RDF statements declaring that there is a resource with URL http://www.xyz.com/jv.html. This resource is the domain of a property named title which has the range "Just Java". The title property is defined in the Dublin Core namespace.

```
jv_html in EduQuery with
    title
        titl : "Just Java"
    namespace
        naml : "http://www.xyz.com/"
end
```

4.   HyperCuP Peer-to-Peer Routing

Obviously, routing in P2P networks is crucial for scaling up the network. Most P2P networks evolve in an unorganized manner, and are prone to suffer from serious scalability problems, limiting the number of nodes in the network, creating network overload and pushing search times to unacceptable limits. In [Schlosser et al., 2002] we describe the HyperCuP topology for P2P networks, which organizes the peers into a deterministic graph topology featuring symmetry (i.e. every node can equally originate searches and broadcasts), low network diameter, a limited node degree (of logarithmic order to the amount of peers in the network) and load balancing of traffic in the network.
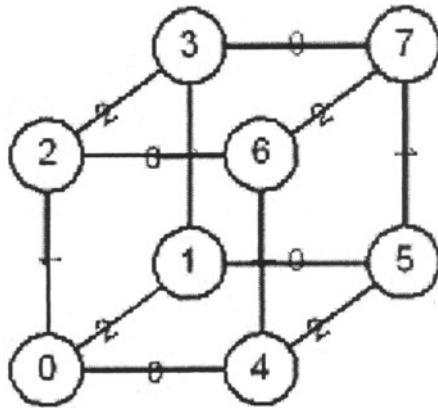

Figure 3.   A Hypercube with Dimension 3

We arrive at this by constructing a hypercube with all available peers, as shown in figure 3. To construct a hypercube of base 2 and dimension d, for each dimension the set of nodes is split into two partitions 0 and 1, which are defined by the d-th digit (0 or 1) in the binary encoding of the node number (from 0 to 2^d-1), giving d partitions/dimensions for the hypercube and thus d neighbors for each node. Broadcasting on this topology guarantees that each node receives messages exactly once, featuring a network diameter logarithmic to the amount of nodes. Searching, i.e. broadcast along only a limited number of nodes, is efficient for the same reasons. [Schlosser et al., 2002] also provides a construction and maintenance algorithm which permits nodes to join and

leave the network arbitrarily (also for number of nodes not equal to 2^d), with a low message overhead (logarithmic to the amount of nodes in the network), and not requiring any central servers or supernodes.

5.   Edutella Annotation and Modification

Annotation Service.   In order to easily provide metadata for a particular document, the Edutella annotation service provides a document viewer and a browser for RDF schemata. Using the browser, a corresponding definition, e.g. Dublin Core, is loaded into the annotation tool to be browsed. Fields for annotation are displayed according to the schema definition and may either be filled by typing or by marking and dragging information from the document viewer. The Edutella annotation service is composed of the Edutella peer libraries and the KAON tool-suite[2] [Maedche et al., 2002] incorporating the Ont-O-Mat Plugin Framework[3] and Annotation application [Handschuh et al., 2001] (cf. Figure 4).
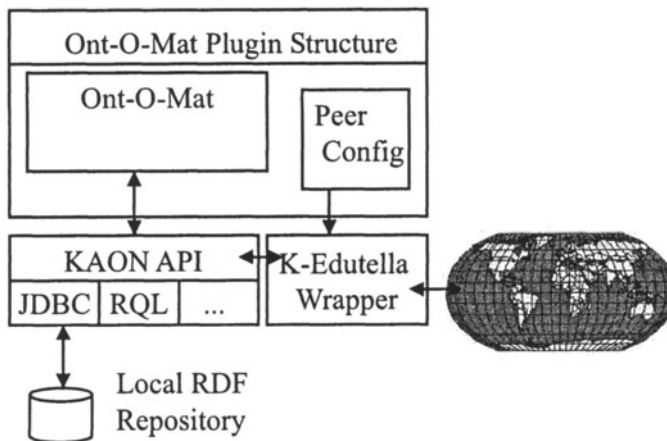

Figure 4.   Ont-O-Mat as Edutella Peer

KAON is a Semantic Web tool suite originally created in isolation of Edutella. The Ont-O-Mat Framework is part of this tool suite and provides a java-based plugin structure which allows for loading services dynamically. One such service is the annotation tool Ont-O-Mat which uses the KAON API to query for RDF schema definitions in order to build up its ontology browser. It queries for instances, attributes and relationships in order to let its users explore the current state of the knowledge base.

The Edutella Wrapper for KAON (K-Edutella Wrapper) is again a KAON plugin. The task of the K-Edutella Wrapper is to wrap the KAON-API for QEL and vice versa. The K-Edutella Wrapper calls JXTA lower levels for services

like registration, pipes, etc. in order to connect to the Edutella network. Using Ont-O-Mat, a user can directly connect the Edutella network in order to query metadata from other peers or provide metadata from his repository.

Modification and Update Language. For this annotation service and also for additional replication functionality between distributed peers, we need a standardized mechanism to communicate metadata changes. In [Nejdl et al., 2002a] we present a basic language designed for communicating metadata changes between distributed RDF repositories, designed for use cases such as the one depicted in figure 5.
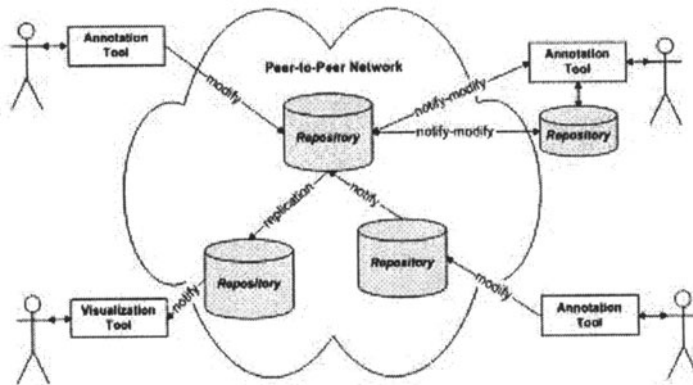

Figure 5. Annotation / Replication Use Case

Our proposal specifies a modification exchange language with subgraph-centered granularity (in contrast to statement- or resource-centered requests), where an RDF query language can deliver the query result as a subgraph of the repository. Therefore we can design modification commands as a combination of a query to specify the affected statements and a specification of the changes to these statements. In this context, an update consists of a query specifying the changed statement(s) and the description of the new statement(s). The repository can then change the selected statements accordingly. This approach can also handle variables in the modification specifications, supports change patterns, enables replacing the object part of a statement without knowing its actual value, and integrates nicely with our query exchange language QEL, which is used to specify the subgraph selection. The language uses modification messages as atomic transaction unit, which may contain multiple commands. The repository can process such a message in one chunk, thereby allowing a stateless modification/update protocol.

6.   Adaptation in Open Corpus Hypermedia

Another important topic for learning materials is personalization to different users. Adaptive hypermedia systems have a longer tradition than the World Wide Web, and have been based on general hypermedia systems, which can be described as a set of nodes (containing multiple forms of media such as text, video, audio, graphics, etc.) connected by links [Rada, 1995]). Adaptive hypermedia systems generally include two forms of adaptation: content level and link level adaptation. Using content level adaptation, the document's content is tailored to the needs of a specific user, for example by hiding specialized information or by inserting additional explanations. Using link level adaptation, the user's possibilities to navigate the hypermedia documents are personalized. Techniques for content level and link level adaptation are described e.g. in [Brusilovsky, 1996].

The World Wide Web is now leading to a new class of adaptive hypermedia systems which we call "adaptive open corpus hypermedia" [Henze and Nejdl, 2002]. The hypermedia repository is no longer a closed system managing only local data, but potentially extends over different sites, using metadata to describe the structure and content of the distributed resources (see also [Henze, 2000, Dhraief et al., 2001]). This is useful especially for learning resources, as authors can add arbitrary learning objects to an open learning repository. However, this often leads to an oversupply of information, and link level adaptation becomes a central issue. For content level adaptation the metadata of a specific learning object must deliver information to redesign the presentation of the learning object itself.

Standardization.   Even though personalization is treated in several standardization working groups, metadata standardization efforts for user-oriented data are still in early stages, compared to standardization efforts for learning object metadata (LOM) and course structure metadata (IMS CP) incorporated in the current SCORM 1.2 standard. Interesting in this context is the IEEE LTSC Learner Model Working group P1484.2, which works on standards for exchanging learner profiles, covering the information dimensions "personal, "relations", "security", "preference", "performance" and "portfolio". Important extensions to the current status of these metadata proposals are in the area of learning processes of individual users or learning styles, which are crucial parts of a learner model and will probably play a role for the development of a standard for course sequencing (IEEE P1484.6 Course Sequencing, no proposed standard so far).

Knowledge Models.   User knowledge is the main source of information for adaptation in educational hypermedia. Based on the user's actual knowl-

edge state, information is selected, examples are proposed, reading sequences are generated and appropriate learning steps are presented to the user.

A knowledge model for OAHS has to be *universally valid* (throughout the different corpora of the learning repository), *expandable* (learning objects on new topics may be added to one of the corpora) and *expressive* (the knowledge contained in one single corpus must be describable as a valid model on its own). Ontologies as formal explicit specifications of shared conceptualizations [Gruber, 1993] meet these requirements and can therefore be used within a knowledge model to describe the domain to be learned by the student. Concepts in the knowledge model then define a *controlled vocabulary* for describing the knowledge of the application domain and can be used for metadata annotation of the learning objects (LOM category General, data element 1.6 Keywords). The learning objects themselves are described and identified by a set of keywords describing their content.

Adaptive Functionality as Queries. Furthermore, once the metadata in an OAHS describe distributed learning resources and are represented in RDF, most adaptive functionalities in an OAHS can be represented as RDF-QEL queries, which can be distributed through the Edutella P2P network.

7.   OAI-P2P: P2P Solutions for Digital Libraries

In the digital library community, the Open Archive Initiative (OAI) has defined a metadata harvesting protocol (OAI-PMH) to achieve interoperability among distributed archives, based on HTTP and XML as well as the Dublin Core metadata scheme [Lagoze and de Sompel, 2001]. To keep the instruction set simple, OAI-PMH calls for a separation between data and service providers. Data providers establish an OAI-PMH-based interface to local digital resources, while service providers (like ARC [Liu et al., 2001] and SCIRIUS[4] provide facilities for searching across multiple archives plus value-added features like ranking and unified access to other sources.

This separation exposes the simplicity of the protocol as the source of its' strength (low barrier to adoption) and its' weakness. OAI-PMH is designed as simple as possible for data providers at the expense of service providers, creating and maintaining an OAI-PMH service provider requires much more resources than setting up a data provider. On the other hand, OAI-PMH offers no front-end services. Data providers offer an interface for metadata harvesting to outsiders but do not have any immediate advantages (like a query service for outside repositories) from their efforts, unless a service provider provides an interface to their data.

OAI-P2P.   In essence, the OAI-PMH defines client-server-relationships, with the data provider and the service provider responding to client requests.

However, digital libraries in most cases act both as a client and as a server, at the same time trying to obtain outside material for inside users and offering inside resources to outside users. Similar initial situations have spawned the emergence of services like Napster and Gnutella, which offer resource sharing by means of peer-to-peer structures. [Ahlborn, 2002] describes an organizational and technical framework which merges the OAI-PMH concept with a true peer-to-peer approach (OAI-P2P). It thus takes the OAI-PMH one step further by extending query services to data providers and by avoiding the dependencies of centralized server-based systems.

In an OAI-P2P system, there is no separation between service provider and data provider (each peer maintains separate subsystems for data storage and query handling). Each query is routed to appropriate peers by the network, no administration is necessary to introduce new peers. Of course such a network still benefits from additional service providers which replicate metadata, thereby enhancing the reliability and performance of the net. However, although performance may suffer if such a peer is discontinued, overall communication and services will stay alive even if a single node dies.

Individual digital libraries may want to decide which other repositories they share their data with and which repositories they want to access. In the OAI framework, choosing the scope of the community is the prerogative of the service provider, who may arbitrarily decide which data providers to include. Using the OAI-P2P approach peers can devise community specific access policies using the peer group concept.

Queries.   Merging data provider and service provider functionality means that an OAI-based peer-to-peer digital library network has to address issues outside the OAI-PMH scope. OAI-PMH does not state how data providers should set up source metadata. Although very small archives can use flat files to store XML-metadata, most institutional data providers use a dedicated relational database from which OAI output is created. On the other hand, current OAI service providers replicate the metadata they have harvested in relational databases to provide for clients' queries. In order to give service provider functionality to each data provider in a P2P context, repositories must be able to pose, process and accept advanced queries, an ability which they used already to build an OAI-compliant infrastructure in the first place. Using Edutella's support for a range of query exchange languages (RDF-QEL-i), OAI-P2P is able to adapt to heterogeneous peers as well as changing demand on metadata scope and query complexity.

8.   ELENA: Smart Spaces for Learning

Which steps are coming next? We are currently starting up the EU/IST funded research project ELENA (under the general topic "ambient intelli-

gence"), aimed at creating a smart space for learning. Smart learning spaces in this context are defined as educational service mediators, which allow the consumption of heterogeneous learning services via assessment tools, learning management systems, educational (meta) repositories and live delivery systems such as video conferencing systems or interactive TV (see figure 6). Central design element of the ELENA smart learning space is a dynamic learner profile, which includes learning history, learner specific information and learning goals. Interconnecting learning services in a smart learning space leads to an empowerment of learners since they become capable of choosing among a variety of knowledge sources in order to achieve their personal learning goals.

Figure 6.   ELENA Architecture

References

[ADL, 2001] ADL, T. T. (2001). SCORM Specification v1.2. http://www.adlnet.org/Scorm/scorm.cfm.

[Ahlborn, 2002] Ahlborn, B. (2002). Oai-p2p: A peer-to-peer network for open archives. Technical report, University of Hannover. submitted for publication.

[Brickley and Guha, 2000] Brickley, D. and Guha, R. V. (2000). W3C Resource Description Framework (RDF) Schema Specification. http://www.w3.org/TR/1998/WD-rdf-schema/. W3C Candidate Recommendation.

[Brusilovsky, 1996] Brusilovsky, P. (1996). Methods and techniques of adaptive hypermedia. *User Modeling and User Adapted Interaction*, 6(2-3):87–129.

[Dhraief et al., 2001] Dhraief, H., Nejdl, W., Wolf, B., and Wolpers, M. (2001). Open learning repositories and metadata modeling. In *International Semantic Web Working Symposium (SWWS)*, Stanford, CA.

[Dornfest and Brickley, 2001] Dornfest, R. and Brickley, D. (2001). The power of metadata. http://www.openp2p.com/pub/a/p2p/2001/01/18/metadata.html. excerpted from the book "Peer-to-Peer: Harnessing the Power of Disruptive Technologies.

[Edutella, 2002] Edutella (2002). The Edutella Project. http://edutella.jxta.org/.

[Gong, 2001] Gong, L. (2001). Project JXTA: A technology overview. Technical report, SUN Microsystems. http://www.jxta.org/project/www/docs/TechOverview.pdf.

[Gruber, 1993] Gruber, T. R. (1993). A translation approach to portable ontology specifications. *Knowledge Acquisition*, 5:199–220.

[Handschuh et al., 2001] Handschuh, S., Staab, S., and Maedche, A. (2001). CREAM - creating relational metadata with a component-based, ontology-driven annotation framework. In *Workshop on Knowledge Markup and Semantic Annotation at the First International Conference on Knowledge Capture (K-CAP'2001)*, Victoria, BC, Canada.

[Henze, 2000] Henze, N. (2000). *Adaptive Hyperbooks: Adpatation for Project-Based Learning Resources*. PhD thesis, University of Hannover.

[Henze and Nejdl, 2002] Henze, N. and Nejdl, W. (2002). Knowledge modeling for open adaptive hypermedia. In *Proceedings of the 2nd International Conference on Adaptive Hypermedia and Adaptive Web Based Systems*, Malaga, Spain.

[IEEE-LTSC, 2001] IEEE-LTSC (2001). IEEE LOM Working Draft 6.1. http://ltsc.ieee.org/wg12/index.html.

[IMS, 2001] IMS (2001). IMS Learning Resource Metadata Specification v1.2.1. http://www.imsproject.org/metadata/index.html.

[Jarke et al., 1995] Jarke, M., Gallersdörfer, R., Jeusfeld, M., Staudt, M., and Eherer, S. (1995). ConceptBase - a deductive object base for meta data management. *Journal on Intelligent Information Systems*, 4(2):167 – 192.

[Jeusfeld et al., 1998] Jeusfeld, M. A., Jarke, M., Nissen, H. W., and Staudt, M. (1998). ConceptBase - Managing Conceptual Models about Information Systems. In Bernus, P., Mertins,

K., and Schmidt, G., editors, *Handbook on Architectures of Informations Systems*. Springer Verlag.

[Kramer et al., 1991] Kramer, B. M., Chandhri, V. K., Koubarakis, M., Topaloglon, T., Wang, H., and Mylopoulos, J. (1991). Implementing Telos. *SIGART Bulletin*, 2(3).

[Lagoze and de Sompel, 2001] Lagoze, C. and de Sompel, H. V. (2001). The open archives initiative: Building a low-barrier interoperability framework. In *Proceedings of the ACM/IEEE Joint Conference on Digital Libraries*.

[Lassila and Swick, 1999] Lassila, O. and Swick, R. R. (1999). W3C Resource Description framework (RDF) Model and Syntax Specification. http://www.w3.org/TR/REC-rdf-syntax/. W3C Recommendation.

[Liu et al., 2001] Liu, X., Maly, K., Zubair, M., and Nelson, M. L. (2001). Arc: an oai service provider for cross-archive searching. In *Proceedings of the ACM/IEEE Joint Conference on Digital Libraries*.

[Lloyd and Topor, 1984] Lloyd, J. W. and Topor, R. W. (1984). Making prolog more expressive. *Journal of Logic Programming*, 3:225–240.

[Maedche et al., 2002] Maedche, A., Staab, S., Studer, R., Sure, Y., and Volz, R. (2002). Seal — Tying up information integration and Web site management by ontologies. *IEEE Data Engineering Bulletin*. http://www.research.microsoft.com/research/db/debull/.

[Mylopoulos et al., 1990] Mylopoulos, J., Borgida, A., Jarke, M., and Koubarakis, M. (1990). Telos: A language for representing knowledge about information systems. *ACM Transactions on Information Systems*, 8(4).

[Nejdl et al., 2001] Nejdl, W., Dhraief, H., and Wolpers, M. (2001). O-Telos-RDF: A resource description format with enhanced meta-modeling functionalities based on O-Telos. In *Workshop on Knowledge Markup and Semantic Annotation at the First International Conference on Knowledge Capture (K-CAP'2001)*, Victoria, BC, Canada.

[Nejdl et al., 2002a] Nejdl, W., Siberski, W., Simon, B., and Tane, J. (2002a). Towards a modification exchange language for distributed rdf repositories. In *Proceedings of the 1st International Semantic Web Conference*. Springer LNCS.

[Nejdl et al., 2002b] Nejdl, W., Wolf, B., Qu, C., Decker, S., Sintek, M., Naeve, A., Nilsson, M., Palmér, M., and Risch, T. (2002b). EDUTELLA: a P2P Networking Infrastructure based on RDF. In *Proceedings of the 11th International World Wide Web Conference*, Hawaii, USA. http://edutella.jxta.org/reports/edutella-whitepaper.pdf.

[Nejdl et al., 2002c] Nejdl, W., Wolf, B., Staab, S., and Tane, J. (2002c). Edutella: Searching and annotating resources within an rdf-based p2p network. In *Proceedings of the Semantic Web Workshop, 11th International World Wide Web Conference*, Honolulu, Hawaii, USA.

[Nilsson and Palmér, 1999] Nilsson, M. and Palmér, M. (1999). Conzilla - towards a concept browser. Technical Report CID-53, TRITA-NA-D9911, Department of Numerical Analysis and Computing Science, KTH, Stockholm. http://kmr.nada.kth.se/papers/ConceptualBrowsing/cid_53.pdf.

[Qu and Nejdl, 2001] Qu, C. and Nejdl, W. (2001). Towards interoperability and reusability of learning resources: A SCORM-conformant courseware for computer science education.

Technical report, Learning Lab Lower Saxony, University of Hannover. submitted for publication.

[Rada, 1995] Rada, R. (1995). *Interactive Media.* Springer.

[Risch and Josifovski, 2001] Risch, T. and Josifovski, V. (2001). Distributed data integration by object-oriented mediator servers. *Concurrency and Computation: Practice and Experience,* 13(11):933 – 953.

[Schlosser et al., 2002] Schlosser, M., Sintek, M., Decker, S., and Nejdl, W. (2002). HyperCuP – shaping up peer-to-peer networks. submitted for publication.

[Silberschatz et al., 2001] Silberschatz, A., Korth, H. F., and Sudarshan, S. (2001). *Database Systems Concepts.* McGraw-Hill Higher Education, 4 edition.

[Sintek and Decker, 2002] Sintek, M. and Decker, S. (2002). TRIPLE—An query, inference, and transformation language for the semantic web. In *Proceedings of the 1st International Semantic Web Conference.* Springer LNCS.

[Wolpers et al., 2002] Wolpers, M., Nejdl, W., and Brunkhorst, I. (2002). O-Telos provider peers in RDF-based P2P environments. Technical report, University of Hannover. submitted for publication.