The Edutella P2P Network

- Supporting Democratic E-learning and Communities of Practice

Mikael Nilsson <mini@nada.kth.se>

Knowledge Management Research Group

Royal Institute of Technology, Stockholm, Sweden.

The infrastructures we use for developing, finding and combining learning objects influence the usage of the material - inflexible frameworks will not support flexible learning. For this reason, it is essential to consider the pedagogical consequences of the design of the technical frameworks that are used in e-learning systems. Much of the current work in e-learning technology targets learning objects stored in LMS (Learning Management System) applications and/or in other centralized servers, often of very large scale. Even though standards such as IEEE LOM increase the interoperability of such systems, they are still mostly information islands. Cross-searching of repositories is not a reality. It has even been said that the Web is still in the "hunter-gatherer phase" with respect to searching. This is certainly true for learning objects. We have not yet reached the goal of a global e-learning society. In addition, many institutions are reluctant to give up control over their learning resources. This is problematic for many central-server based methods of learning resource sharing, (e. g., e-learning "portals".) Such portals are costly and difficult to maintain.

Edutella takes a different approach. It is one piece in an e-learning infrastructure with a decentralized vision. By encouraging sharing among small-scale content repositories, anyone can participate in the exchange and annotation of e-learning resources. By allowing anyone to participate, the learner is given more control over their learning process, leading us one step closer to the dream of a learner-centric educational architecture.

Edutella is a peer-to-peer (P2P) network for exchanging information about learning objects (and not for exchanging content). It is built with semantic web technology applying the latest P2P research. This chapter will discuss the technologies that make Edutella possible, explaining the vision and importance of the project, and how applications can use it.

The Edutella project is being developed by a number of institutions - among others: the Learning Lab Lower Saxony, the KMR Group at KTH, the Uppsala Database Laboratory, Stanford Infolab, AIFB at University of Karlsruhe, and the UNIVERSAL project - and it is still expanding. The latest developments can be found at http://edutella.jxta.org.

Edutella Technology

By using a distributed technology, Edutella enables institutions and individuals to actively participate in a global information network, without losing control over their learning resources. Edutella connects highly heterogeneous peers (heterogeneous in uptime, performance, storage size, functionality, number of users, etc.). The goal of the Edutella project is to make the distributed nature of Edutella services (e. g., repository search) completely transparent to Edutella clients.

The first building block of Edutella is an open-source peer-to-peer technology called JXTA-, initiated by Sun Microsystems. JXTA is a generic P2P protocol, designed to be used in many diverse kinds of P2P applications, focused on interoperability, platform independence and ubiquity.

The second building block of Edutella is RDF (Resource Description Framework), which is a framework for representing information in the Web. It has been developed by the World Wide Web Consortium (W3C). The RDF specifications provide a highly sophisticated, lightweight framework for exchanging ontology-based knowledge, containing facilities for combining resource descriptions using different vocabularies and from different sources. It will be seen how the

decentralized nature of the RDF metadata descriptions plays a central role in the Edutella platform.

To show the kinds of queries Edutella manages, consider the following Edutella query, constructed in the Conzilla concept browser:

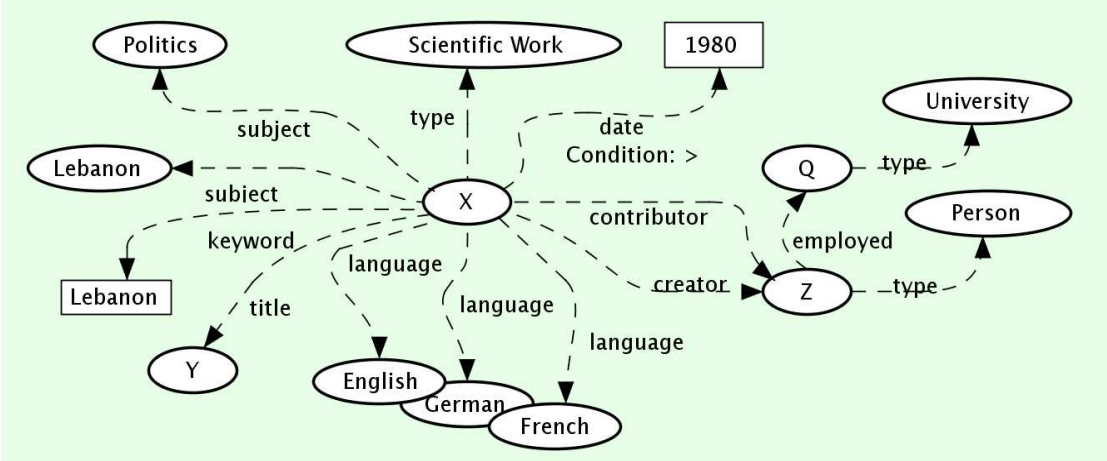

Figure 1. Edutella query in Conzilla

In Figure 1 above, X represents the resource that is searched for. The arcs are properties of that resource. In plain English, the query asks for (counter-clockwise)

scientific works on the subject of politics, having Lebanon as subject or keyword, with a title (Y), written in English, German or French, created or contributed to by a Person (Z), employed at a University, and created after 1980.

(There are several occurrences of "or" in this transcription. However, this information is not explicit in the figure, but is represented separately. See http://www.conzilla.org)

Edutella takes queries of the above complexity, distributes them to peers capable of answering the query, collects the answers and returns them to the originator. It is possible that parts of the answers are located on different peers. In the example, the university employee information is perhaps not located on the same server as the resource metadata. Edutella will be able to handle these kinds of situations transparently.

Note that Edutella deals with metadata about content, not with content itself. Access to educational content is not always as simple as downloading a file - it might include logging in to a web service, starting a certain application with specific parameters, As Edutella uses RDF, each resource must have a URI (Uniform Resource Identifier). The route from that URI to the resource itself is not determined by Edutella. It may be an HTTP URL, so that your Edutella-aware application can point your browser in the right direction. Or it might just be a URN (Uniform Resource Name), uniquely naming the resource but not locating it, and you must go through some sort of lookup service to find it.

Nodes in an Edutella network

Edutella adds a search service to the JXTA platform, so that any node, or peer, that carries metadata about some resources, can announce an Edutella search service to the network. When looking for information on Edutella, your question will be routed to peers that can answer your query, and they will return matching results to you.

There are actually three types of roles to fill in an Edutella network: provider (provides a query service), consumer (asks questions) and hub (manages query routing in the network). An Edutella network will contain many types of peers which may combine several of the roles.

Hubs are typically set up to increase performance in the network. Most providers will not need to care about hubs at all, as they operate transparently in the Edutella network.

Examples of providers, exposing data to the Edutella network, could be:

- A traditional LMS system at an educational institution;
- A modern RDF-based repository such as UNIVERSAL , OLR or SCAM;
- A metadata harvester that collects information from legacy archives, such as OAI archives or Z39.50 sources;
- A mediator database such as AMOS , that searches a number of databases in combination, while only exposing one query service to Edutella; orAny other kind of database containing learning object metadata.

Many other kinds of metadata providers can be imagined. To be a provider, all that is required is that you are able to answer questions formulated in the Edutella query language. Any kind of information source can be given an Edutella interface.

Examples of consumers that use Edutella to find information could be:

- The "search" tool in an LMS system that uses Edutella to get answers;
- A generic self-contained search tool, such as Conzilla, or a domain specific search tool such as the SWEBOK example application;
- An end-user applications that uses Edutella to enhance the user experience with metadata information (such as "related material");
- An augmented-reality system that displays and uses metadata for objects in three-dimensional space (real or virtual);
- A web portal that includes an Edutella search interface;
- A mobile device (PDA, cell phone, etc.) that gathers information from Edutella to enhance your stay in Rome;
- A smart software agent that gathers relevant information from Edutella to help construct a learning environment; and
- A crawler or push-based system such as CourseWare Watchdog, that uses Edutella as an additional information source.

It should be evident from this list that Edutella support can be added to many kinds of software. And as Edutella supports any kind of metadata expressed in RDF, all kinds of information can be distributed, and not only the pure learning object metadata.

The Vision behind Edutella

Edutella is driven by a vision of a global Democratic Information Network - democratic in the sense that anyone is allowed to say anything about anything. This kind of vision is not new. The Internet has been designed as a peer-to-peer network where anyone can connect to anyone, and that is one of the main reasons for its success. In the same way, the success of the WWW through leveraging hypertext is fundamentally dependent on a peer-to-peer model, where anything may link to anything. This creates a global democratic web, where there is no single point of control, no middle man in control of the network.

However, the web has developed into a predominantly client-server based system, which mainly relies on centralized information handling, something that is at odds with basic Internet technology. This trend is even more evident in the case of e-learning systems, where large-scale databases of learning objects are becoming the standard.

Peer-to-peer networks can be a way out of that trap. Edutella makes it possible for anyone, even with very limited technical and financial resources, to participate in the exchange and annotation of learning resources.

For Edutella, this vision means that anyone must be able to attach any metadata to any learning object. What makes this such an important feature? We will now look into the design goals of Edutella that enable a different kind of e-learning infrastructure.

Design Goal 1: Subjectivity in Metadata

Many metadata-aware systems only contain indisputable information such as title, author, identifier, etc. (most Dublin Core elements are of this kind). Learning objects also need many other kinds of metadata, such as an indication of the granularity of objects, pedagogical purpose, assessments and learning objectives, etc. However, many implementers are skeptical about using such metadata.

One of the reasons for this skepticism is the fact that properties of that kind do not represent factual data about a resource, but rather represent interpretations of a resource. When metadata is treated as authoritative information about a resource, adding descriptions of subjective features becomes not only counter-productive, since it excludes alternative interpretations, but may also be dishonest or authoritarian, forcing a subjective interpretation on the user. This creates unnecessary conflicts of interest and is unfortunately hindering the adoption of metadata technologies.

Edutella takes the position that this problem is partly due to lack of technological support for a different model. When metadata descriptions are instead properly annotated with their source, creating metadata is no longer a question of finding the authoritative description of a resource. Multiple, even conflicting descriptions can co-exist. This amounts to a realization that metadata descriptions are just as subjective as any verbal description. We must allow people and institutions to express different views on learning objects.

It is a fact of life that consensus on these matters will likely never be reached, and the technology must support diversity in opinion, not hinder it. Meta-metadata (information about metadata) and subjective metadata is thus of fundamental importance for a metadata architecture. In a democratic network, ‘objectivity’ is defined by consensus, not by authority. Metadata needs to be a part of that consensus building process.

Naturally, the problem of supporting this fundamental subjectivity is not trivial. By designing Edutella on top of the Semantic Web framework, the built-in support in RDF for meta-metadata will make this task surmountable. Imagine, as a simple example, adding a link called "Who said this?" to each search result. Another possibility is to add functionality to search using only trusted sources. This example emphasizes the need for networks of trust and digital signatures of metadata, in order to ensure the sources of both metadata and meta-metadata. Supporting webs of trust will be a fundamental part of the Semantic Web infrastructure, and thus of Edutella.

Design Goal 2: A Metadata Ecosystem

Implementing metadata as authoritative, objective information about a resource, consisting of facts that do not change, also has the effect of efficiently hindering context-dependent metadata. How do you describe a resource if you don' t know what its intended use is? For example, a single piece of media like a photograph can have different meaning when used in a History context than when used in a Photography context. These contexts may very well not be known when the resource is published, and new uses of resources may arise long after publication. So the choice is to fix a context at the outset, or not describing any context-specific information at all.

Many resources that are useful in learning (such as the material in libraries) are not even designed to be learning objects. Forcing the creator to annotate them using learning object metadata descriptions is unreasonable and often unrealistic.

In Edutella, metadata can be handled as a distributed work in progress, where updating and modifying descriptions is a natural part of the metadata publishing process. There is no central repository where your metadata changes need to be pushed - all metadata is stored at the provider, and there can be several providers supplying information around a single resource.

Treating metadata as a work in progress and allowing subjective metadata leads to a new view of metadata. Metadata is information that evolves, constantly subject to updates and modifications. Competition between descriptions is encouraged, and thanks to RDF, different kinds and layers of context-specific metadata can always be added by others when the need arises. Any piece of RDF metadata forms part of the global network of information, where anyone has the capability of adding metadata to any resource. Edutella then handles combining resource metadata using different vocabularies and coming from different sources.

In this scenario, metadata for one resource need not be contained in a single RDF document. Translations might be administrated separately, and different categories of metadata might be separated. Additional information might be contributed by others.

Consensus building then becomes a natural part of metadata management, and metadata can form part of an ongoing didactic discourse. The result is a global metadata ecosystem, a place where metadata can flourish and cross-fertilize, where it can evolve and be reused in new and unanticipated contexts, and where everyone is allowed to participate. In this way, Edutella provides support for a bottom-up conceptual calibration process, which builds consensus within communities of practice.

Design Goal 3: Extensible Syntax and Semantics

In developing and applying metadata standards for learning objects, important considerations include interoperability and extensibility. Interoperability in this context means that different systems are able to exchange information about learning

objects without requiring complex translation tools, while extensibility means that they are able to incorporate other metadata elements and vocabularies than those explicitly specified in the standard. Both issues are very important for Edutella, as interoperability enables cross-searching of repositories, and advanced extensibility is needed to support domain and application specific additions to the metadata.

Edutella uses RDF for metadata expressions in order to be maximally compatible with these two principles. It makes interoperability simple, as RDF provides a single framework for expressing any kind of metadata, while leaving the flexibility for defining a custom vocabulary. RDF also includes powerful facilities for extensions. These extensions come in two kinds:

1.Structural extensions. This includes adding completely new metadata elements to resources. This is built into RDF itself, and can be done in the same metadata document ("model" in RDF terms) or in a separate one.

2.Semantic extensions. This includes refining existing elements and vocabulary terms, the way "abstract" refines "description" in Dublin Core, or the way "Digital Text" is a kind of "Text" in the case of a learning resource type. Expressing this in RDF is done in the RDF Vocabulary Description Language (also known as RDF Schema).

The need for extensions will explode with the number of domain and application specific standards that are developed. Most deployments will have a need for extensions of many kinds, both domain specific and application specific. The problems with mixing metadata vocabularies can therefore be expected to increase. However, current metadata standards in wide use in the e-learning domain, notably XML versions of Dublin Core and IEEE LOM, do not support a common model for extensions.

Edutella avoids many parts of this problem by relying on the built-in mechanisms of RDF and RDF Schema. Supporting mix-and-match vocabularies and supporting semantic extensions are design goals at the very core of Edutella.

The vocabularies most frequently used (separately and in combination) within Edutella at the time of writing are:

- Simple Dublin Core
- Dublin Core Qualifiers
- Vcard
- IEEE LOM
- IMS Content Packaging

as well as a number of locally developed taxonomies, vocabularies, refinements and element sets.

Using Edutella - What can I do?

Some of what Edutella wants to accomplish has now been shown, and what technologies are used to implement the Edutella visions. It remains to understand how Edutella is supposed to be used, and how it can support practical work in e-learning. The following scenario highlights some of the possibilities. While it is not realistic in every detail, it is hoped that it will demonstrate the different ways in which Edutella-enhanced tools can enrich the learning experience. The readers are also encouraged to

add their own visions to this picture - Edutella is an infrastructure on top of which many kinds of functionality can be added.

A Story about an Edutella user

You are studying Taylor expansions in mathematics. Your teacher has not provided the relevant links to the concept in your concept browser, Conzilla, so you first enter "Taylor expansions" in the search form. The result list shows that Taylor expansions occurs in several contexts of mathematics, and you decide to have a look at Taylor expansions in an approximation context, which seems most appropriate for your current studies.

After having studied the background material on the different kinds of approximations for a few hours, you decide you want to see if there are any appropriate learning resources. Simply listing the associated resources turns out to return too many, so you quickly enter a query for "mathematical resources in Swedish that are related to Taylor expansions, and are on the university level and part of a course in calculus at a Swedish university". Finding too many resources again, you add the requirement that an older student at your university must have given a good review of the resource. You find some interesting animations provided as part of a similar course at a different university, which has been annotated in the portfolio of a student at your university, and start out with a great animation of three-dimensional Taylor expansions. The animation program notes that you have a red-green color blindness and adjusts the animation according to a specification of the color properties of the movie, which was found together with the other descriptions of the movie.

After a while you are getting curious. What, more precisely, are the mechanisms underlying these curves and surfaces? You decide you need to more interactively manipulate the expansions. So you take your animation, and drag it to your graphing calculator program, which retrieves the relevant semantic context from Conzilla via the application framework, and goes into Edutella looking for mathematical descriptions of the animation. The university, it turns out, never provided the MathML formulas describing the animations, but the program finds formulas describing a related Taylor expansion at an MIT OCW course site. So it retrieves the formulas, opens an interactive manipulation window, and lets you experiment.

Your questions concerning Taylor expansions multiply, and you feel the need for some deeper answers that the computer cannot give you. Asking Edutella for knowledge sources at your own university that have declared interest in helping out with advanced Calculus matters, you find a fellow student and a few math teachers. Deciding that you want some input from the student before talking to the teachers, you send her some questions and order your calendaring agent to make an appointment with one of the teachers in a few days.

A week later you feel confident enough for changing the learning objective status for Taylor expansions in your portfolio from "active, questions pending" to "on hold, but not fully explored". You add your exploration sequence, the conceptual overviews you produced in discussion with the student and some annotations, to the public area of your portfolio. You conclude by registering yourself as a resource on the level "beginner" with a scope restricting the visibility to students at your university only. This way, your knowledge is made available both as annotations to Edutella, and as a real-life contact.

This scenario is not a complete fantasy. Tools to enable this kind of learning experience via Edutella are being designed right now, andresearch is underway to make them even better. Some of the important features of Edutella can be seen being used in this scenario:

- Distributed material and distributed searches; mixtures of metadata schemes (for example, personal information and content descriptions) being searched in combination;
- Machine-understandable semantics of metadata (calendaring information, animation parameters, finding the right kind of resources);
- Human-understandable semantics of metadata (contexts, persons, classifications);
- Tool interoperability - any tool can use the technology;
- Distributed annotation of any resource by anyone, using digital portfolios;
- Personalization of tools, queries, and interfaces, affecting the experience in several ways; and
- Competency declarations and discovery for personal contacts.

Conclusions

Learning, just like other human activities, cannot and will not be confined within rigidly defined boundaries such as course systems. There is a strong need for more decentralized structures. Moreover, a learning environment has to support trust building and rich forms of communication between teachers and learners as well as between learners. In order to be powerful, the environment must be inspiring and trigger curiosity for the learning task. Semantic Web technologies form a basis for realizing a multitude of fascinating e-learning visions, by giving software access to the semantics of your material. Edutella is a way to support the introduction of such technologies in e-learning systems.

Although much of the present development within e-learning is driven by the so-called knowledge economy, there are more fundamentally important issues for the future; namely, how to provide access to knowledge for people who cannot afford to pay. Our efforts within the Edutella project is driven by the overall vision of a global knowledge community, where relevant information and efficient support for the knowledge construction process is freely available for all.

For more information, see the web site of the KMR group at http://kmr.nada.kth.se.

References

Web sites

Learning Lab Lower Saxony	http://www.learninglab.de
the KMR Group at KTH	http://kmr.nada.kth.se
the Uppsala Database Laboratory	http://www.dis.uu.se/~udbl/
Stanford Infolab	http://www-db.stanford.edu
AIFB at University of Karlsruhe	http://www.aifb.uni-karlsruhe.de
UNIVERSAL project	http://nm.wu-wien.ac.at/research/
JXTA Organization	http://www.jxta.org
RDF Information	http://www.w3.org/