The Courseware Watchdog: an Ontology-based tool for finding and organizing learning material

Julien Tane, Christoph Schmitz, Gerd Stumme, Steffen Staab, Rudi Studer

Learning Lab Lower Saxony (L3S), Expo Plaza 1, D–30539 Hannover, Germany;
www.learninglab.de; <lastname>@learninglab.de
and
Institute for Applied Informatics and Formal Description Methods (AIFB),
University of Karlsruhe, D–76128 Karlsruhe, Germany; www.aifb.uni-karlsruhe.de/WBS/

Abstract. Topics in education are changing with an ever faster pace. E-Learning resources tend to be more and more decentralised. Users need increasingly to be able to use the resources of the web. For this, they should have tools for finding and organizing information in a decentral way. In this, paper, we show how an ontology-based tool suite allows to make the most of the resources available on the web.

1 Introduction

In later years, mobile technology and internet access have improved in such a way that it has become realizable to have resources stored remotely. In the E-Learning domain, the use of notebooks or mobile devices implies a new way of managing resources. The goal is, therefore, to make the best possible of the resources of the network. [16] predicts the role of mobile devices in education and argues that E-Learning has to be seen as a part of the general framework of knowledge management. To achieve this, it is important to integrate the technologies of these two domains.

Among the current knowledge management techniques,ontologies play a greater role than ever. Current research on ontologies has shown that they facilitate the retrieval, interaction and management of resources, for some examples see [20] or [10]. In the E-learning domain, standard shemas have been developed to help describe learning objects[1]. Although these developments are a good start, there is a need for a more comprehensive approach which integrates the content, structure and evolution of the learning material. We show here how ontologies help finding and organizing distributed courseware resources by offering a common framework for the retrieval and organization of courseware material. We illustrate this with a scenario.

In the first section we briefly expose the role of ontologies for E-Learning. The following section will presnet a scenario displaying how a courseware watch tool will be used. This will lead us to an overall presentation of the integrated architecture of our courseware watchdog. In the four subsequent sections we will address the specific modules of the courseware watchdog and how they address the problem. Finally we will sum up our results and hint at further research goals.

2 Semantic Web and E-Learning

On a personal computer, it is possible to organize one's resources according to one's need. In the case of remote resources, it is not possible anymore, since their storage is not under the control of the user. Through the use of hypertext, it is possible to link to remote material

[1] For more information on these standards, please see http://ltsc.ieee.org/index.htm.

and retrieve it when needed. However, the particular problem of finding and organizing this remote material is crucial, if one to make the most of remote material. Since forcing the storage position is neither possible nor desirable, the only remaining alternative is querying. But, no matter how good the indexing mechanism of search engine are, it is not possible to retrieve learning material in the web using only keyword queries. Solving this issue requires to be able to query semantically for resources.

[2] show that standards like LOM or Dublin Core are getting importance. They provide increased information on the learning material that is to be found in the web. However, their simple structure prohibits their use for modeling more complex knowldege. [19] explains how Semantic Web technologies based on ontologies can improve different aspects of the management of E-Learning resources. Indeed, ontologies are a means of specifying the concepts and their relationships in a particular domain of interest. Web Ontology languages , like OWL, are specially designed to facilitate the shaing of knowledge between actors [18] in a distributed environment. We wish to emphasize here on diverse advantages.

From the modelling point of view ontology languages are not only able to integrate[2] LOM and Dublin Core metadata , but also allow for the extension of the description of the learning objects[3] with non standard metadata, thus giving users and groups of users more flexibility when sharing resources.

Ontologies were also successfully used for integrating resources in heterogenous repository networks as well as in multiple applications: annotation, content browsing, content structuration.

Research in the E-Learning domain shows that standardisation is needed[4], but we think these should be integrated in common framework, which we illustrate with a scenario and the description of our courseware watchdog.

3 Scenario and User Requirements

To illustrate the ultimate goal and purpose of our tool, we will describe in this section a scenario. It will show the different tasks that need to be addressed when trying to find and organize courseware material.

3.1 A Professor Preparing his Lectures

Professor Meyer is a university professor at a German university in the domain of computer science. He is in charge of some courses every semester. His main field of activities are: Data Mining and Knowledge Management. Since these fields of studies evolve very rapidly he has to be aware of the latest evolution of these domains.

At the beginning of the semester break, Professor Meyer prepares the two lectures and two seminars which he will give during the next semester. He already has material from previous lectures but he feels that there is still room for improvement in the different lectures.

These different courses are:

- “Introduction to Computer Science” designed for freshmen
- a “Knowledge Discovery” lecture for more advanced students.
- a seminar on “Knowledge Management”
- a seminar on peer-to-peer and web services.

---

[2] See http://wwww.imsgloabl.org/metadata/ ond RDF(S) LOM Binding.

[3] For the purpose of the following scenario, the simpel case of an ontology integrating LOM, Dublin Core and the ACM taxonomy should be sufficient.

[4] See for instance the efforts of the Learning Technology Standard Commitee.

Professor Meyer expects to find a lot of material accessible on the web for the first lecture and he already has some scripts which are more or less ready to be used. He has already annotated part of the material with educational markup using the LOM and Dublin Core standards, and using for the domain of computer science the well known ACM taxonomy.

Browsing Model and Content: Professor Meyer starts his tool for finding and managing remote resources. He planned the construction of the two courses in the following way: first, he needs to have an overview to plan what the different lectures should contain. He looks for an ontology on the different domains of his lectures. These ontologies contain both the most important concepts and instances of the domain.

Furthermore, he might be sharing the ontology with colleagues or reusing on the respective domain, he retrieved from a pool of existing ontologies. Therefore, he needs first to familiarize himself with its content and the meaning of the diverse concepts and relations. Because most user disagree on certain aspects of the modeling.

Of course, he can also access usual documents (HTML pages or PDF files through the use of the interface). This allows him to control the documents not by their position but rather by their conceptual representation as long as the relevant parts are represented in the ontology model.

Finding Relevant Material He now wants to find new material. For this, there are different approaches: either look for it in the world wide web or use structures that provides semantic metadata about learning material. We suppose, he has tools for this ready. Both tools use an ontology, which will in the end improve the interaction with the learning material.

He loads the two given ontologies which were stored in Professor Meyer's ontology registries as 'LOM + Computer Science".

Querying a Network of Semantically Annotated Material: Suppose our professor has access to an Edutella[5] network [17]. Edutella is a peer-to-peer framework, where the different peers provide semantically annotated metadata on learning material. It also allows for the integration of web services to gain access to material offered by libraries or similar repositories, see for example [1].

In order to find new relevant material, Professor Meyer first needs to define a query using an Edutella query interface[6]. So Professor Meyer looks for a lecture on the topics "Algorithmics" or "Knowledge Discovery". Using the query interface, he defines a query:

> look for a "Lecture" which "hasTopic"[7] "Algorithmics" or which "hasTopic" "Knowledge Discovery" and retrieve also for each match the values of properties "dc:title" and "dc:author"

He sends the query to the network and gets an answer. He can then browse the lectures having as topic "Algorithmics" or "Knowledge Discovery". He can then send more queries to the Edutella network to get more information about the specific lectures or authors that are of interest to him.

[5] In this scenario we will take edutella as a prototypical distributed network of semantically (basically LOM + Dublin Core +...) annotated learning material. Another example could have been POOL, see [9].

[6] there exists already different edutella query interfaces such as Conzilla, or the one integrated in the courseware watchdog in this paper

[7] this has been simplified for sake of simplicity. But the interfaces allow to do this in a very simple ways

Finding Learning Material on the Web: Professor Meyer knows some internet web site for the topics that are of interest to him. He is quite certain some interesting material would be accessible there, had he only the time to browse the sites.

The courseware watchdog assorts him with a focused crawler which follows the links starting from his start pages, and retrieves relevant material. To describe his topic of interest, prof. Meyer uses an ontology. Using the browsing interface of the ontology, he can select the topics and elements he is most insterested in.

He decides to send two crawlers to search new material on Knowledge Management and Peer-to-Peer. For this, he selects the corresponding concepts as well as other concepts that seem to be relevant in both domains. Then he sends the focused crawlers by giving them start pages: for example homepages of the European projects Ontoweb and SWAP. The crawler will follow the links of the web until it finds web documents with enough elements corresponding to its query. The crawler uses the ontology for guiding the search by expanding the query through diverse ontological means. The retrieved results can finally be browsed by using the same ontology.

Cluster the Documents: Professor Meyer now wishes to organize the documents he has retrieved. Ideally, he wants to group similar documents together and structure the documents according to certain criteria. For this, he uses the Subjective Clustering mechanism which allows him to cluster the documents using their conceptual similarities. By choosing an ontology specific to the domain of interest, he will be able to cluster related documents closely together since it will make use of the conceptual information coded into the ontology to perform better than common clustering techniques.

Moreover, it is important for him to be able to understand the clusters, for this he needs special vizualizations techniques, which allow him to understand why the documents have been grouped together.

Evolution: Professor Meyer wants to be aware of the evolution of the vocabulary of his field and be able to augment his ontology of the domain. For this, he can use tools to find frequent terms out of a corpus of documents. These terms can be either single terms or multiword terms (for example: “knowledge discovery”, or “computer science”). By having a look at the frequent terms he should be able to find out new techniques or methods which might be of interest to him. If he decides that some terms interest him, he will be able to integrate them into the ontology.

3.2 User Requirements

From the preceding scenario, one can extract the diverse tasks that need to be supported by a courseware watchdog. These can be summed up in the following list.

- understanding the ontology and browse the content
- querying semantically annotated resources repositories
- retrieving relevant material through crawling
- organizing the documents according to the ontology
- updating the ontology and knowledge base according to current data

Since ontologies can be used in all these tasks to improve user interaction, find or structure the data, it is desirable to integrate them in a common framework.

It should be noted that the two retrieval approaches: querying repositories and retrieving elements through crawling can be considered as complementary. While one is working on metadata described in a specific semantic model, the other tries to use the resource and its

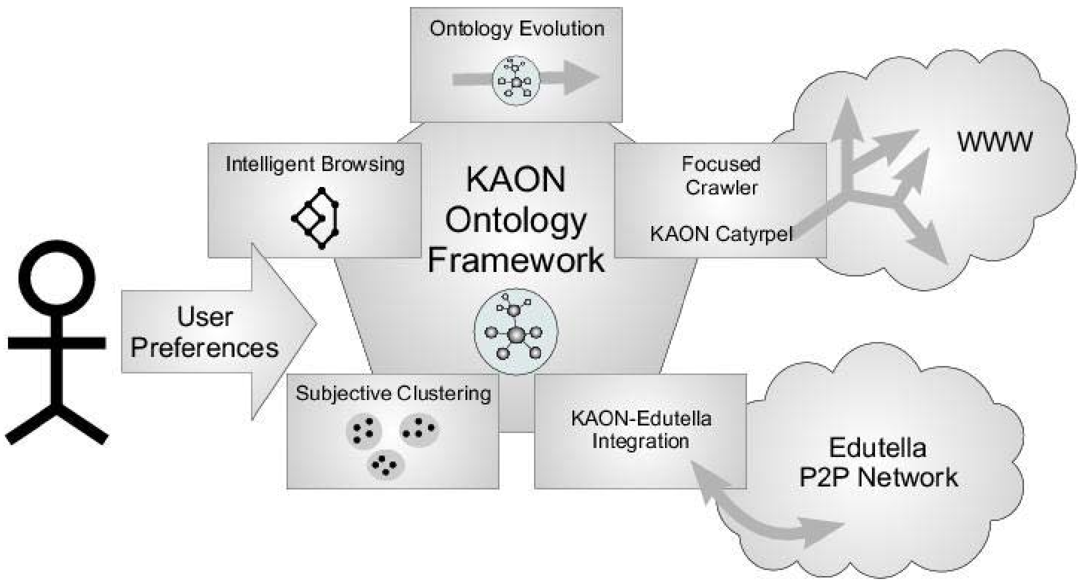

Fig. 1. The components of the Courseware Watchdog.

relation to a semantical model to select it. In both cases, the interaction with the model is necessary.

The use of the semantic model makes it then simple to use a conceptual browsing material to interact with the diverse tools make use of the browsing interface to start the retrieval, or to visualize the results. This will allow the use of one ontology and therefore a specific semantic for two different kind of queries

Moreover, the scenario shows that these five tasks are not completely independent. Not only do they all share the ontology of the domain, but they also share their results. Indeed, the results of a search[8] can be browsed or clustered, or either new concepts and instances can be extracted from the documents. The information about these documents will then be stored in the knowledge base and the user does not have to know the position of the specific resources.

If the resources are integrated into the ontology, they can be then manipulated semantically. And the metadata attached to them can be reused for other needs, thus reducing the modelling parts. In the rest of this article, we will describe our courseware watchdog framework, starting with a general description of the framework, before considering each module independently.

4 Courseware Watchdog

The courseware watchdog described in this paper addresses the preceding requirements by using a comprehensive approach which exploits concepts from the Semantic Web, such as ontologies, in an E-Learning scenario [19]. It is part of the PADLR framework (Personalized Access to Distributed Learning Repositories) that builds upon a peer-to-peer approach for supporting personalized access to learning material.[9]

When developing the Courseware Watchdog, we aimed at addressing the different problems made appearant by the previous scenario. The tasks to be solved are addressed by different

[8] Which could be performed in the edutella or network in the web
[9] http://www.learninglab.de/pdf/L3S_padlr_17.pdf

modules. One important goal was to use a single semantic model for the different tools. We tried to integrate the different tools together. We show that their combination offers the user a single simple tool for tasks depending on each other. The Courseware Watchdog consists of the following components (see Figure 1):

- *Visualization and intelligent browsing* allows the browsing of the model and knowledge base in order to improve the interaction between the user and the content.
- A *focused crawler* will find related web sites and documents that match the user's interests. The crawl can be focused by checking new documents against the user's preferences as specified in terms of an ontology.
- Integration into the *Edutella framework* enables to query for metadata on learning objects with an expressive query language
- As different users may have different points of view, *subjective clustering techniques* are used to generate subjective views onto the documents.
- In order to reflect changes and trends within the field of interest, ontology learning methods will be employed to facilitate *ontology evolution*.

All the modules mentioned are built on top of an ontology framework named KAON. For more details on the ontology model, as well as on the KAON API we refer to [7], or to the KAON Developer documentation[10]. However, it should be noted that more integration issues have been addressed.

Integration in the courseware watchdog is done at different levels:

- at the semantic level - through ontologies
- at the web structure level - the structure of the graph of web document is stored in an ontology
- at the structure level of the corpus - the different algorithm for the clustering and ontology evolution use the same corpus model

This common integration model allows to use both the browsing and the querying of the resources available or discovered. Or it allows the interaction with the results of algorithm. For instance, it is then possible to use the clustering results as input to the ontology evolution.

5 Browsing of Watchdog Data

As shown in the scenario, the interaction of the user with the ontology is crucial for all ontology-based tools. In the courseware watchdog, this is done using the browsing component. By displaying specific hierarchies[11] as lattices, it gives the user simple entry points on the data accessible.

Formal Concept Analysis [8, 21] is a conceptual clustering technique which allows the display of hierarchies of concepts. The courseware watchdog follows a recent implementation a Conceptual Email Management system (CEM) which supports exactly this navigation in collections of emails [6]. When applied to learning material, the multiple inheritance within this hierarchy provides a rich conceptual landscape for navigating and retrieving the educational media.

Through Formal Concept Analysis, it is possible to display multiple inheritance as well as multiple instantiation in a very easy way, and permits also diverse views on data. For example, 2 displays the hierarchy of subconcepts of “algorithm”. One can see that there are instances of “sorting algorithms” which are also “recursive” and “parallel” in the current

[10] http://kaon.semanticweb.org.

[11] For example, it displays the concept hierarchies or topic hierarchies.

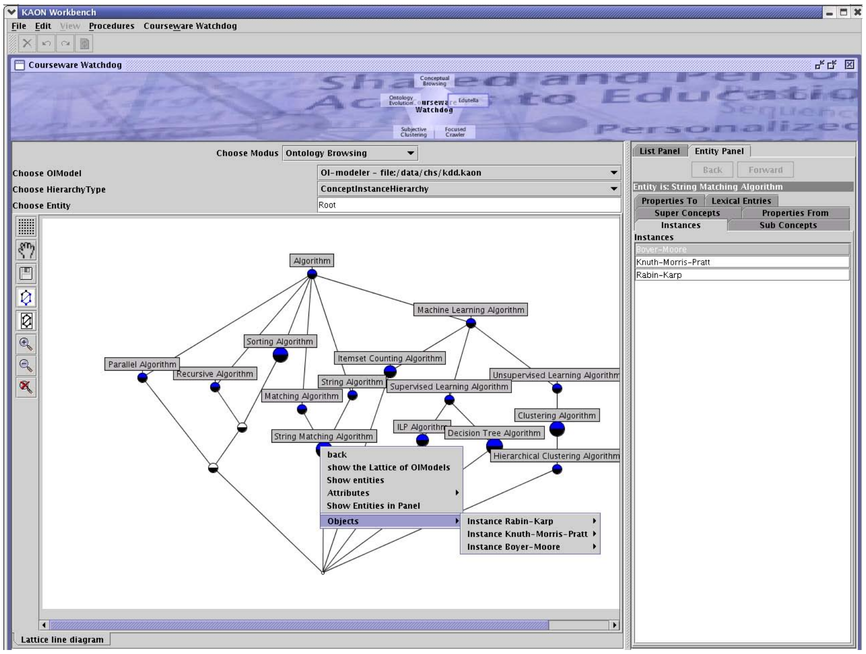

Fig. 2. The browsing interface of the courseware watchdog

knowledge base associated with the ontology. Through the relational browsing (i.e: simple technique allowing to jump from the display of an instance or concept to another instance or concept by following the relation they have) it is possible to navigate through the relation of the ontologies, and then display different kinds of hierarchies according to one's needs. By clicking on the concept "String matching algorithms", you will be able to find its instances, for example "Boyer-Moore" algorithm, find the lectures which refer this algorithm, by following the relation "isReferedBy" of the instance "Boyer-Moore"(selected in the right panel of 2.).

As generic way of interacting with the ontology, this component plays a central role in the courseware watchdog, since it allows, additionally to the browsing, the querying or selection of entities which can be then used in the crawling, clustering or evolution process.

6 Focused Crawler and Edutella

In this section, we will describe the two retrieval components which allow the user to find material according to his interest. In both cases, the ontology browsing can be used to define the elements that should be looked for. Thereby the focused crawler is able to look for material in the web without prior annotation[12], whereas the edutella approach presupposes that the learning objects are semantically annotated.

6.1 Focused Crawler

A web crawler is a program that collects data from the web automatically by following links extracted from web documents. Thus, a portion of the web is traversed in a breadth-first manner, usually without regarding the relevance of the collected documents with respect to the user's needs. In order to restrict the traversal to material relevant to the user, the crawling

[12] But it is also able to cope with annotated material

process can be *focused*[13]. Focusing in this context means preferring those links in the crawling process that appear to be pointing to relevant documents.

Our focused crawler [14] is an ontology-based focused crawler embedded in the KAON environment for ontology-based tools. With it, the user can specify topics of interest in terms of an ontology. The user's preferences, i. e. entities in the ontology, and the ontology itself are then used to compute the relevance of documents and hyperlinks. After simple linguistic preprocessing (HTML tag removal, stemming etc.), lexical entries of the ontology are matched against the text, and a relevance score is computed.

The structure of the ontology allows determining a useful measure of relevance, even if no exact match can be found. Several relevance measures can be used, which compared favourably against breadth-first search or simple keyword-based measures.

For example a user can indicate that he is interested in the topic of "Machine Learning Algorithm". When searching for "Machine Learning Algorithm", the focused crawler may for instance come along the web pages of a course on "Knowledge Discovery". The crawler will then update the ontology, and will return the page with a certain score.

Of course, there is a need in the end to use and understand the results of the crawling. For this, two different approaches can be used. First of all, the results of the crawling are put into an ontology-driven corpus model. This ontology can be used to interact with the user.

6.2 Integrating the Edutella Peer–to–Peer Network

The courseware watchdog includes the possibility to access the Edutella peer-to-peer network. Peer-to-peer (P2P) networks are decentralised networks, which allow for publishing and searching resources based on direct collaboration between its nodes.

The Edutella network applies the P2P paradigm to the exchange of structured information about available learning resources.[14] A common data model facilitates the integration of data sources such as relational database systems or XML and RDF repositories. Thus, all of these can act as Edutella peers.

For instance, our lecturer may query Edutella. He will then be able to locate relevant material on the P2P network which has been provided by other Edutella users. This way, he may for instance come across further lectures and tutorials about the topics he is interested in.

In the courseware watchdog, the Edutella module has two functionalities. First, it can serve as a provider in the network. That is, it can propose diverse ontologies or metadata repositories. Thus, a professor can offer his courses online.

Second, the user can send queries to the Edutella network which he has previously defined by using the ontology browsing interface. The user does not need to know the Edutella query language to refine queries. He has a repository of template queries which he can adapt to his own need by simply selecting the concept, property or instance in the ontology which fits his need and replace the free suitable variable of the query (cf Figure 3) by the URI of the resource. This, of course makes it a lot easier to use for simple users, who surely would not want to learn any query language.

Of course, because of the integration of the whole in a common data model it is possible to visualize the results using the browsing interface.

7 Subjective Clustering

As shown in the scenario, a lecturer has access to a lot of course material which he needs to organize (For example, lectures he retrieved through the focused crawler or Edutella). Of

[13] There are in the literature other focused crawler, see for example [3].
[14] http://edutella.jxta.org

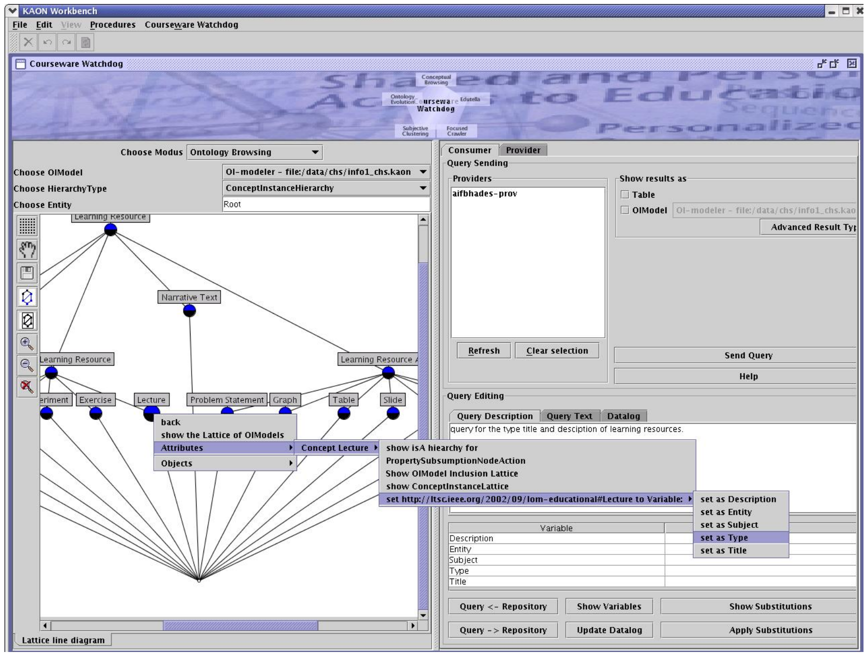

Fig. 3. Refining a query.

course he wants to group together documents according to their topics and similarities [12]. Because users typically do not want to exactly specify their complete profile and because users' profiles tend to change rather often (which is a major problem for recommender systems), we want to provide to the user views onto existing educational media. Conventional clustering techniques provide a first answer for this purpose.

Typically, however, clustering is used to give a single, "optimal" view on (learning) components. This is not suitable to account for the plurality of views that exist when looking at educational media. We have recently developed clustering mechanisms that allow to provide *subjective views* onto documents collections [11], which are based on an underlying ontology.

For instance, one view may concentrate on differences and similarities of the content of learning material, while another view may concentrate on its presentation form, or on the levels of skills and experiences needed. The lecturer can then use the first view to select the material which addresses the topics which are most relevant to his planned course. He might then use the second view in order to see how the material is distributed over different types of material like presentation slides, exercise sheets, or online demonstrations.

Following [13], we display the distribution of the most relevant terms of the diverse clusters through the use of a lattice displaying the diverse combinations of terms occuring in the clusters. This combination of the browsing and the clustering results helps the user to understand better the results of the clustering process and select in a simple way the lectures which interest him.

This allows a very simple interaction with clustering results and achieves the goal of helping the user in organizing his learning material. For example, Figure 4 shows how the user having selected two documents, he can relate them to the concept "Decision Tree algorithm" through the relation "topicOf". In the context of the managing of resources, he will then have the possibility to keep only the metadata of the documents that has clustered.

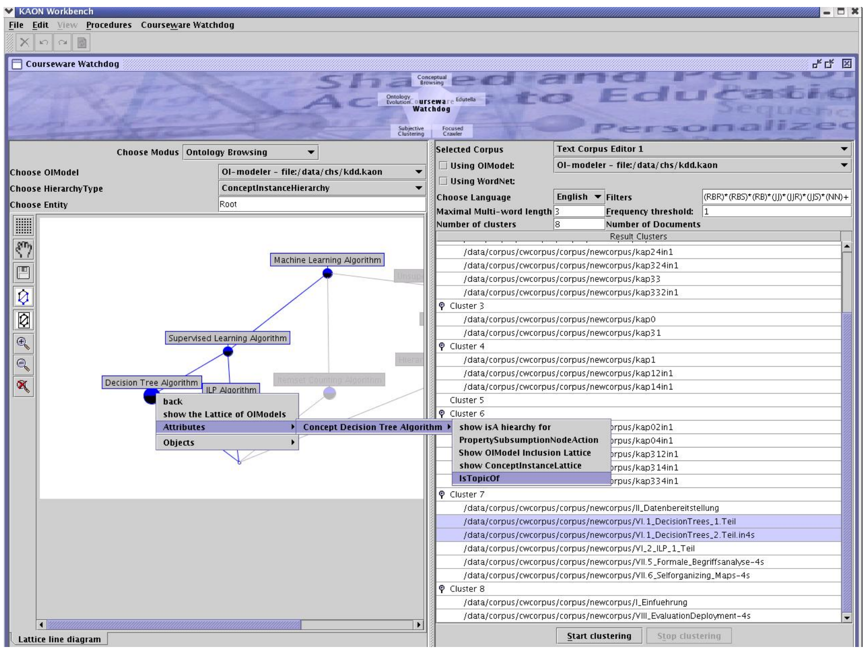

Fig. 4. Simple annotation helped by clustering.

8 Ontology Evolution

The courseware watchdog as presented in this paper so far builds heavily on a proper ontology that reflects what the user is interested in. However, over time such interests will invariably change together with the teaching/learning subject itself. Therefore, the ontology and the topics represented therein need to be updated. One must deal with several requirements incorporated in such updates:

Modifying the ontology: The ontology must remain consistent at all time we use the evolution functionalities of the KAON API, which insure that changes to the ontology will not corrupt it.

Introducing new concepts: The first requirement is about (i) recognizing that a new concept (e.g. a new topic) has appeared in the course material available in the network or on the Web, (ii) inserting this concept into the right place of the taxonomy, and (iii) linking it via further relations to other concepts.

We use methods in [15] to find relevant concepts. Moreover, we are working ontechniques to create ontologies semi-automatically. These will couple current hierarchy building techniques as we presented in [4] and [5] with the courseware conceptual browsing.

For instance, Web Services are today an emerging topic, and will probably have to be included in future courses on the Semantic Web. Hence 'Web Services' will be recognized as a term that denotes a new concept, since it occurs frequently in documents on the Semantic Web. It can be inserted into the concept hierarchy (e.g. as a subarea of computer science). It also must be related to other disciplines (e.g. to business process modeling and E-Business). For this, we use again the display of the browsing interface, where the user can select the place to insert the concept or instance. Then he can relate the new instances or concepts to other concepts or instances.

Although the difficult problem of ontology learning is not solved, we offer a tool that gives the user possibility to adapt his ontology and instance base according to the interesting concepts or instances found in the documents.

9 Conclusion

The courseware watchdog is a comprehensive approach for supporting the learning needs of individuals in fast changing working environments, and for lecturers who frequently have to prepare new courses about upcoming topics.

As shown in the paper, the courseware watchdog addresses the different needs of teachers and students to organize their learning material. It integrates, on the one hand, the Semantic Web vision by using ontologies and a peer-to-peer network of semantically annotated learning material. On the other hand, it addresses the important problems of finding and organizing material using semantical information. Finally, it offers primtive solutions to the problem of evolving ontologies.

The components of the courseware watchdog need further improvement. For instance, focused crawling has to be improved by offering further measures for computing the relevance of documents based on the ontologies and available metadata, and ontology evolution needs further techniques for better reflecting changes in the underlying learning material,such as concept drift detection.

The courseware watchdog indicates how a Semantic Web based approach is better able to meet the retrieval ans management of remote resources, by providing tools for discovering and organizing these.

Acknowledgements

We thank all members of our groups AIFB and FZI for intensive discussions and contributions, especially Marc Ehrig, Andreas Hotho, Boris Motik and Philipp Cimiano.

This research is partially funded by the German Federal Ministry of Education and Research (bmb+f).

References

1. Wolf Siberski Benjamin Ahlborn, Wolfgang Nejdl. Oai-p2p: A peer-to-peer network for open archives. In *ICPP Workshops*, pages 462–468, 2002.
2. Jan Brase and Wolfgang Nejdl. *Ontologies and Metadata for eLearning*, pages 579–598. Springer Verlag, 2003.
3. Soumen Chakrabarti, Martin van den Berg, and Byron Dom. Focused crawling: A new approach to topic-specific web resource discovery. *WWW8 / Computer Networks 31(11-16)*, pages 1623–1640, 1999.
4. Philipp Cimiano, Julien Tane, and Steffen Staab. Automatic acquisition of taxonomies from text: FCA meets NLP. In *Proceedings of ECML/PKDD Workshop on Adaptive Text Extraction and Mining*, Cavtat-Dubrobnik, Croatia.
5. Philipp Cimiano, Julien Tane, and Steffen Staab. Deriving concept hierarchies from text by smooth formal concept analysis. In Data Mining Fachgruppe Machinelles Lernen, Wissensentdeckung, editor, *Proceedings of GI Workshop Lernen -Wissen -Adaptivität (LLWA)*, pages 72–79, Karlsruhe, Germany, 2003.
6. Richard Cole and Gerd Stumme. CEM - a Conceptual Email Manager. In Bernhard Ganter and Guy W. Mineau, editors, *Proc. ICCS 2000*, volume 1867 of *LNAI*, pages 438–452. Springer, 2000.
7. Errol Bozsak et al. KAON – Towards a Large Scale Semantic Web. In G. Quirchmayr (Eds.) K. Bauknecht, A. Min Tjoa, editor, *: Proc. of the 3rd Intl. Conf. on E-Commerce and Web Technologies (EC-Web 2002)*, pages 304–313, 2002.
8. Bernhardt Ganter and Rudolf Wille. *Formal Concept Analysis: Mathematical Foundations*. Springer, Berlin – Heidelberg, 1999.

9. Marek Hatala and Griff Richards. Pool, pond and splash: A canadian infrastructure for learning object repositories. In *Proceedings od the 5th IASTED Int. Conference on Computers and Advanced Technology in Education (CATE 2002)*, pages 54–59, Cancun, Mexico, 2002.
10. Ian Horrocks and Jim Hendler, editors. *The Semantic Web - ISWC 2002. Proceedings of the First International Semantic Web Conference*. Springer Verlag, Sardinia,Italy, 2002.
11. Andreas Hotho, Alexander Maedche, and Steffen Staab. Ontology-based text clustering. In *Proc. of the Workshop "Text Learning: Beyond Supervision" at IJCAI 2001. Seattle, WA, USA, August 6*, 2001.
12. Andreas Hotho, Steffen Staab, and Gerd Stumme. Explaining text clustering results using semantic structures. In *Principles of Data Mining and Knowledge Discovery, 7th European Conference, PKDD 2003*, Dubrovnik, Croatia, 2003.
13. Andreas Hotho and Gerd Stumme. Conceptual clustering of text clusters. In G. Kokai and J. Zeidler (Eds.), editors, *Proc. Fachgruppentreffen Maschinelles Lernen (FGML 2002)*, Hannover, 37-45.
14. Alexander Maedche, Marc Ehrig, Siegfried Handschuh, Raphael Volz, and Ljiljana Stojanovic. Ontology-focused crawling of documents and relational metadata. In *Proceedings of the Eleventh International World Wide Web Conference WWW-2002*, May 2002. (Poster).
15. Alexander Maedche, Viktor Pekar, and Steffen Staab. Ontology learning part one — on discovering taxonomic relations from the web, 2002.
16. Hermann Maurer and Marianne Sapper. E-learning has to be seen as part of general knowledge management. In *Proceedings of ED-MEDIA 2001*, pages 1249–1253, Charlottesville, USA, 2001. AACE.
17. Wolfgang Nejdl. Semantic Web and Peer-to-Peer Technologies for Distributed Learning Repositories. In *17th IFIP World Computer Congress, Intelligent Information Processing / IIP-2002*, 2002.
18. Rudi Studer Steffen Staab, Hans-Peter Schnurr and York Sure. Knowledge processes and ontologies. *IEEE Intelligent Systems*, 16(1), 2001.
19. Ljiljana Stojanovic, Steffen Staab, and Rudi Studer. E-learning based on the semantic web. In *WebNet2001 - World Conference on the WWW and Internet*, Orlando, Florida, USA, 2001.
20. Rudi Studer and Steffen Staab, editors. *Handbook on Ontologies in Infornation Systems*. Springer Verlag, Berlin – Heidelberg, 2003.
21. Gerd Stumme and Rudolf Wille, editors. *Begriffliche Wissensverarbeitung — Methoden und Anwendungen*. Springer, Heidelberg, 2000.