[BRZSt07] *Sonja Bergsträßer, Christoph Rensing, Birgit Zimmermann, Ralf Steinmetz;* **Building a Representation of Learning Resources to Support their Re-Purposing;** World Conference on Educational Multimedia, Hypermedia and Telecommunications (EDMEDIA 2007), Vancouver, Juni 2007

Building a Representation of Learning Resources to Support their Re-Purposing

Sonja Bergsträßer[1], Christoph Rensing[1], Birgit Zimmermann[1,2], Ralf Steinmetz[1]

[1] KOM Multimedia Communications Lab
Technische Universität Darmstadt
Merckstr. 25, 64283 Darmstadt, Germany
{sonja.bergstraesser, birgit.zimmermann, christoph.rensing, ralf.steinmetz}@kom.tu-darmstadt.de

[2] SAP AG
SAP Research CEC Darmstadt
Bleichstr. 8, 64283 Darmstadt, Germany
birgit.zimmermann@sap.com

Abstract: The Re-Use of Learning Resources is an important issue and it is addressed in several actual papers. Re-using Learning Resources in a new context makes it necessary to adapt them to fit the new requirements. Our goal is to support users to perform these adaptations of Learning Resources, because the adaptation of Learning Resources to a new context is not easy to perform. Re-Purposing comprises challenges like to deal with multiple files in multiple formats. To hide this from the user there is need for an abstraction of the underlying details. With a representation which is including only the information the user needs and which is abstracting from the obstacles an adaptation of Learning Resources becomes a possible task even for novice users. This paper points out the process how a representation is developed in general and how this representation development process is applied to build a content representation of Learning Resources which is used to support Re-Purposing.

1 Introduction

The production of new Learning Resources is cost and time consuming. Therefore the Re-Use of existing Learning Resources becomes more and more important. To re-use an existing Learning Resource in a new context or scenario it first has to be re-purposed to suit the new requirements. Re-Purposing is the transformation of a resource to suit a new context. This implies that the author who wants to re-purpose a resource has a new purpose in mind. Re-Purposing tools should support an author and help him to transform the resource according to the author's intention. Just editing the resource with one or even many specific tools does not help the authors in their work.
Our goal is to develop a Re-Purposing approach which supports the author in transforming a Learning Resource to suit a certain context and guides him through the Re-Purposing process. A Re-Purposing tool should enable the author to pursue his intention instead of forcing him to think of all issues and perform them one-by-one in small editing steps.
Re-Purposing is a process consisting of three main issues. First of all a Learning Resource has to be modularized into suited parts, so called modules. After that these modules have to be adapted regarding to the new context. The adaptation includes different changes and transformations of the module, for example an adaptation of the design or a translation. Finally the single modules the new Learning Resource should consist of have to be aggregated into a new Learning Resource.
But Re-Purposing of Learning Resources is not easy to support, different challenges arise. First of all the Learning Resources can be very complex, consisting of multiple files in multiple formats. Also the Re-Purposing process itself is complex as well. For the most adaptations like an adaptation of the semantic density experiences are necessary.

To adapt those Learning Resources usually not only different tools for different adaptations are needed (e.g. a graphics tool for layout adaptation and a word processor for translation), also different editors have to be used for editing the different formats (e.g. a text editor for a plaintext document and a html editor for a text in a document written in html). Additionally every single file contained in the Learning Resource has to be edited manually in the majority of the cases.
To ease the Re-Use of Learning Resources different tasks have to be considered. In most of the scenarios when Learning Resources are re-used, the author is a domain expert and has the didactic basics to design a Learning Resource, but only little technical knowledge. Therefore it is crucial for this author when he should be able to re-use existing Learning Resources to get support for the technical realization of the changes he wants to perform. Additionally it can be very helpful to give some guidance to this author, how the adaptation he wants to perform is executed best and what other things have to be considered resulting out of already performed changes.
Based on these requirements we have created a framework which is able to support the different adaptations and can handle the different formats. The framework is working resource spanning and includes semi-automation of adaptation processes. This concept is designed to enable novice users to perform Re-Purposing of Learning Resources and to lower the effort for expert users.
One of the base components of this framework is a representation of the Learning Resource that should be re-purposed. The representation is abstracting from the details of the Learning Resource to uncouple the user from document boundaries and format specifics. To design an appropriate representation we first analyzed the design process and considered all necessary design criteria. As a result we will introduce a generic design process in chapter 2. In chapter 3 we show how we used this generic process for creating a representation of Learning Resources which is suited to support Re-Purposing. After that we will shortly introduce the implementation of our framework in Chapter 4 and finally talk about conclusions and future work in chapter 5.

2 A generic process for the design of representations

While talking about representations first has to be defined what a representation means for us. A representation is a model of something and includes information about the thing it is representing. But the representation is not the thing itself, it only knows about certain properties of the thing which it is representing.
In "What is a Knowledge Representation" (Davis, Shrobe, Szolovits 1993) it is stated that this question is answered best by looking at the different roles of a knowledge representation (KR). These roles are:

- "A KR is a surrogate"
  "A KR is a set of ontological commitments"
- "A KR is a fragmentary theory of intelligent reasoning"
- "A KR is a medium for efficient computation"
- "A KR is a Medium of human expression"

A knowledge representation is not only the abstraction of a thing but defines much more. The representation is used to think about something, it is an internal substitute of an external thing. Normally before taking some action in the real world, we first consider what we will do using an abstraction of the real thing in our mind. Furthermore the representation defines how to view the world and which results we can and should get from what we know. This is also a very important point because beside the fact that a representation always represents only certain properties it also defines which conclusions we can draw and which information we can get using the representation. So for different tasks different representations are needed. Representations are applied in many areas because they are used to compute with them (e.g. a climate model is needed when computing the climate development) and they are a medium of expression and communication.
But nevertheless the only completely accurate representation of an object is the object itself. There is no one perfect representation of an object. Because of that and the strong correlation between the representation and the application it is used in, every challenge needs an appropriate representation.
We analyzed how such a representation can be designed and what are the different points that have to be considered. Based on this analysis we developed a generic representation development process which can be used for the design of specific representations (Figure 1).
The representation development process consists of five steps and can be described using the questions a representation modeler is confronted with during the process. The first three steps are the definition phase of the representation. The last two steps are the creation and evaluation.

Definition of the representation:

- What should be represented?
  What are the goal(s) of the representation?
  What is the vocabulary of the representation?

Creation and evaluation:

How should the representation be done?
Does the model suits its purpose?

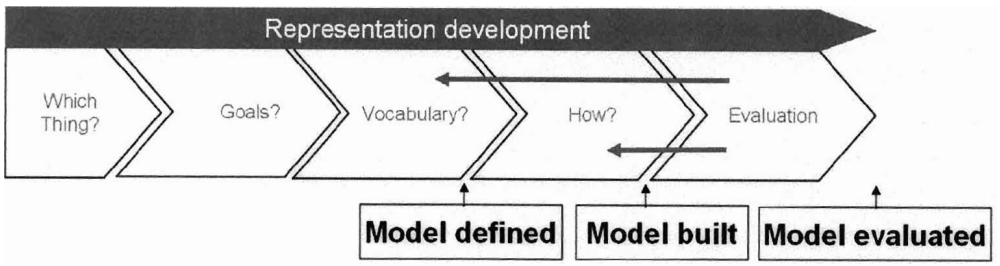

Figure 1: Generic representation development process

In the definition phase the first thing which has to be considered is what we want to represent. The thing which should be abstracted must be specified, it must be clear what are its properties and features.
In the second step the information which is needed about the thing has to be defined. The application which should use the representation must be specified to know which of the properties of the defined thing are needed and therefore have to be considered in the model. The model is the resulting representation when abstracting a certain instance of the defined thing. For example when we want to build a representation of Learning Resources the abstraction of a certain WBT (Web Based Training) is the model of exactly this WBT.
The last step of the definition phase is the specification of the vocabulary which is used in the representation. The vocabulary has to be suited to describe the properties identified in the second step and their relations so that a model has the power to describe all needed information.
After the definition phase what can be described and how it can be described is defined. Now the next question for the modeler is how to do the mapping from the real thing into the model. Methods and algorithms must be defined which are suited to get the required information about a certain object and express them in the language of the representation.
If all this is done then it is possible to create models of certain instances of the defined thing. So the last step to do is to evaluate the created models. Therefore it has to be checked whether a model holds all required information for the goals defined in the design phase. The drawbacks identified in the evaluation have to be investigated. Drawbacks can result for example out of an unsuited vocabulary or a wrong or bad mapping. Through an iterative improvement of the crucial parts the resulting model can be improved until it fulfils the requirements given through the goal(s) defined in the second step of the definition phase.
This general development process can be applied for the creation of different representations. The next chapter describes in detail how we applied this general process to develop a representation of Learning Resource which is suited to support Re-Purposing.

3 A representation of Learning Resources to support Re-Purposing

As stated in the beginning the challenge in supporting Re-Purposing is to span the gap between the user who wants to perform different adaptations while re-purposing a Learning Resource and the Learning Resource which consists of multiple files in multiple formats. This can be solved by abstracting the user from the details given in the Learning Resources and presenting only the relevant information to her. Therefore we are using a representation to support the user in Re-Purposing of Learning Resources.

We developed our Learning Resource Content Representation (LRCR) according to the workflow given in the last chapter. So now we will go into detail about the single steps of the development process.

3.1 What should be represented?

First of all it must be specified what exactly is it, we want to build a representation of. In our case we want to represent Learning Resources, which are defined as: "A Learning Resource is a digital resource used for E-Learning" (see Rensing et al. 2005). For the implementation in the Re-Purposing framework (see chapter 4) we focus on Learning Resource Content Packages (IMS 2006b). Content Packaging is a transport format for the packaging of content for interoperability which was defined by the IMS Global Learning Consortium (IMS 2006a) and is used in SCORM (Shareable Content Object Reference Model) (SCORM 2004). The Shareable Content Object Reference Model is an ADL (Advanced Distributed Learning) (ADL 2006) standard for E-Learning courses and is widely used in the E-Learning community. The IMS Content Package can include multiple files in multiple formats, e.g. html, xml, jpeg, or MS Office formats. Contained in the Content Package are the single resources, an organization describing the structuring of the resources and LOM (Learning Object Metadata) (IEEE 2002) about the whole course and the single resources. As we will describe later it is important for the representation to consider which features are explicitly given through the way the Learning Resource is composed (for example the organization which is already given can be transferred and used in the representation).

3.2 What are the goals of the representation?

In the second step of the development process the target tasks must be defined to know which information about the Learning Resource is needed. As stated before every representation is build for a special purpose because the universal representation is impossible to create. Our goal is to support Re-Purposing of Learning Resources. To do this we need an abstraction of the Learning Resource which can provide all the required information needed in the Re-Purposing process. The Re-Purposing process consists of three parts. These parts are modularization, adaptation and aggregation.
We have done different analysis to find out which information about the Learning Resources is relevant when doing modularization, adaptation and aggregation of Learning Resources. For example we have done a user survey about Learning Resource adaptations (Zimmermann et al. 2006a). There we have analyzed how adaptations are performed, if standard workflows for adaptations exist, what the relevant information while performing an adaptation process is and so on. Important for the user are for example the different elements which are contained in the Learning Resource as well as a description of these elements. The needed information is typified in the representation through the used vocabulary which is defined in the next step.

3.3 What is the vocabulary of the representation?

After defining what exactly should be represented and which information about the defined thing is needed, the next step is to define the vocabulary of the representation. The vocabulary must be suited to describe the Learning Resource but it also must suffice the specified goals (for example it must be possible to describe the information about the Learning Resource which is needed for the adaptations).
A conceptualization can be defined to specify the vocabulary and the relations which are used in the representation. This specification can be done using an ontology. An ontology is a sharable conceptualization of a knowledge domain (see Gruber 1993). It includes concepts, which are the entities of a knowledge domain, and relations, which are semantic interconnections between concepts.
We have identified several requirements an ontology for the Re-Purposing of Learning Resources must fulfill. For example a concept for every relevant element of a Learning Resource is needed and the description has to be format independent. After analyzing related work we could not find an ontology which suited our requirements so we started to build our own Content Ontology for the Re-Purposing of Learning Resources. (Bergsträßer et al. 2006) For the ontology creation we choose an iterative development process. In the first phase we did a requirement analysis, specified the characteristics and collected concepts for the ontology. After that an initial Content Ontology was build. Then the iterative improvement phase followed by collecting criticism and doing an evaluation by integrating the Content Ontology into the Re-Purposing framework.

While investigating the Learning Resources we identified different characteristics of a Learning Resource which are of importance. The first one is the structure that means the arrangements and connections of the different included resources (the organization of the Learning Resource in the Content Packages) and the organization and structuring inside the resources. The second one is the semantic which is the meaning of the different parts a resource consists of (e.g. definition, example, introduction, summary, etc.).
With the creation of the ontology the definition phase of the representation development process was completed.

3.4 How should the representation be done?

To build an actual model of a certain Learning Resource we need to define the algorithms for the mapping from the real word into the representation. The mapping algorithms should extract the required information out of the Learning Resource and use the specified vocabulary to map it into the representation. For the mapping of the Learning Resources into the Learning Resource Content Representation, we are using three different mapping methods. As the base for the representation we are using the structural concepts because there are more obviously given in the Learning Resources. We use parsing of the Learning Resource to extract the specified structure and elements. After parsing the Learning Resources a simple model based on the organization of the different resources and the internal composition of the resources is created. The second part of the mapping is done as an enrichment of the simple model by adding semantic information to the entities in the model. There are two kinds of mapping for the enrichment. The first one is to use semantic analyses to explore the meaning of an identified entity and the second one is user interaction. User interaction means that the user declares the meaning of a certain entity (e.g. an image is a logo).
Now we want to give an example of how the Learning Resource Content Representation of a specific Learning Resources is created by looking at a small cut out of the model (Figure 2).

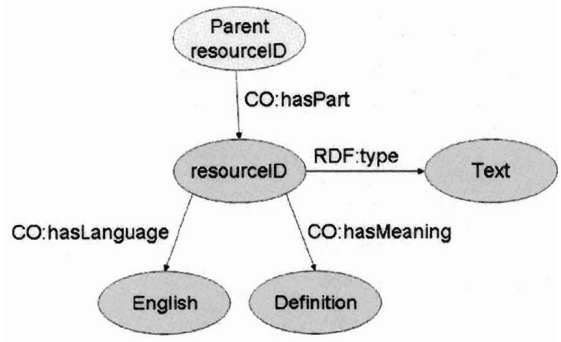

Figure 2: Example for an English text which is a definition

In Figure 2 an example how the representation of an English text could look like is given. Every resource is addressed by an unique resource ID. The example resource in this case is contained in another resource ("Parent resource"). The example resource is of the type "Text" and its language is "English". By analyzing the resource also its property was determined and the representation has been enriched with this information (the resource is a "Definition").
All the concepts and relations we are using in the created model are defined in the Content Ontology and only the vocabulary which was defined there can be used in the model. All the information which we have included in the model can now be used in the Re-Purposing process.
After building the model the next step is to check if the model which was created is really suited for the identified purpose.

3.5 Does the model suits its purpose?

The last step of the representation development process is the evaluation. The evaluation must not only be understood as an evaluation of the created representation, it is a review of the whole development process. By looking at the finally created models there must always be in mind that they depend on the different decisions taken during the development process.
Because the representation was created for a specific purpose to evaluate the representation it must be checked whether the model includes all information needed for the specified goals. During the evaluation not only the model itself is revised additionally it is checked if the concepts and the relations defined in the conceptualization are sufficient. Also the mapping is checked because a correct mapping is crucial for a good representation. The evaluation can be used to find drawbacks and to improve the vocabulary, the mapping algorithms and as a conclusion the created models of the Learning Resources.
To evaluate the representation we have implemented it as part of a Re-Purposing tool which we developed based on our Re-Purposing framework.

4 Implementation

The Re-Purposing framework (Figure 3) consists of three layers (Re-Purposing Layer, Abstraction Layer and Physical Layer).

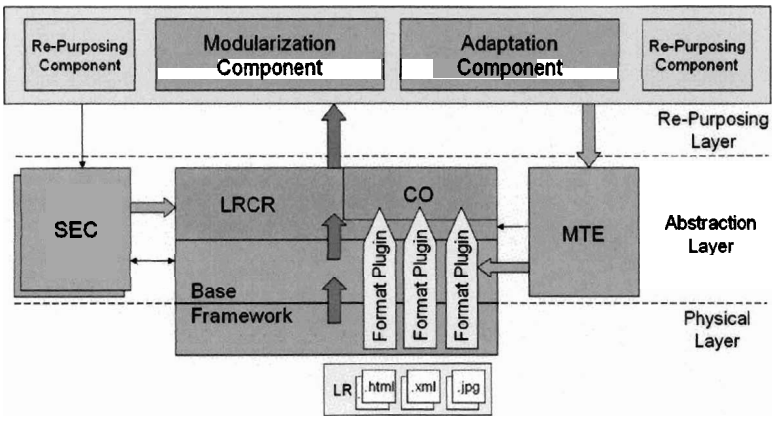

Figure 3: The Re-Purposing framework

The top level layer (Re-Purposing Layer) is composed of different Re-Purposing components which interact with the user (for example the Adaptation Component, which is responsible for supporting the user in performing adaptations of a Learning Resource or the Modularization Component, which performs the modularization of the Learning Resources into appropriate modules). Functionalities of this layer which are related to the Learning Resource Content Representation are for example to find out which information the user needs and to provide this information to her.
The second layer (Abstraction Layer) contains the core components of the framework. For example the Learning Resource Content Representation (LRCR) and the Content Ontology (CO) are a part of this layer, as well as the Semantic Enrichment Component (SEC) which is in charge of performing different semantic analysis of the Learning Resources.
The third layer is the "Physical Layer". There all the "physical" interactions with the Learning Resources (LR) take place (read/write, etc.). For further information about the implementation of the Re-Purposing framework see (Meyer et al. 2006).

4.1 Implementation of the Content Ontology

The Content Ontology was defined in the Web Ontology Language (W3C 2004a) under the use of the ontology creation tool Protégé (Stanford 2005). The Content Ontology is included in the base framework of the Re-Purposing tool and provides the vocabulary which can be used in the Learning Resource Content Representation. The concepts defined in the structure part of the ontology are the base types for the Content Representation of Learning Resources. The Content Representation is enriched with the concepts defined in the semantic part of the ontology and with additional attributes. The semantic analyzing methods which are performed by the Semantic Enrichment Component (SEC) are also based on the Content Ontology.

4.2 Implementation of the Learning Resource Content Representation

The Learning Resource Content Representation is part of the base framework and implemented as an RDF-model (Resource Description Framework) (W3C 1999). The RDF-model holds the actual representation of the Learning Resource. Different mapping mechanisms are implemented to create the model. Mainly parsing of the Learning Resource and semantic enrichment of the model are used to map the information contained in the Learning Resources into the model. The parsing of the Learning Resource is done using format plug-ins to deal with the different formats of the resources and map them into the format independent vocabulary defined in the Content Ontology. The semantic enrichment is done by the Semantic Enrichment Component and through user interaction which takes place on the Re-Purposing layer. For the semantic enrichment we focus on text analyzing methods to annotate the different text fragments of the Learning Resource with their meaning. For example we developed heuristics for the analysis of introductions, definitions, examples and summaries and methods for topic detection. For the graphical elements we mainly use user interaction. This means for example images which could be a logo are presented to the user and he can then decide whether they are or not.
The created model is used by the other components in the framework to search for required information and to find certain elements (e.g. definitions). To query the model there are some frequently used queries implemented and beside this a RDQL (RDF Data Query Language) (W3C 2004a) interface is provided.

4.3 The Re-Purposing tool

With the implemented Re-Purposing tool we focus on the support of users with a pedagogical background (e.g. teachers, professors) who don't want to be confronted with technical issues and users which are no experts in doing Re-Purposing but have to Re-Use and adapt Learning Resource (for example in small and medium size enterprises). Also we want to support users with technical knowledge by reducing the effort needed for Re-Purposing of Learning Resources for them.
Our Re-Purposing tool covers the whole Re-Purposing process. The first step in a Re-Purposing process is the modularization. Using the modularization component a Learning Resource can be segmented into modules of different granularity. A modularization of the Learning Resource is proposed by the tool and can be adjusted using defined parameters or can be done manually based on the information about the structure of the Learning Resource and the semantic of its elements (for example parts which are or include definitions are tagged). To generate the semantic tagging the information in the Learning Resource Content Representation is used. After the modularization modules can be adapted to suit the new context. The support of the adaptation process is based on patterns (Zimmermann et al. 2006b) and includes guidelines, step by step workflows and semi-automation. For the adaptation of the Learning Resource all relevant information about a module is provided. Therefore again the information in the Learning Resource Content Representation is used and presented to the user. For example for a translation the language or languages which are used in the module and the meaning of the elements is provided (for example to enable a translation with source language English and target language German, but all definitions should remain in English). Furthermore there is information about the topic of the module or parts of it and so on. In the last step, the aggregation of the modules into a new Learning Resource, all the provided information again can be used to find the fitting modules and finally achieve the Learning Resource the user had in mind.

5 Conclusion and Future Work

We have developed a generic representation development process and applied it to create a representation of Learning Resources which is suited to support Re-Purposing. For a proof of concept we have implemented the Content Ontology and the Learning Resource Content Representation into a Re-Purposing tool we have developed based on the introduced Re-Purposing framework. The Content Ontology provides the vocabulary definition for the representation and the Learning Resource Content Representation provides the model of a Learning Resource the user wants to re-purpose.
After implementing the Learning Resource Content Representation and the Content Ontology we have started the evaluation phase. We have performed different tests in our application scenario. The results show that information which is provided through the Learning Resource Content Representation can be used to support the user in doing modularization and adaptation of the Learning Resources.
The additionally provided information can help the user in the modularization of the Learning resource. The user is able to find the parts which are relevant for him and also the parts he wants to adapt during the Re-Purposing process more easily.
Because the implementation of the adaptation is pattern based and includes the information provided through the Learning Resource Content Representation the user can be guided and supported exceedingly through the different adaptation processes. The patterns are also format independent and they are easy to understand. They can provide the expert knowledge which is necessary to perform the adaptation processes (Zimmermann et al. 2006c). Because they are written in natural language they are well to understand and simple to provide. Patterns contain descriptions in form of problems, solutions and examples so the presented information can be adapted to users with different experience levels. Using the combination of patterns and the Learning Resource Content Representation we have achieved a user support which aids the users in performing adaptation of Learning Resources.
For the future we have planned some further improvements and user studies to qualify the profit of our approach.

6 References

Advanced Distributed Learning Initiative (ADL). (2006).

Advanced Distributed Learning Initiative (ADL). (2004): Sharable content object reference model (SCORM) 2004.

Bergsträßer, S., Faatz, A., Rensing, C., Steinmetz, R. (2006): A Semantic Content Representation Supporting Re-Purposing of Learning Resources. I-Know, September 2006.

Davis,R., Shrobe, H., Szolovits, P. (1993): What is a Knowledge Representation. AI Magazine; 14 (1): 17-33.

Gruber, T. R. (1993): Toward principles for the design of ontologies used for knowledge sharing. In N. Guarino and R. Poli, (Eds.), International Workshop on Formal Ontology, Padova, Italy.

IEEE LTSC (2002): Draft Standard for Learning Object Metadata.

IMS Global Learning Consortium. (2006a).

IMS Global Learning Consortium. (2006b): IMS Content Packaging Specification V1.1.4.

Meyer, M., Hildebrandt, T., Rensing, C., Steinmetz, R. (2006): Requirements and an architecture for a multimedia content re-purposing framework. In Proceedings of the First European Conference on Technology Enhanced Learning, October 2006.

Rensing, C., Bergsträßer, S., Hildebrandt, T., Meyer, M., Zimmermann, B., Faatz A., Lehmann, L., Steinmetz, R. (2005): Re-Use, Re-Authoring, and Re-Purposing of Learning Resources - Definitions and Examples. Technical Report KOM-TR-2005-02, Darmstadt University of Technology.

Stanford (2005): Protégé 3.1.1 an ontology editor and knowledge-base framework.

W3C (1999): Resource Description Framework (RDF).

W3C (2004a): RDQL - A Query Language for RDF.

W3C (2004b): Web Ontology Language (OWL).

Zimmermann, B., Bergsträßer, S., Rensing, C., Steinmetz, R. (2006a): A requirements analysis of adaptations of re-usable (e-learning) content. In AACE EDMEDIA, June 2006.

Zimmermann, B., Rensing, C., Steinmetz, R. (2006b): Patterns for Tailoring E-Learning Materials to Make them Suited for Changed Requirements. In: VikingPLoP 2006, Helsingör, Dänemark, October 2006.

Zimmermann, B., Rensing, C., Steinmetz, R. (2006c): Format-übergreifende Anpassung von elektronischen Lerninhalten. In: Mühlhäuser, Max; Rößling, G., Steinmetz, Ralf (Edits.): DeLFI 2006 - die 4. e-Learning Fachtagung Informatik 2006, p. 15--26, GI Lecture Notes in Informatics , September 2006.