[BFRSt06] *Sonja Bergsträßer , Andreas Faatz, Christoph Rensing, Ralf Steinmetz;* **A Semantic Content Representation Supporting Re-Purposing of Learning Resources;** 6th International Conference on Knowledge Management (I-Know 06), Graz, September 2006, S.

A Semantic Content Representation Supporting Re-Purposing of Learning Resources

Sonja Bergsträßer
(TU Darmstadt, Germany
Sonja.Bergstraesser@KOM.tu-darmstadt.de)

Andreas Faatz
(SAP Research CEC Darmstadt, Germany
Andreas.Faatz@sap.com)

Christoph Rensing
(TU Darmstadt, Germany
Christoph.Rensing@KOM.tu-darmstadt.de)

Ralf Steinmetz
(TU Darmstadt, Germany
Ralf.Steinmetz@KOM.tu-darmstadt.de)

Abstract: Because of the costly production of Learning Resources the Re-Use of existing Learning Resources becomes more and more important. But reusing Learning Resources in a new context makes it necessary to adapt them. We have developed a framework as base of a Re-Purposing Tool to support users to perform adaptations of Learning Resources in different dimensions which are crucial when using existing Learning Resources in a new context. The adaptation of Learning Resources is not easy to perform and 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 model 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 content representation which is used in the framework to abstract from the given Learning Resources as well as the content ontology which the content representation is based on.

Keywords: Content Representation, Ontology, Re-Use, Re-Purposing, Learning Resource
Categories: H.3.1, H.3.2, H.3.3

1 Introduction

The creation of high quality Learning Resources is a cost and time consuming task. Especially the creation of multimedia content can be arbitrarily complex. Many Learning Resources have already been created and used in the specific scenarios they have been designed for. To use them in a new scenario especially in a new teaching or learning context they have to be adapted to the new context. The context can be described by different characteristics of the participants in a learning or teaching process and of the learning or teaching process itself. These adaptations have to be done manually operated until now, which is almost consuming like the production of

new Learning Resources. So many resources are lost. To keep these resources we want to support these adaptations by a tool and developed corresponding concepts and a tool framework. We call it Re-Purposing of Learning Resources which is defined as the adaptation of a digital resource used for E-Learning to a new context, especially a new learning objective or target group [Rensing, 05].

Re-Purposing of Learning Resources is very complex. A Learning Resource even if developed according to standards like SCORM [SCORM,04] can include resources in different formats, e.g. HTML, XML, MS Office, Flash, etc. Also Re-Purposing is dealing with various kinds of adaptations: design adaptations, technical adaptations and content adaptations. For our analysis of adaptation processes based on expert interviews see [Zimmermann, 06]. To adapt Learning Resources in a consistent way and independently from all these different requirements a representation is needed which can cover the main aspects common in all the different flexible characteristics of Learning Resources. For this reason we have developed a conceptualisation of Learning Resources as a base for a representation of Learning Resources which is used to guide the user through the Re-Purposing process and present all relevant information to her.

This paper mainly focuses on the conceptualization used as base for the Learning Resource Content Representation (LRCR). Therefore we start in chapter 2 with a brief introduction of our Re-Purposing framework and the characteristics of the content representation. Chapter 3 is dealing with related conceptualizations. In chapter 4 we will show details of our content ontology and finally give a summery of the paper in chapter 5.

2 A Learning Resource Content Representation as part of a Re-Purposing framework

To support Re-Purposing of Learning Resources we are developing a framework for a Re-Purposing tool. The main challenge of this framework is to deal with the multiple characteristics of Learning Resource and adaptations. On one side of the framework there are the Learning Resources which should be Re-Purposed and the framework has to manage all the different files contained in the Learning Resource as well as the different formats which these files can have. On the other side of the framework is the user who wants to perform adaptations of different kinds, e.g. translation, change of the corporate design or increasing of the semantic density of the Learning Resource [Zimmermann, 06]. To uncouple the adaptations from the different files and formats and hide the physical details from the user a representation of the content of the Learning Resources is needed. Based on the representation we can support the user and guide her through the different kinds of adaptations she wants to perform. So with our Re-Purposing tool even novice users can do a Re-Purposing of Learning Resources. We are now introducing our framework and the requirements we identified for the Learning Resource Content Representation.

2.1 A Re-Purposing Framework for Learning Resources

The framework consists of different layers. The first layer of the framework is the Re-Purposing Layer. It includes the Re-Purposing components of the framework. The Re-

Purposing components are interacting with the user and guiding as well as supporting her during the different steps of a Re-Purposing process, like modularization, adaptation or aggregation. Hidden from the user is the second layer of the framework, the Abstraction Layer. The core components located in the Abstraction Layer include all mechanisms needed to build, extend and query the representation of the Learning Resource, the conceptualization which is the basis of the representation and the management for the modifications which will be performed on the Learning Resource. The third layer is the Physical Layer. The Physical Layer includes all the read and write actions performed on the Learning Resources (figure 1). Details about the framework can be found in [Meyer, 06].

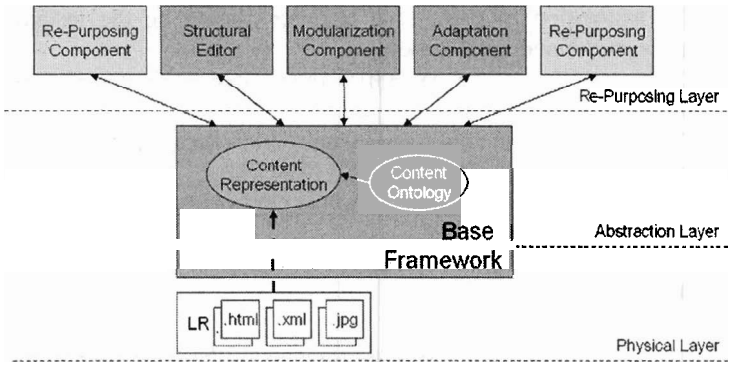

Figure 1: The Re-Authoring framework

2.2 Requirements for the Content Representation

To support the user in a Re-Purposing process a representation of the processed Learning Resources is needed. This representation has to deal with all the elements a Learning Resource may consist of. It needs to be able to deliver information about the Learning Resource to the components in the Re-Purposing Layer, for example the occurrence of company logos for an adaptation of the corporate design. That means we need a mapping from the Learning Resource into a model which can provide all required information. We identified several requirements which such a model must satisfy. The model should give a format independent representation of a Learning Resource and the model should allow a mapping of an arbitrary Learning Resource. The model should contain all information which is relevant for the adaptation components and queries on the model should be possible and easy to perform. Based on the requirements we decided to use the Resource Description Framework (RDF) [RDF, 99] for the LRCR. With an RDF-Model it is possible to create a content representation of a Learning Resource which fulfils all afore mentioned requirements on the model.

2.3 Requirements for a Ontology as base for the LRCR

After specifying the LRCR there was a lack of descriptive power to generate a model of a Learning Resource. A conceptualization of Learning Resource which can be re-

purposed as a base for this model is required. A conceptualization of a knowledge domain can be expressed with an ontology [Gruber, 93].

In a Re-Purposing process resources in multiple formats are involved. It must be possible to find an instantiation of a concept defined in the ontology for all of them. Also the meaning of the different parts of a Learning Resource should be describable if it is relevant for the Re-Purposing process. The requirements for the ontology are based on the requirements of the LRCR, because the ontology is the base for the model of the Learning Resources. The main requirements for the ontology are:

- The ontology should be the base for a format independent creation of Learning Resource models.
- The annotation of parts of a resource should be the instantiation of the concepts defined in the ontology.
- The vocabulary for a semantic representation of the resource should be provided.
- The ontology should also contain attributes which can be used as properties.
- Queries on the instantiation should be possible.
- The ontology should be extendable to future developments.

3 Related Work

To find an ontology which suits the requirements of the LRCR we looked at existing conceptualisations. We did not take domain ontology's into consideration, because we need a conceptualization of Learning Resources and not one of a scientific knowledge domain. Structured Writing [Horn, 98] is a way to produce text, which comes along with a mark-up of its building blocks (semantically and rhetorically). Nevertheless, structured writing is an approach on the producing side (new content), Re-Use is only possible, if every author obeys this modular approach from the beginning. The Darwin Information Typing Architecture (DITA) [DITA, 05] was designed for a flexible Re-Use of technical documentation. DITA is implemented specified via XML. Its central element is the conceptual unit called "topic", which refers to a thematic entity. However the shift from technical documentation to the application domain of E-Learning goes along with obstacles for the concept of "topic" in the sense of DITA. It does not fully match the structures we obtain from pedagogical and especially rhetorical and layout driven modularisation. ALOCoM [ALOCoM, 05] extents DITA, the formats of this extension are ontology related [OWL, 04], but the conceptual extension merely refers to the sense of slide presentations.

4 An Ontology for the LRCR

We developed our own ontology for describing Learning Resources, because there is no ontology which fulfils our requirements. Now we will introduce this ontology in detail.

4.1 Ontology development

Ontology development is an iterative process. Based on the collaborative method of ontology development [Holsapple, 02] we designed an iterative development strategy suited for our scenario. We decided to develop the ontology in the Ontology Web Language (OWL) [OWL, 04] using the ontology creation tool Protégé [Protégé, 05].

4.2 Semantic and structure

One of the main challenges during the ontology development process was the decision how to handle the conflict of having two different kinds of information. We found concepts which describe structural features of Learning Resources and concepts regarding to semantic interpretation. In related work there is no separation of concepts of this two kinds, the conceptualizations mix both occurrences. We decided to separate structure and semantic in our conceptualisation since they express different aspects of a resource. Additionally there are also parts in Learning Resources which do not belong directly to the content but they are attached information about the content. So we have three different kinds of contemplable information included in a Learning Resource. Therefore the top level concepts for describing Learning Resources in our ontology are StructureComponent, SemanticEntity and AttachedInformation (figure 2).

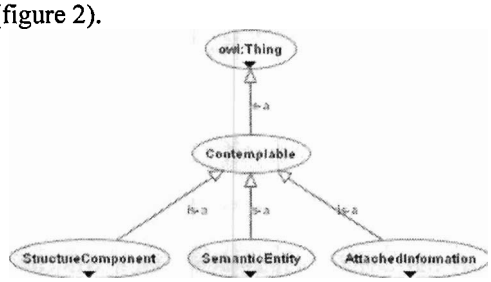

Figure 2: Top level concepts for describing Learning Resources

Definitions of top level concepts:

- ***StructureComponent*** **is a** component which includes structural information. For example a text fragment or a table.
- *SemanticEntity* is an entity which has a certain meaning. For example a definition or an example.
- *AttachedInformation* is attached information which includes information about the content but is not part of the content. For example metadata or a style sheet.

4.3 Degree of concretisation / A levelled conceptualisation

After our first collection of concepts we had specified the top level concepts and collected a lot of very concrete concepts, for example image, table, example, title or logo. These concrete concepts are of many different kinds for example containers or elements as structural concepts and pedagogical or labelling or graphical as semantic concepts. We have defined different levels of sub-concepts to group the concepts

which are of the same kind and also to allow different levels of concretisation when instantiating the concepts of the ontology. When the ontology is instantiated to represent a certain resource it is not always wanted to represent the detailed parts of the Learning Resource with their concrete meaning. Especially for the Re-Purposing of large Learning Resources, the point of interest may lie in the approximate meaning, in the first step. Later in the Re-Purposing process often only delimited areas are of special interest and a more concrete description of the parts in this special area is needed. For the structural concepts a levelled conceptualisation is quite sophisticated, but for the semantic concepts it is very difficult. We will now show by example how we defined different levels of concepts in our ontology (figure 3).

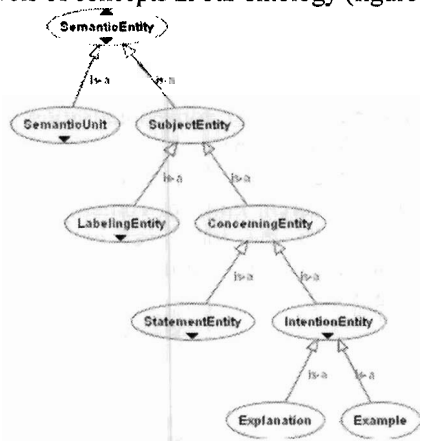

Figure 3: Levelled conceptualization

We will show the levelled conceptualization with the concept "Example". The concept "IntentionEntity" summarizes all specific concepts which are used in a Learning Resource with a special intention of the author. This means if the author of a Learning Resource wants to give an example of something she uses the concept example. The same for other concrete concepts like "Explanation" or "Description". On the next level the concept "ConcerningEntity" summarizes all concepts which belong to a certain subject and include content about this subject. This is true for all "IntentionEntitys" but also for concepts which are used to make statements ("StatementEntitys"). A concrete "StatementEntity" is for example a definition. The next level is the concept "SubjectEntity", which is an Entity which belongs to a certain subject. This is true for the "Example" but also for "LabellingEntitys" like "Titel". The top level concept "SemanticEntity" now summarizes all "SubjectEntitys" and also "SemanticUnits" which are semantic containers which include semantically correlated content, for example an introduction.

4.4 Properties and the relation of concepts

We have defined a set of properties to specify attributes to the concepts and to relate different concepts and resources which each other. We are now introducing some of the concepts which are defined in the content ontology. The semantic model is based

on the representation of the structure of the Learning Resource. So in the creation process of the semantic model we can use this information to annotate the resources in the model with the concepts representing structural properties. To include the information in the semantic model we use the RDF property RDF:type with the range StructureComponent. The other properties we can divide into three sets according to the different applications they are used for. The first class of properties are the has-part and referencing properties to indicate the relations between the resources in the model which are representing different parts of the Learning Resource. The second class of properties are used to specify different attributes to the resources in the model, for example the has-language property can be used to indicate the language of a resource representing a textual part of the Learning Resource. Information about the language can be extracted out of the metadata included in the Learning Resource or can be gained through a semantic analysis of the content. The last set of properties includes properties which annotate the resources with semantic concepts. As showed in [Chapter 4.2] we separated the concepts of structure and semantic and though we need properties to include the semantic information. For example we have specified a has-meaning property to relate semantic meaning to the resources in the model. Now we can represent a definition given in a text which is written in English as follows (figure 4):

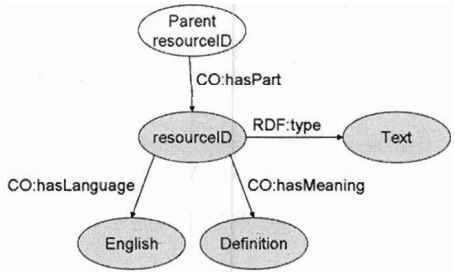

Figure 4: A representation of a definition in the Semantic Model

5 Conclusions and Future Work

After the development process had been advanced far enough we integrated the content ontology into the Re-Purposing framework. Based on the framework we realized three adaptations for two different formats so far. One is the adaptation to corporate design. The user is guided through the different steps of the adaptation process. If a logo has to be changed as part of the adaptation to corporate design the framework lists all images which are part of the Learning Resource to the user. Therefore the adaptation component (see figure 1) asks the LRCR to list all relevant images. After selection of all logos the system exchanges every occurrence of the logo independently from the format of the image and of how the image is included in the Learning Resource. The user does not care about any format specific detail how the exchange has to be performed. During the work on and with the framework we evaluated and improved the content ontology and the usability of the generated content representation. Using the Re-Purposing framework for different adaptations

of Learning Resources from different authors in different domains shows that the LRCR we are generating is suited as a format and file independent representation of Learning Resources and the Re-Purposing components can get all required information out of the model. The content ontology which is the base of the LRCR provides the required vocabulary which is needed to describe the Learning Resources, as far as we evaluated it until now, but evaluation and improvement are not finished.

Acknowledgements

We would like to thank Birgit Zimmerman, Marek Meyer, Tomas Hildebrand and Knut Manske for discussions. The work was founded by SAP Research CEC Darmstadt.

References

[ALOCoM, 05] The ALOCoM Ontology, 2005, http://memling.cs.kuleuven.ac.be/alocom/documentation.php

[DITA, 05] Darwin Information Typing Architecture (DITA), Updated 28 Sep 2005, http://www-128.ibm.com/developerworks/xml/library/x-dita1/

[Gruber, 93] T. R. Gruber, 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, August 1993

[Holsapple, 02] C.W. Holsapple, K.D. Joshi, A collaborative approach to ontology design, Communications of the ACM, February 2002, Vol. 45, No. 2, 42-47

[Horn, 98] R.E. Horn, Structured writing as a paradigm, In Instruct. Devel.: the State of the Art, Englewood Cliffs, N.J, 1998

[Jena, 05] Jena-2.3: "A Semantic Web Framework for Java", October 2005, http://jena.sourceforge.net/

[Meyer, 06] M. Meyer, T. Hildebrandt, C. Rensing, R. Steinmetz, Requirements and an architecture for a multimedia content re-purposing framework, in Proceedings of the First European Conference on Technology Enhanced Learning, October 2006

[OWL, 04] Web Ontology Language (OWL), 2004, http://www.w3.org/TR/owl-features/

[Protégé, 05] Stanford: "Protégé 3.1.1 an ontology editor and knowledge-base framework", 2005, http://protege.stanford.edu/

[RDF, 99] Resource Description Framework (RDF), 1999, http://www.w3.org/RDF/

[Rensing, 05] C. Rensing, S. Bergsträßer, T. Hildebrandt, M. Meyer, B. Zimmermann, A. Faatz, L. Lehmann, R. Steinmetz, Re-Use, Re-Authoring, and Re-Purposing of Learning Resources - Definitions and Examples, Technical Report KOM-TR-2005-02, http://www.kom.e-technik.tu-darmstadt.de/publications/abstracts/RBHM+05-1.html

[SCORM, 04] Sharable Content Object Reference Model (SCORM), 2004, http://www.adlnet.gov/scorm/index.cfm

[Zimmermann, 06] B. Zimmermann, S. Bergsträßer, C. Rensing, R. Steinmetz, A Requirements Analysis of Adaptations of Re-Usable (E-Learning) Content, in Proceedings of World Conference on Educational Multimedia, Hypermedia, June 2006