Improving Authoring-by-Aggregation and Using Aggregation Context for Query Expansion

Marek Meyer[1,2], Christoph Rensing[2], and Ralf Steinmetz[2]

[1] SAP AG, SAP Research CEC Darmstadt, Bleichstr. 8, 64283, Germany
marek.meyer@sap.com
[2] KOM Multimedia Communications Lab, Darmstadt University of Technology,
Merckstr. 25, 64832 Darmstadt, Germany
{rensing,steinmetz}@kom.tu-darmstadt.de

Abstract. Authoring-by-Aggregation is an authoring paradigm for creating new Learning Resources by composing several smaller ones. However finding suitable Learning Resources that can be reused in a particular course is a time-consuming task. The user has to query a repository and review the resulting list of Learning Resources whether they are applicable. In this paper we present a new method for narrowing the result set by taking the aggregation context into account. Context features are used as additional query attributes, leading to more precise query results. This paper also presents an improved Authoring-by-Aggregation process and a corresponding implementation, which facilitates the creation of new Learning Resources.

1 Introduction

The use of E-Learning for education and advanced training has grown over the past years. Many E-Learning technologies are based on digital materials, also called Learning Resources. Some of these materials are produced based on didactic guidelines, following best practices. Especially large content producers rely on content models as blueprints for didactically well-structured courses. There are different kinds of authoring process; one of them is the Authoring-by Aggregation (AbA) paradigm.This paper is mainly based on the AbA reference process introduced by Hoermann [1].

Learning Resources that are supposed to be reused have to be made accessible to other users. Learning Object Repositories (LOR) are storage systems for Learning Resources. A LOR typically offers retrieval interfaces, which can be used to search for and obtain Learning Resources. Recent initiatives work towards standardization of these interfaces, for instance the Simple Query Interface (SQI) [2] that is based on web service technology. However, most known implementations of SQI use a very limited query language, which allows to search for plain text search terms only, but not for specific metadata fields.

Authoring-by-Aggregation systems are currently designed only to work on a single - primarily local - repository. Access to different remote repositories is not yet supported. Thus, a user who wants to reuse Learning Resources from other

repositories has to search for them manually, download them and afterwards load them into his authoring tool. However, it would be more intuitive if the retrieval of Learning Resources were integrated into the authoring tool.

Current authoring tools and repositories have shortcomings regarding efficiency of AbA. It is time consuming to search a repository for many different small contents, which might be reusable for authoring a new Learning Resource.

In order to improve efficiency of Authoring-by-Aggregation, this paper proposes the use of aggregation context information for rewriting queries. Furthermore, an improved AbA process has been developed, which enables more intuitive and efficient retrieval and integration of existing Learning Resources.

The structure of this paper is as follows. Section 2 presents a definition of aggregation context and explains how it can be determined and used. An improved AbA process, which will be used as foundation of our prototype is shown in section 3. An implementation of the approach is presented in section 4. Finally, some conclusions and an outlook are given.

2   Context in Authoring Environments

As mentioned in the previous section retrieval methods for Learning Object Repositories still lack some functionality. This section focusses on how the aggregation context of an authoring process can be used for query refinement.

Search results of LOM queries may become quite large. Especially if a user searches for multiple Learning Resources one after another, the required time adds to the production costs and may make reuse of Learning Resources inefficient. Query expansion - rewriting a query based on additional knowledge - could improve the quality of retrieval results and thereby reestablish the economical benefit of reuse.

Context information from authoring environments and relevance feedback mechanisms are promising candidates for query expansion. Context is defined here as any information that is known about the author, his authoring environment and the tasks he is performing. Different types of context information are imaginable, particularly system context, explicit project context and aggregation context (see Fig. 1).

As system context we subsume all information about the user, the tools and systems he uses and what he has generally done in the past. Exemplary context information is for instance technological restrictions of his authoring system. Explicit project context is all information, which has been explicitly specified by the author about the project he is currently working on. Project here typically refers to a particular Learning Resource the author is working on. Examples for project context are the intended course language, target document formats, target group (age, role, difficulty level, interaction level, etc.). Also, limits for the total amount of learning time or acquisition costs may have an impact on which Learning Resources are suitable for aggregation. Aggregation context is implicit information about the current project, which is deduced from the contents already existing in the project. Aggregation context information may be deduced
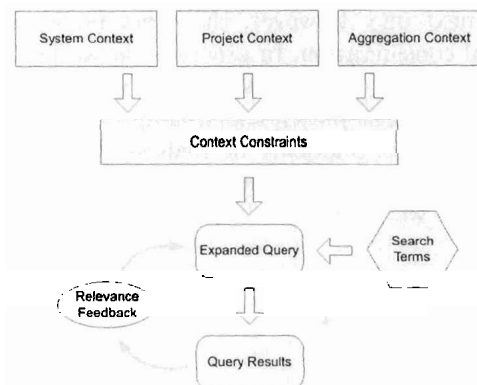
Fig. 1. Usage of context information as input for query expansion

either from the contents itself or their metadata records; in practice, metadata is easier to use. The more Learning Resources are aggregated in a project over time, the better the aggregation context may automatically be determined.

A similar approach has already been proposed by Sanchez and Sicilia [3]. They apply the design-by-contract paradigm from object oriented software construction to the composition of Learning Resources. Their goal is to automatically select and aggregate Learning Resources for a given learning target. However, because of the design-by-contract paradigm the approach supports only a binary matching: either a Learning Resource fits or it does not. Furthermore, a formal contract has to be specified for each Learning Resource; thus, the method is not applicable to plain LOM metadata. In practice, Learning Resources - and their metadata records - are rarely perfectly fitting. But Learning Resources can be adapted to fit into the new course [4]. Therefore, fuzzy queries, which produce a ranked result set, are better suited than strict matching.

3   Improved Authoring-by-Aggregation Process

Authoring-by-Aggregation is a lightweight rapid authoring approach, which is based on recursive aggregation of modular contents. One shortcoming of existing AbA tools is the missing support of didactic design and more general the non-existing separation of different authoring phases.

Therefore, this section presents an improved authoring process, which supports multi-phase authoring. The process is based on Hoermann's process [1]. It is assumed that a Learning Resource is created by a single author at a time. Existing Learning Resources are supposed to be stored in a repository and specified by adequate metadata.

In total the new process consists of five authoring phases: the didactic design phase, the retrieval and replacement phase, the adaptation phase, the content authoring phase and the publishing phase. These phases may be arranged in a strictly sequential way. Each phase has to be completed before the author

may proceed to the next one. However, the strict implementation is mainly thought for theoretical consideration. In practice, the author will not always be able to finish the phases one after another. Imagine an author has designed a course structure, replaced most placeholders, adapted the contents and is now creating the missing contents. Suddenly he realizes that an exercise is missing in his structure. The strict process would not allow to go back and change the structure. Therefore we define a relaxed process, which allows iterations of authoring phases.
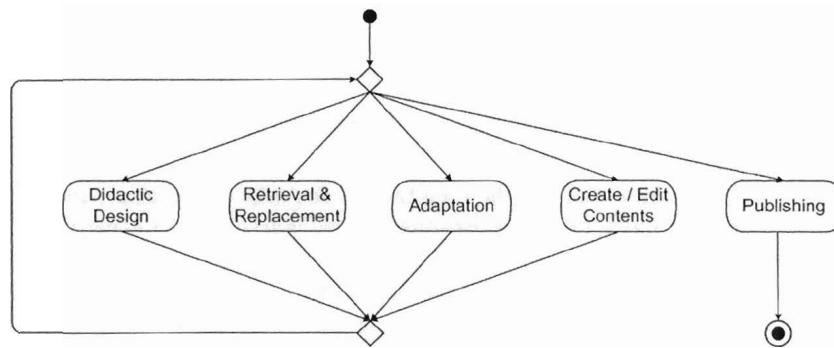

Fig. 2. Relaxed Authoring-by-Aggregation process

In the relaxed process, the first four phases may be repeated in any order. The author may choose at any time to either edit the course structure, to replace placeholder items with existing Learning Resources, to adapt a Learning Resource or to create own contents. When the author has decided that the Learning Resource is finished he invokes the publishing phase and thereby ends the process. The relaxed process is shown in Fig. 2.

4   Implementation

Having presented theoretical approaches in the previous sections, this section shows an actual implementation. The Authoring-by-Aggregation process, combined with context-based query expansion, has been implemented in the Content Sharing project [5]. It has been integrated into the Content Sharing Module Editor, which is an re-authoring tool for SCORM-based Learning Resources. In the following, the term *module* is used for SCORM-based Learning Resources, which comply to the Content Sharing specification. Particularly, the specification extends SCORM by enabling aggregation of SCORM packages by reference [6]. Besides aggregation, the tool also supports adaptation of Learning Resources to new contexts [4].

The AbA process has been realized by extending the SCORM editor of the re-authoring tool. Didactic design is supported by creating empty placeholder items in the SCORM manifest of a module. Thus, the structure of a module is
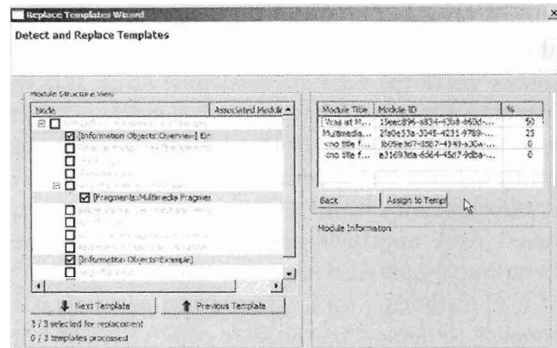
Fig. 3. Retrieval and replacement of Learning Resources

first created as a skeleton SCORM manifest without linked content files. Each placeholder element has three attributes, which may be specified by the author: keywords, granularity and didactic type. Keywords describe the topic of an element. Granularity and didactic type specify its didactic function. Didactic types can be chosen from Meder's didactic ontologies [7]. Granularity and didactic type may be selected graphically by drag&drop from a template collection.

The placeholder attributes are utilized for the retrieval of existing modules. A replacement wizard presents a list of all placeholders to the user and lets him assign existing modules to them (Fig. 3). Whenever a placeholder is selected, the connected repository is searched for suitable modules. The placeholders are then replaced by references to the chosen modules. For those placeholder elements that could not be replaced by existing modules, new content files may be created.

The retrieval of Learning Resources is based on the attributes of placeholder elements, which have been explicitly specified by the author. However, with a growing amount of modules, which have already been aggregated, there are more and more implicit constraints that impact the suitability of other modules. For instance, if most existing modules are known to be specifically designed for schoolchildren, a module from adult education will less likely fit in. Many other metadata can be used similarly to judge the suitability of further modules, such as language, format or difficulty level.

For the current implementation the four LOM fields language, end user role, typical age range and difficulty have been chosen to be determined as aggregation context. As a method for formally determining the aggregation context, a selection of metadata fields of present modules is analyzed for frequent values. A threshold of 75% is applied for considering a repeatedly occurring metadata value as context constraint.

5   Conclusion and Outlook

This paper has introduced a number of contributions to the idea of reuse of e-Learning contents. The intention of the presented approaches is to improve

the usability of reuse processes for authors in semi-professional environments. Based on a critical study of existing approaches for Authoring-by-Aggregation and Learning Resource retrieval an improved authoring process for AbA has been developed. Learning Resource retrieval still lacks some features which are common in (Multimedia) Information Retrieval. One of these features is query expansion - modifying queries for improving the performance. This paper has shown how context information can be used for query expansion.

Both, the improved AbA approach and query expansion based on context information have been implemented within the Content Sharing project as integral components of a re-authoring tool. The implementation is only considered as a first proof of concept to demonstrate the feasibility of the approach. Further usability studies of the re-authoring tool are planned for the future.

However, still a lot of research has to be done in the area of Learning Resource retrieval. In particular, transferring methods from Multimedia Information Retrieval to the area of e-Learning may lead to further improvements.

Acknowledgments

References

1. Hoermann, S.: Wiederverwendung von digitalen Lernobjekten in einem auf Aggregation basierenden Autorenprozess. PhD thesis, TU Darmstadt (2005)
2. Simon, B., Massart, D., van Assche, F., Ternier, S., Duval, E., Brantner, S., Olmedilla, D., Miklos, Z.: A simple query interface for interoperable learning repositories. In: Proceedings of the 1st Workshop on Interoperability of Web-based Educational Systems, Chiba, Japan, CEUR, pp. 11–18 (2005)
3. Sanchez, S., Sicilia, M.A.: On the semantics of aggregation and generalization in learning object contracts. In: Proceedings of the 4th IEEE International Conference on Advanced Learning Technologies (ICALT '04), IEEE Computer Society Press, Los Alamitos (2004)
4. Zimmermann, B., Rensing, C., Steinmetz, R.: Format-übergreifende Anpassungen von elektronischen Lerninhalten. In: Proceedings of the Deutsche e-Learning Fachtagung Informatik (DeLFI) 2006 (2006)
5. Content Sharing project: [last accessed, 2007-03-21] Project website http://www.contentsharing.com
6. Meyer, M., Rensing, C., Steinmetz, R.: Supporting Modularization and Aggregation of Learning Resources in a SCORM Compliance Mode. In: Proceedings of the 6th IEEE International Conference on Advanced Learning Technologies (ICALT '06), pp. 933–935. IEEE Computer Society Press, Los Alamitos (2006)
7. Meder, N.: Didaktische Ontologien. Globalisierung und Wissensorganisation: Neue Aspekte für Wissen, Wissenschaft und Informationssysteme 6, 401–416 (2000)