Components for a System to Support Customisable Training Material

Asta Bäck[1], Timo Järvinen[1], Helene Juhola[1],
Michael Reiche[2], Klaus Kreulich[2], and Arved Hübler[2]

[1] VTT Information technology
[2] Chemnitz University of Technology

Abstract. This paper describes the possibilities and advantages of electronic publishing for the apprenticeship and advanced vocational training of employees in the field of digital printing. Based on the work of an EU-funded project, the technical and technological concept as well as the strategy for the production of high-quality learning contents is discussed. Current standards, like XML and SCORM, and their application in the project are examined. An outlook demonstrates how electronic publishing can be employed profitably for future training purposes in the print and media industry.

1 Introduction and Background

The work presented in this paper has been carried out in a European project funded by the EU Leonardo da Vinci II programme (a programme supporting vocational training). The project 'A System for Creating and Offering Customisable Training Materials Applied to Digital Printing' (CustomDP) is a pilot project as it includes system development.

The project consortium works with three groups of partners: system developers, educators, who do also act as content creators, and **business associates**. VTT Information Technology coordinates the project and is, together with Chemnitz University of Technology, responsible for the system development. The University of Art and Design Helsinki has created the design for the user interface and the page layouts.

Politis Research from Greece, the GOC-Centre for Education and Labour Market in the Graphic Arts and Media Sectors from the Netherlands, the KTH Royal

Institute of Technology, **a division** of Media Technology and Graphic Arts, from Sweden and the Finnish institutes AEL Centre for Technical Training and EVITech Espoo-Vantaa Institute of Technology will be involved in the functional and technical specification of the system as well as in the creation of content; furthermore, they will test the system and its content in their regular teaching operations. The European Graphic Arts/Media Industry Network (EGIN) will assist in disseminating the results. Moreover, companies such as the paper mill M-Real and the Finnish publisher and user of digital printing technology Sanoma WSOY Oyj are also members of the project consortium and support it with their specialist knowledge.

A lot of system development has been done in the field of eLearning and many systems and platforms are available in the market. One might ask whether there is any point in creating yet another one. The CustomDP system is not intended as a complete eLearning environment; it focuses on supporting the management of learning and training materials as small and independent units, also known as learning objects. The CustomDP system does also include the functionality necessary to combine chosen learning objects into aggregations that meet the needs of different learners and learning circumstances and to provide the content either as web page or PDF file.

The main justification for the project is therefore to test how the idea of learning objects works in practice, both from content creators', teachers', and learners' point of view, and to identify what kind of system support is needed to manage and utilize these objects. If successful, the learning object approach will facilitate creating and updating learning materials. In vocational training, problems related to finding relevant and up-to-date training materials are acute; this is particularly true for technical fields as technology changes quickly and training needs vary extremely. Therefore, the project fits well into the scope of the Leonardo programme.

Since the project group consists of institutions and educators working in the printing industry, the preferred subject area is digital printing. In the coming years, many people will need to be trained to master this new technology, which already gains more widespread use. As soon as the metadata vocabulary is modified to the meet the requirements of other subjects, these may be included in the database as well.

As demonstrated above, the objectives of the project are multilayered. It aims at improving the quality of apprenticeship, advanced vocational training, and university education in the field of digital printing technologies and thereby contributes to the establishment of **new job profiles** in the European print and media industry. The discussion of such sociocultural aspects is one of the aims of this conference. As new concepts and technologies for planning, creating, and utilizing learning contents are applied in the project, the technical aspects of implementing them are being discussed below.

2 Standards and Ideas employed in the Project

2.1 Learning Objects

Learning objects are regarded as solution to content creation in educational applications. Content that is organized in small and independent units can be combined in various ways; the costs for content creation are thereby reduced and the offer of learning content increased. Unfortunately, there is no definition providing a precise characterization of a learning object. The IEEE Learning Technology Standards Committee (LTSC), for instance, defines it as any entity, digital or non-digital, that can be used, reused, or referenced in technology-supported learning [3]. According to this definition, almost anything can be considered a learning object.

An important effort to facilitate content creation based on learning objects and its sharing is the US DoD's Sharable Content Object Reference Model (SCORM) initiative [1] (DoD, Department of Defence). The US DoD requires substantial eLearning and distance learning material and has started to develop a reference model that shall enable the sharing of eLearning content between applications. SCORM recognizes three main levels: assets and sharable resources (typically single files, such as image or audio files), Sharable Content Objects (main receptacle for the training content), and Aggregations (define how the content will be presented to users, a "table of contents"). However, SCORM deals with digital content only.

Ideas presented in SCORM were adopted for this project. The definition of learning object worked with here includes non-digital content items, such as printed books, of which only metadata is available; these are referred to as 'referenced content'.

2.2 Metadata

One of the main challenges during the specification phase of the project was the development of a concept to describe and manage learning objects in order to create context-specific aggregations of them. In recent years, there has been a lot of research into the field of personalized teaching materials and computer-based training that could be benefited from. As a result of these investigations some metadata vocabularies supporting the reusability of learning objects have been proposed. A summarizing overview of learning and training metadata was published by the European Committee for Standardisation Learning Technologies Workshop (CEN/ISSS WS/LT) [2].

The assumedly most widely applied metadata vocabulary is the standard for "Learning Object Metadata" (LOM) [3], which was released by the IEEE Learning Technology Standards Committee (LTSC). This standard provides a conceptual structure for metadata, a description of each entry consisting of element name, definition, data type, and field length. LOM can be used as starting point for

planning learning objects. However, it does not include any description concerning the implementation and further processing of metadata. In order to support this next level of learning applications, the IMS Global Learning Consortium published an implementation guide [4] with instructions for the technical use of LOM-based metadata.

A further development of the IMS implementation guide is the Sharable Content Object Reference Model (SCORM) specified within the framework of the ADL Initiative [1]. It "defines a Web-based learning 'Content Aggregation Model' and 'Run-time Environment' for learning objects". According to SCORM, metadata elements are grouped into the following categories:

1. General information about a learning object, for example, its name, free text description, and keywords.
2. A life cycle category describing the change history of a learning object: **Who created the content of the object, and when?**
3. A Meta-metadata category describing the life cycle of the metadata itself.
4. A technical category supporting information about the data format and the physical location of the content; possible data formats are XML, PDF, PowerPoint, etc.
5. An educational category grouping the educational and pedagogical characteristics of a learning object, such as "typical learning time", "typical learner age", "level of difficulty", etc.
6. A copyrights category describing, in a very basic way, restrictions in terms of intellectual rights and/or costs for using the learning object.
7. A relations category providing information about relations to other learning objects, like a link between an article and a slide show or between the original and the translated versions of an object.
8. Annotations. Metadata to be filled in by the user, who gives feedback concerning the application and utility of the learning objects in his/her context.
9. A classification category to describe domain-specific classification schemes.

The CustomDP approach to providing reusable learning objects utilizes the results of the introduced specifications as far as possible, but some adaptation was necessary to meet the specific requirements of the project and the area of application, digital printing.

2.3 XML

XML has gained a lot of importance both as medium-neutral content format and as format for data exchange. SCORM and many metadata languages have been expressed in XML so that data can be exchanged automatically between applications. XML is mainly applied in the creation of content, but also metadata can be imported in XML.

3 From the Idea to the Technical Concept

3.1 Workflows

The most important workflows are creating content, importing content, validating content and metadata, creating aggregations, and viewing content.

According to a DTD that has been developed in the project, the preferred format for content creation is XML; a more detailed explanation of this will follow in chapter 3. The system itself does not directly support content creation. Two main types of content can be imported into the system: learning objects created according to our DTD together with files that are referenced to in the XML text and single resource files (so-called Sharable resources) that can be referenced to in future XML files. Furthermore, it is possible to create metadata for (**the referenced content of**) training material which cannot be imported into the system. In this way the system can thus also be used as a databank for all types of training material.

When an XML file is imported, the content is checked and metadata created by the author is extracted and stored in the database. Being its author, the content creator has special knowledge about the content and must therefore be given a chance to describe the particular resource.

However, the content cannot be solely the content creator's responsibility. Once the learning object is in the system, its content and metadata are therefore approved by a so-called validator. After the validator's approval of content and metadata, the new content becomes visible to other users. The metadata plays an important role in the system because content can be found by searching it.

Creating aggregations, i.e. the selection and organization of learning objects into a useful training material package, is a key task in the system. The system supports arranging the content into a hierarchical structure. The top level is a course, which consists of at least one module. A module contains at least one topic, which consists of the actual learning objects, or content. In creating an aggregation either existing ones are used as a starting point or an entirely new one is generated.

There are two main ways of viewing content: as HTML or as PDF. When a user wants to see an aggregation, the aggregation structure is displayed; the user may either navigate through the material in the default order or jump directly to the part that s/he wants to see first.

3.2 System Architecture

The CustomDP system architecture is illustrated in **Figure 1**. Components in grey boxes run on the server side. If needed, they can be distributed in a number of computers. The actual content (the learning objects) is stored in a file system. However, the major data storage is a relational database. Java servlets do not use the database directly. Rather, they use Database API (DB API) written in

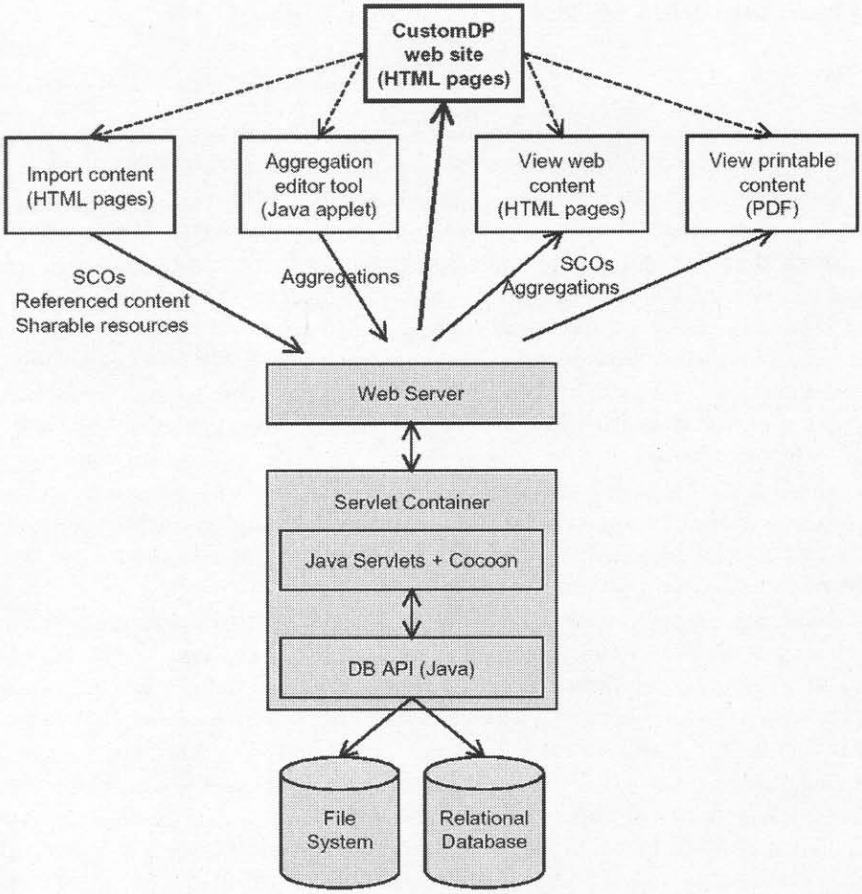

Fig. 1. *The CustomDP system architecture.*

Java to access the database. The DB API hides database details from the servlets and offers an easy-to-use API to servlet programmers. In addition to CustomDP servlets, the Cocoon web publication framework runs on the servlet container to generate printable PDF documents.

The CustomDP web site consists of a set of HTML pages. Most of them are generated dynamically by Java servlets. The core services of the web site are those related to learning objects, which are imported by using a set of HTML pages. The Aggregation Editor Tool is a Java applet for creating and modifying content aggregations. Learners can view aggregations by means of HTML pages. Moreover, aggregations can be downloaded as printable PDF documents.

4 Creating Content for the System

The most important delivery channel for the CustomDP content is the browser. Therefore, any content that can be viewed or used by means of a browser can be included in the system. However, in order to be able to also serve other channels, XML was chosen as main **authoring choice (?)**. From XML, content can be automatically converted into other formats. The XML file is used as the container for the content: it includes the text and links to other files in browser-supported formats. From the point of view of educational content, the XML file is a learning object, which teaches something and should not be directly dependent on other objects.

Above all, the DTD was required to be easy to learn, i.e. without intensive training, and to include the elements necessary to present the educational structure of the content. Some training material used by project partners was analyzed in order to identify what kind of structures and elements would be required. Furthermore, metadata was to be included in the XML file in order to enable the authors to create metadata as part of the content creation task.

Existing DTDs were analyzed to determine if they met the requirements for this projects. The main options were Learning Material Markup Language (LMML) [7] and XHMTL [8]. Regarding tasks and exercises, IMS QTI [6] and its lighter version IMS QTILite [5] were analyzed.

However, since none of the existing alternatives met the requirements, a new markup language had to be defined. Its main structure is close to XHTML, but the required metadata and educational elements, including elements needed to mark up the most important exercise types, were added. Even though, due to its relatedness to HTML, the basic structure of this DTD is easy to understand, the content creator is required to put some effort into learning the utilization of all its features. Customizations were made to a XML editor tool to offer a convenient authoring environment to content creators; the result resembles the familiar office authoring environment.

5 Presenting the Content to Users

Using a Document Type Definition based on experiences with existing DTDs assures the creation of high-quality learning contents in XML. The XML concept includes the separation of content and layout, i.e. authors do not explicitly give direct presentation-related information, which is an important prerequisite for the production of different renditions of the content. The consistency of the different renditions is assured only if there is just one source file for the content.

The renditions are created by using style sheets. Stylesheets can be created to produce different presentations of the same XML source and to generate different

output formats, e.g. HTML or PDF. Learning objects could therefore be presented differently depending on the learner and the content. The metadata includes information, such as the degree of difficulty, the learning context, the age of the user, or the degree of interactivity of the content, which could be employed in the selection of the stylesheet for creating the rendition as well as to make it more pleasant and effective to use the content.

In the current version of the system, the main rendition is HTML for browsers. The learner uses an intuitive classical interface, which designers developed particularly for the CustomDP system. The interface is divided into a navigation frame on the left and the actual content on the right. **The user may choose the user interface language from the project language (?)**, which facilitates the use of the system. The employment of intuitive graphic elements supports an easy operability. Furthermore, the use of typographic and layout elements, like colours, pictures, animations, etc. contributes to the learning success.

The HTML or PDF are generated from the original XML content by means of XSLT style sheets at run-time; updates to the content are thus immediately available to the user. Actuality of the system is crucial for its acceptance. The style sheets for the generation of content are exchangeable; the layouts of the renditions can therefore be easily changed.

Another advantage is the strict separation of content and layout, which, in addition to enabling the generation of currently known output formats, like HTML or PDF, allow the support of future channels and renditions. The contents created by using the CustomDP tools are therefore future-proof.

6 Outlook

A practical test of the system and the content created as small learning objects is currently being carried out. When analyzing the user feedback, the focus will be particularly on the following issues:

- the ease of creating aggregations including finding content
- the viewing of the content
- the quality of the content
- the usability and the reliability of the system as a whole.

Only after this practical test period, it can be determined if the metadata and the DTD for the actual content serve the needs of real users. The results will contribute to creating metadata vocabulary for educational content and to defining what a learning object should be.

The future development of the system should include better support for copyright management and different business models for using the content. The system itself could be either utilized to become an important content and metadata

databank relating to some specific areas, for example, printing technology, or be developed further to include content of many different fields. Each subject requires its own subject-specific metadata section. In future, as soon as the standards and recommendations which will gain wide-spread practical support are determined, additional support for content import and export can be included.

References

1. Advanced Distributed Learning. Sharable Content Object Reference Model, Version 1.1. Dodds, P. & al. (ed) 223 pp. 2001. (http://www.adlnet.org/ADLDOCS/Document/SCORM1.1.pdf)
2. CEN/ISSS WS/LT Learning Technologies Workshop. A Standardization Work Programme for "Learning and Training Technologies & Educational Multimedia Software"
3. IEEE Learning Technology Standards Committee (LTSC). Learning Object Metadata (LOM). Draft Document v2.4. 16 November 1998 (http://ltsc.ieee.org/doc/wg12/LOMdoc2_4.doc)
4. IMS Learning Resource Meta-data Specification. (http://www.imsproject.org/metadata/index.html)
5. Smythe Colin, Shepherd Eric: IMS Question & Test Interoperability QTILite Specification, Final Specification Version 1.1. IMS Global Learning Consortium, Inc. 9.4.2001, http://www.imsproject.org/question/qtilite03.HTML. (16.4.2002).
6. Smythe Colin, Shepherd Eric: IMS Question & Test Interoperability: ASI Best Practice & Implementation Guide, Final Specification Version 1.1. IMS Global Learning Consortium, Inc. 9.4.2001 / 24.10.2001, http://www.imsproject.org/question/qtibpig03.html.
7. Süß Christian, Winter Stefan: Learning Material Markup Language 1.1. Universität Passau 2000, http://daisy.fmi.unipassau.de/pakmas/lmml/11/doc/en/html/ (16.4.2002).
8. W3C HTML working group: XHTML^{TM} 1.0: The Extensible HyperText Markup Language, A Reformulation of HTML 4 in XML 1.0. Word Wide Web Consortium (W3C) 26.1.2000, http://www.w3.org/TR/xhtml1 (16.4.2002).