MODELLING, AUTHORING AND PUBLISHING THE "DOCUMENT ANALYSIS" LEARNING OBJECT

Alexandre Flament (alexandre.flament@imag.fr)
Emmanuelle Villiot-Leclercq (emmanuelle.villiot-leclercq@imag.fr)
Laboratoire CLIPS, Fédération IMAG –
BP 53, 38000 GRENOBLE Cedex 9 - FRANCE
http://www-clips.imag.fr/arcade/

KEYWORDS: Learning object, authoring, publishing, XML standards

Abstract

This article describes the modelling and implementation of a document analysis learning object. Actually, the objective of this research is double : providing models and tools to teachers allowing them to produce learning objects and building a publishing chain which can be applied to other kinds of learning objects. Implementation choices rely on interoperability and use of standard.

INTRODUCTION

The LOG, Lycée Ouvert de Grenoble, is an e-learning structure intended for high-school students. A team of LOG teachers wishes to produce a large variety of exercises on document analysis. They favour several criteria: interoperability, reusability, and separate expression of both the document content and the scenario. The project OASIS (Open Architecture and Schools in Society) helped the LOG to be integrated in a numerical support, the *Cartable Electronique de Savoie â* using the portfolio metaphor. It is a personal and collective space for teachers, students, and parents. However, for teachers, the problems of structuring and producing their exercises on document analysis remain.

Our work on document analysis modelling started at the end of 2002. Villiot-Leclercq (2003) describes our approach and Analydoc, the model of document analysis. In the present paper we focus more on implementation in the context of the Cartable Electronique de Savoie® which will be used by LOG as a platform. This work allows us to build a publishing chain which can be applied to other kind of learning objects and other e-learning contexts.

In the first part, we describe quickly the Analydoc model upon which the process of the publishing chain of learning object has been built. In the second part, we explain in detail the process of this publishing chain of learning object, and in the last part, the technological constraints of the implementation.

DESCRIPTION OF THE ANALYDOC MODEL

The first object of this research was to help designers to structure their contents and to reuse these contents in scenarios fitted to the needs of learners. In the first time, we worked with a team of teachers of LOG who want to structure and to design a specific pedagogic activity: an activity of analysis of documents. These teachers wanted to propose global analysis activities and to be able to organise the tasks not as linearly as in a textbook. Furthermore they wished to be able to use an interactive designing process compatible with their pedagogical objectives, and

to adapt easily their previous productions to the needs of different students. In interaction with teachers, we modeled this activity and created the model AnalyDoc.
This question of structuration contents is linked to the definition of numerical contents standards. About the definition of these standards, the challenge is to allow the reuse of contents and their sharing in differents contexts and with different actors. Crozat (2003) describes four axis of formalisation: packaging of contents, indexation of content, structuration of contents and tracing of activities linked with contents.
Our choice is focused on the structuration of contents and on the process of the edition of learning objects. A lot of approaches were possible to build a model which structures an activity of analysis of document and the resources linked.
The learning object approach is a solution advocated by the ARCADE team of the CLIPS-IMAG laboratory (Grenoble). A learning object is a numerical entity allowing the learner to undertake an autonomous activity of training, independent of its context of use (David, 2003). According to the MARS model (Pernin, 1995), a learning object is composed of four elements which are complementary: the model (M) describes the system used by the learning object, the representation (R) is the student interface, the scenario (S) is the dynamic part of the learning object, and the associations (A) define links between the model, the representation and the scenario.
Another approach is EML, Educational Modelling Language (Kopper, 2001), which is more focused on activity. “An EML is a semantic information model and binding, describing the content and process within a ‘unit of learning’ from a pedagogical perspective in order to support reuse and interoperability” (Rawlins, 2002). Nevertheless, specifications of EML are very large and it is difficult for teachers to structure their activity due to this formalism.
As we wanted to work in a micro level and focus both on pedagogical contents and activities, we have combined these two approaches to build our model of analysis of documents. Document analysis involves a variety of tasks: reading, observation, interpretation, research, writing, exploration etc. In fact, the model is composed of the content (the resources like texts, pictures, questions etc.) and of the scenario which allows teachers to structure these resources in tasks according to their pedagogical objectives. This structuration is based on EML.

The content (model) component

The content is composed of two parts: general information and resources. General information globally describe resources: title, thematic of study, creation date, language, keywords… The second part is a list of typed resources. The resource types come from a review of LOG practice with Learning Space. Some invariants have been be detected. To prepare a document analysis activity, teachers need to have a first resource type: the document itself (text, image and, in the future, video and sound). The second type of resource is questions: closed question, open question, multiple choice questions. The third type of resource is aimed to help the student achieve the proposed tasks: complementary resource, guide, and « Did you know it? ». Further resource types are glossary and bibliography.

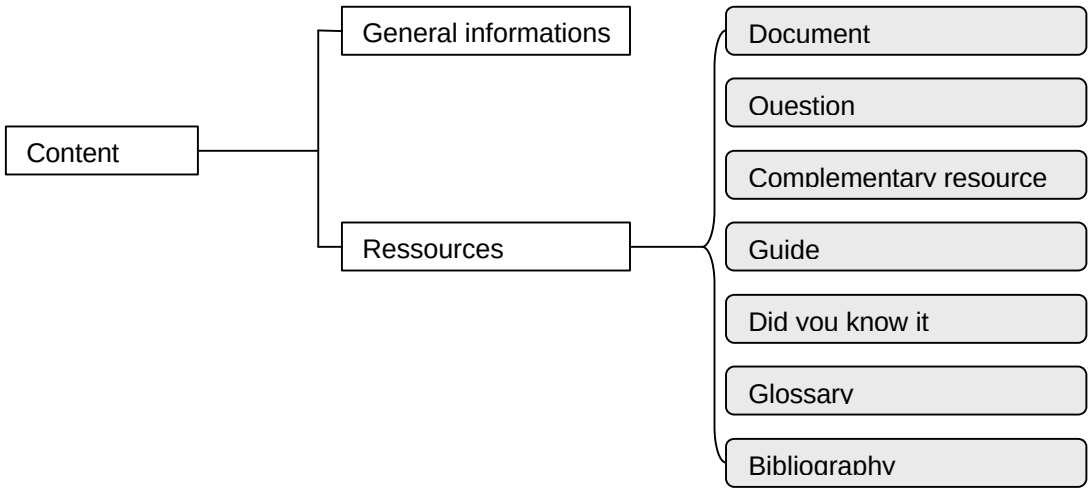

Figure 1 : Content structure

The scenario component

The Scenario is the dynamic part of a learning object. It describes how the activity will be proposed to the student. A scenario contains two kinds of elements: tasks and scenarized resources (see figure 2). A scenario lists different tasks that are organised as a sequence or as a selection. Each task contains instructions and scenarized resources. Scenarized resources are composed of a reference to a content resource, and of information about this resource, such as duration of accessibility.

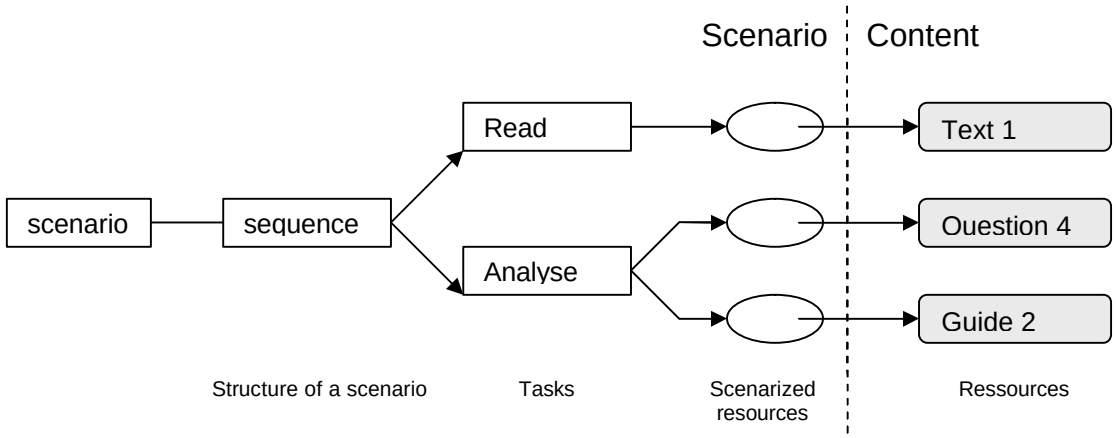

Figure 2 : Example of a scenario

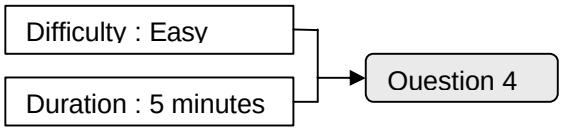

Figure 3 : Details of a scenarized resource

```
…
<RESSOURCES>
    …
    <RESSOURCE ID="Question4" TYPE="OPEN_QUESTION" >
        <REFERENCE HREF= "content.xml#Question4" />
        <AVAILABLE TIME=”300” />
        <DIFFICULTY VALUE=”EASY” />
    </RESSOURCE>
</RESSOURCES>
<STRUCTURE>
    <SEQUENCE>
        …
        <TASK NAME=”ANALYSE”>
            <USE RESSOURCE=”Question4” />
            <USE RESSOURCE=”Guide2” />
            <CORRECTION BY="STUDENT"/>
        </TASK>
    </SEQUENCE>
</STRUCTURE>
…
```

Figure 4 : XML sample of scenario

Actually, the first objective was to provide models and tools for teachers allowing them to produce the learning object “document analysis”. But, at the moment, from a specific model as Analydoc and specific learning objects, the objective of this work is to build a publishing chain which can be applied to other kinds of learning objects.

FROM TEACHER TO STUDENT: EDITION PHASE AND PUBLICATION PHASE

Describing Analydoc as semantic model changes the method of production for the teacher: he does not edit his document with a WYSIWYG application (What You See Is What You Get) like in Word but with a WYSIWYM application (What You See Is What You Meant). This concept was introduced by Power (1998). The creation of the learning object is made in two phases: first the teacher edits the content followed by the scenario. Then in the second phase, the teacher publishes his work. This method can be compared to the LaTeX environment: the author edits LaTeX document with LyX, for example, then he produces a PDF document.

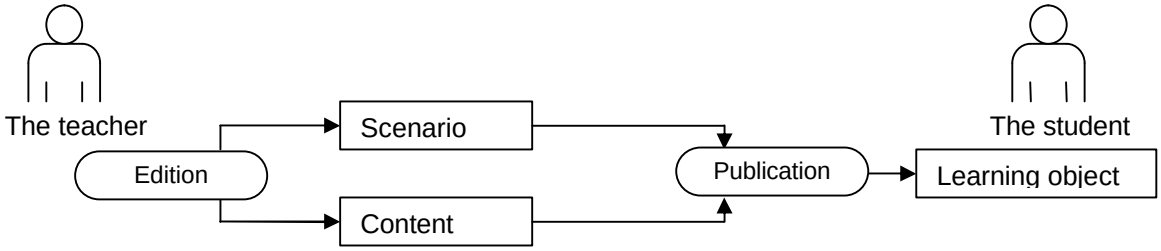

Figure 5 : From teacher to student in two steps

This approach depends on the distinction between content and scenario (actually, they are two separate files). This distinction allows teachers to produce different scenarios from the same set of resources. This approach suppresses much of these technical parts for the teacher: rewriting resources or copying in new files, defining student interfaces with HTML pages,….

Interoperability

Interoperability is an important criterion for learning objects: the learning objects should exist independently of platform. That’s why we decided to use XML technologies: XML to store objects, XML Schema to describe the grammar of content and scenario,and XSL to describe how to transform a teacher object into a student object.

DESIGN OF EDITION

The editing process described here is not specific to a particular model or schema.

The Edition document

The goal of *Cartable Electronique de Savoie â* is to give access to data and tools without worrying of the machine type. It means that all authoring work must be done on the web without installation of applications on client side, except common plug-ins.
The W3C (World Wide Web Consortium) released a candidate recommendation in August 2003 about a new kind of web form. XForms can describe forms for the web, but contrary to traditional HTML, content, model and constraints are clearly described. XForms does not describe how to display the form. XForms is rather intended to be integrated into other mark-up languages, such as XHTML or SVG. XForms will be part of the XHTML2 recommendation. By using XForms, we can produce or edit documents satisfying different grammars like DTD or XML Schema.
What we call an edition document is a document which contains XForms element inside an HTML document. An edition document allows a full or partial edition of the schema. Two schemas exist for Analydoc: the content and the scenario. Compared to an applet, the development and maintenance of an edition document is much easier. Moreover producing a document and not an application guarantees better continuity.

Integration of the edition document in the platform

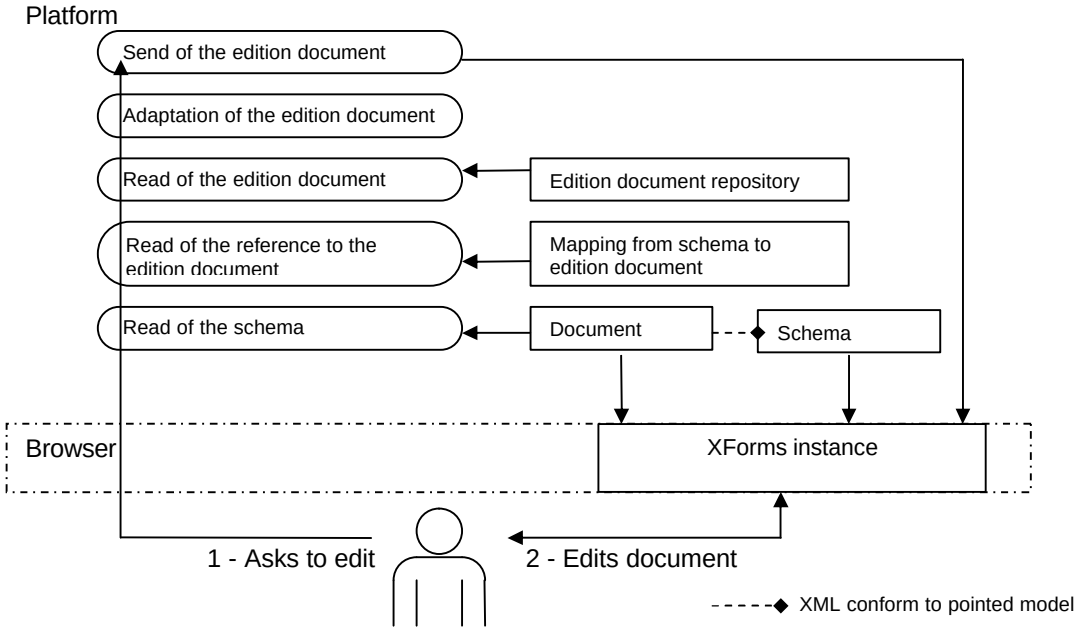

Figure 6 : Authoring tool in a platform

The edition document allows the user to edit one specified document. There are two implications: first, the platform must find a way to adapt edition document for the document the user wants to edit. Second, the platform must find which edit document to use for a document; this can be done using a reference to XML Schema in a document.
The figure 6 shows the relation between a user who wants to edit a document, browser, and platform. When the user wants to edit a document, five steps are done by the platform:

- Read reference to the schema, commonly named public identifier, is read from the document.

- Use mapping to get a reference to the edition document using reference to schema
- Get the edition document from a repository
- Adapt the edition document for the document to edit: only the URL of the document to edit should be added.
- Sent to the user the edition document.

DESIGN OF PUBLICATION

To publish a learning object, we start from its scenario. With this scenario and its content, we produce the student learning object for a student in XHTML format.

The publication starts with computation of scenarized content. This last one is the scenario in which all references to content are resolved. From the resulting document (scenarized content in figure 7), the XHTML version may now be produced. Another output format could also be generated, for instance PDF using XSLFO.

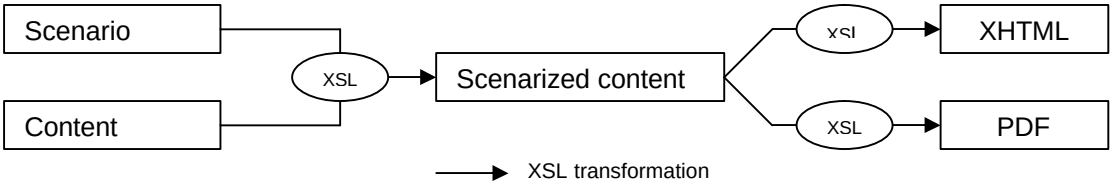

Figure 7 : Publication: From scenario and content to student learning object

CONCLUSION

By modelling content and scenario separately, we facilitate the reuse of resources by the teacher. Moreover, automatic transformation to create student learning objects helps the teacher to focus on his/her expertise field and to be free of technical constraints.
The use of standards (W3C recommendations) produces learning objects that respect several criteria: interoperability, reusability, adaptability. Moreover these standards make learning objects independent of the tools used.
From a specific need, we generate a production chain which can be used with many pedagogic objects. The edition document allows for simple design and maintenance. One perspective is the automatic production of the edition document from a schema and a graphic charter.

REFERENCES

CROZAT S., MAJADA M., SPINELLI S. (2003). SCENARIsup : Un projet pour la gestion des contenus pédagogiques numériques dans l'enseignement supérieur. CNUR'2003, Montpellier, France, Octobre 2003.
DAVID J-P. (2003). Modélisation et production d'objets pédagogiques : ressources numériques, xml et éducation : Revue Sciences et Techniques éducatives, volume 9, non-standard edition Hermès.
POWER R, SCOTT D, EVANS R (1998). What you see is what you meant: direct knowledge editings with natural language feedback: 13th European Conference on Artificial Intelligence".
RAWLINGS A, VAN ROSMALEN P, KOPPER R, MODRIGUEZ-ARTACHO M, LEFRERE P (2002). Survey of Educational Modelling Languages (EMLs), CEN/ISSS WS/LT Learning Technologies Workshop
VILLIOT-LECLERCQ E, DAVID J-P, FLAMENT A, GUILLOUX C (2003). Une démarche de modélisation d'un Objet Pédagogique « analyse de document» : Hypertextes hypermédias, H2PTM'03, Hermès-science edition.