Multimed Tools Appl (2010) 47:257–278
DOI 10.1007/s11042-009-0322-8

Focusing on content reusability and interoperability in a personalized hypermedia assessment tool

Fotis Lazarinis · Steve Green · Elaine Pearson

Published online: 10 July 2009

Abstract This paper presents the development of a modularized hypermedia testing tool, called iAdaptTest, based entirely on e-learning specifications and discusses how this architecture improves the reusability and the interoperability of the learning data. All the categories of data,—that is—topics, user profiles, testing data, adaptive rules, and testing results are coded in XML format complying with Topic Maps, IMS LIP and IMS QTI. The data are stored in distinct files and can be independently shared across different educational applications. The paper concludes with an evaluation study concerning the creation of formative and summative assessments for adult seminars. Through focused interviews, the participants of the study identified the ability to share information and the multi-criteria adaptation options as the most important features of the system. Further, in the second phase of the evaluation the files produced were shared with other educational applications and thus it was verified that the learning data could be imported and rendered correctly.

Keywords Adaptive hypermedia · Interoperability · Reusability · Assessment · Educational technology · Personalisation · Multimedia applications · XML

1 Introduction

Computer-based testing tools are becoming increasingly important for both learning and assessment [6, 30]. Testing tools are utilized in secondary education, in vocational training and in higher education in an attempt to minimize the test administration efforts and costs, maximizing at the same time the potential for more immediate, accurate and multimedia

F. Lazarinis (✉) · S. Green · E. Pearson
School of Computing, University of Teesside, Middlesbrough, Tees Valley TS1 3BA, UK
e-mail: f.lazarinis@scm.tees.ac.uk

S. Green
e-mail: s.j.green@tees.ac.uk

E. Pearson
e-mail: e.pearson@tees.ac.uk

enriched assessments. Technology for assessment spans from scanning technology to support marking to computerized adaptive testing systems. With the increased utilization of technology for both learning and assessment, the need to efficiently re-use and share the data accumulated in educational tools has been more demanding than ever. Content reusability is a key aim for educational technologists, especially in testing systems where several categories of data are involved.

CETIS[1] (Centre for Educational Technology Interoperability Standards) states that standard ways of describing educational material are needed so that they can be easily searched for and located. E-learning specifications refer to the standardization of XML structures which are used to describe various aspects of the learning procedure. IEEE[2] (Institute of Electrical and Electronics Engineers) and IMS Global Learning Consortium[3] (Instructional Management Systems project) are examples of organizations and consortiums which developed XML learning standards for describing, among others, e-lessons, user profiles, e-portfolios and testing data.

The development and diffusion of e-learning standards raised new research questions related to how we can efficiently utilize them in order to increase the reusability and syntactic interoperability of learning content. We believe that in order to greatly improve these "abilities" we need educational tools that are highly modularized and truly conform to e-learning standards. The output of each module should comply with a distinct e-learning standard rather than a mixture of XML elements from different standards as suggested in some studies [9]. In this way the output of each module could be reused and shared separately with other educational applications.

In this paper we present the development of iAdaptTest, a modular adaptive testing tool conforming to the Topic Maps [38], the IMS Learner Information Package [20], and the IMS Question and Test Interoperability [21] standards. The output of each module conforms fully to the latest version of the specification and is stored in a different XML file. As the e-learning industry continues to expand every day, and the methods and tools necessary to create and maintain content and infrastructure applications become more complicated, there is an inherent need for multimedia educational applications to interoperate and exchange data in order to better support the needs of learners and educators. The iAdaptTest is a modular system which can be used either as a standalone desktop application or can be published on the Web. Its modular nature supports both the independent use of the authoring tools to create interoperable content or the combination of the tools in a sequential mode in order to create complex standards-based adaptive assessments. Further, the tool is domain independent and can be used for various educational aims. The main aim of the paper is to present the architecture of the testing tool and discuss how this architectural design focuses on the reusability and interoperability of learning content. On the basis of an evaluative study we discuss the added value of the tool in terms of its adaptability and content reusability. The evaluation presented below was based on the use of the tool as a pilot study with a small number of educators and considers the perceived ease of use of the tool together with the assessment strategies and rules employed. In conclusion we consider how iAdaptTest system might be improved in the future to help maximize interoperability and improve pedagogic value to educators.

The remaining of the paper is organized as follows. In section 2, a number of studies related to the utilization of e-learning standards in assessment tools are discussed. Section 3

[1] http://www.cetis.ac.uk
[2] http://www.ieee.org
[3] http://www.imsglobal.org

presents the basic modules of iAdaptTest and discusses the factors that influenced their design. Section 4 presents the results of an evaluative study and section 5 summarizes and gives directions for further research.

2 Background

Content reusability is a novel aim in educational technologies. One of the most well known learning standards for coding learning data is Sharable Content Object Reference Model [37]. A number of tools abiding to the specification or presenting architectural models have been proposed for SCORM [24, 31]. These papers discuss the need to create interoperable systems and techniques for exploitation of the learning standards; additionally they critically review some of the problems of utilizing this learning standard.

The Question and Test Interoperability (QTI) specification describes an XML based technical format for the coding and exchange of assessment content from individual questions through to complete tests. QTI, which is often referred to as IMS QTI, structures material into assessments, sections, and items. An item is the formal name for a question and assessment is the terminology used for a test. The latest version of IMS QTI supports simple and composite questions and question repositories. QTI was first published in 1999 and has already been matured to a new version 2.1.

A number of QTI aware tools have already been created. In [2] a QTI assessment tool used for assessing physics is presented and utilized in the construction of real assessments. Another study discusses the integration of a QTI v1.x editor and player in a Web portal [14]. These tools support a limited number of questions of the QTI v1.x standard. Rules for processing the results are also defined using the QTI's XML elements. The QTI assessments are integrated into the Web portal using the IMS Learning Design [18]. In [29] an extensible QTI engine is presented. Another tool, QED (Quiz Engine Developer), is an open source and platform independent quiz developer [22]. This is a standards compliant system which permits the user to develop questions at the lowest level of granularity (termed Quiz Objects). An architectural design for resource effective processing of QTI v2.1 on resource-limited mobile devices is presented in [45].

IEEE Public and Private Information [17] and IMS LIP are XML based learning standards for learner profiles. IMS LIP is more general since it contains the attributes of IEEE PAPI. Tools utilizing these standards have already been published [5, 32]. It has also been proposed to combine elements from the two standards to provide an expanded user profile [9].

The need to re-engineer existing tools to conform to e-learning standards has also been identified in the literature [40]. In one of our older works a hypermedia environment for managing online courseware based on adaptable templates is utilized in courseware construction in a number of adult training seminars [26]. During the evaluation of the system, the participants identified as important the need to import their courses to other e-learning tools and thus it was concluded that the tool needs to be re-engineered to become compatible with a learning standard.

In [34] a layered approach for a SCORM compliant assessment tool is presented. The need to extend SCORM for assessment authoring is also discussed in [4]. In [1] the IEEE LOM (Learning Object Metadata) standard is used to describe learning objects. The authors present a framework utilizing additional standards to deliver personalized lessons.

Van Rosmalen et al. [43] introduce a standards-based model for adaptive e-learning and investigate the conditions and tools required by authors to implement this model. The presented model combines IMS LD and IMS QTI in order to provide personalized content.

However, the authors do not present or evaluate an actual implementation and the mixture of elements of the two e-learning standards makes their independent utilization more difficult. The combination of re-usable, though not standards based learning objects was also proposed in [25]. In [12] the creation of adaptive learning systems is based on the combination of component technology, semantic metadata, and adaptation rules. However, the authors do not discuss the issues of re-usability and interoperability or how they code their data.

Paramythis and Loidl-Reisinger [33] examine the sufficiency of existing e-learning standards for facilitating and supporting the introduction of adaptive techniques in computer-based learning systems. Their analysis has pointed out that existing standards do have some provisions for adaptation, but require substantial extensions to accommodate common practice in adaptive learning environments. In [10] an adaptive learning environment, called Personal Reader, is presented where the components realizing the adaptation services require meta-information about courses, learning resources, and about learners. The metadata are coded in Dublin Core [11] and IEEE LOM [16].

In [41] it is argued that standards like IMS LD cannot cover alone the requirements of modelling adaptive learning applications and should be used in combination with other software components. A system for constructing education and entertainment experiences by relating TV and learning content is presented in [35]. The authors construct personalized activities by proposing the adaptation of learning contents by defining an extension to the SCORM specification called A-SCORM (Adaptive-SCORM).

In [28] adaptive testing functionalities are discussed by proposing new extensions to IMS QTI for item parameters and interfacing parameters' characterisation. Boticario and Santos [3] discuss the use of IMS standards for developing a general framework and a running architecture to support the full life cycle of adaptive learning systems. An adaptive authoring system is extended in [13], by adding compatibility with IMS QTI and other IMS standards. The authors claim that in this way the data of their adaptive tool can be shared across different tools.

The above selected publications show that there is an established base of researchers working in the area of e-learning standards. Some work has also been done on integrating learning standards into adaptive tools. Most of the previous research studies focus mainly on a specific learning standard and try to create models or tools which abide by this specification. However, it can be concluded from these studies that working with only one standard limits the required options for modelling the various data involved in building adaptive learning applications. Therefore developers resort into less generalized solutions or extensions to the standards which in turn undermine data re-usability.

As is discussed in the next sections, our tool is a complex adaptive testing system with various categories of data. In order to improve the ability to share content with other educational applications we believe that the data of the system should completely abide by one or more learning specifications. Further, data abiding to different standards should be stored separately so that they can be utilized individually. The architecture of the system should be extensible to easily absorb future expansions which will be driven by alterations of the e-learning specifications. By complying with these demands we can increase the interoperability, shareability, and durability of the system.

3 iAdaptTest: an interoperable adaptive testing tool

As discussed in the previous sections, the main concerns of the present work is to increase content reusability and interoperability among educational applications. To achieve these

novel goals the system is developed as a collection of separate modules (see Fig. 1). Each module of the authoring subsystem conforms to a specific XML based standard. The test execution and data presentation subsystems take as input all these data categories but they update mainly the learner profile of educators.

Users (learners and educators) access the tools of the authoring environment through a unified interface which allows the easy and uniform utilization of the modules of the system. The tools are integrated as different tabs in a single window, sequenced according to their role in the editing of the adaptive tests. For example, the topic management tool is the first tab, followed by the learner profile management module and so on. Through this interface users can start, for example, the question editor assessment, the topic management tool, etc. That way the users do not need to be aware of the system's architecture or have any knowledge of the technical details of the testing tool.

Each tool describes semantically its data using a distinct e-learning standard. The output of each tool is stored in a different repository. This way the output of each tool can be shared autonomously with other educational applications. For example, the topic map could be reused in a learning management system to describe the topic hierarchy of the online lessons. The questions and assessments can be shared with other assessment tools and the user knowledge accumulated in the system could be shared with other learning applications understanding the used learner profiling e-learning standard. This distributed architecture improves content reusability as it is possible to share different parts of our data with various e-learning applications. All the tools presented in the next sections are developed in Visual Studio.Net. So they can be used either as standalone Windows based applications or they can alternatively, through the options of the development environment, be published as a Web application. The XML data are stored as text files. Each assessment is stored in separate directories. As explained, although there are four discrete tools in the system, these are initiated through a simple interface and they communicate through their data, i.e. the output data of one tool are the input data of the next tools. Of course this is transparent to end users who see only the final XML files stored in a single directory.

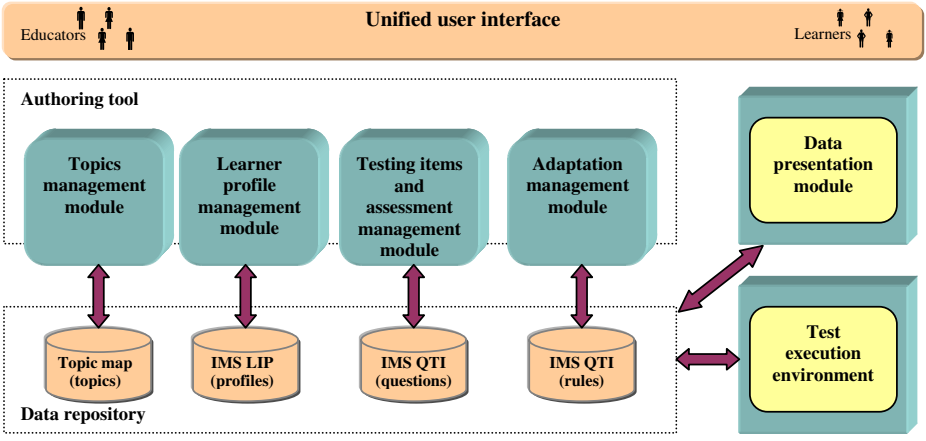

Fig. 1 High level architectural view of the iAdaptTest system

3.1 Authoring system

3.1.1 Topic management

Figure 2 shows the interface of iTM, a visual tool which allows educators to develop and graphically display the topic hierarchy. Educators can then visually associate their testing materials with the appropriate topic and subtopics. The topics' data need to be represented in an XML based standard to promote the goals of interoperability and content shareability.

A suitable specification for coding the topics' data is topic maps. Topic maps are an ISO standard for the representation and interchange of knowledge, with an emphasis on how easy it is to find the right information [39]. A topic map can represent information using topics, associations, and occurrences. They are thus similar to semantic networks and both concept and mind maps in many respects. Topic maps have a standard XML based interchange syntax called XML Topic Maps (XTM) and query and schema languages are being developed within ISO. Topic maps for e-learning are proposed in other studies as well [8]. In our tool the topics and subtopics network is important for estimating the knowledge of students to specific topics through their performance. This knowledge is overlaid in their profile and propagates from the sub-topics up to the main topic.

The graphical tool presented in Fig. 2 supports the creation, update and import of topics and presents the dependencies in a graphical mode. The XML output of the tool (see Fig. 3) is validated against the DTD (Document Type Definition) of Topic Maps [38]. The purpose of a DTD is to define the legal building blocks of an XML document. It defines the document structure with a list of legal elements and attributes. So the produced XML files should be validated using the DTD schema to ensure compatibility with the specification. As shown in Fig. 4, the output of each authoring module is validated against the DTD of a specific XML specification. In this manner, we ensure that the content can be shared across educational applications that share an understanding on the specific content. XML schemas are also used to describe the structure of XML documents and since they are more expressive than the DTDs, they could be used as an alternative for validating the well-formedness of the produced XML files.

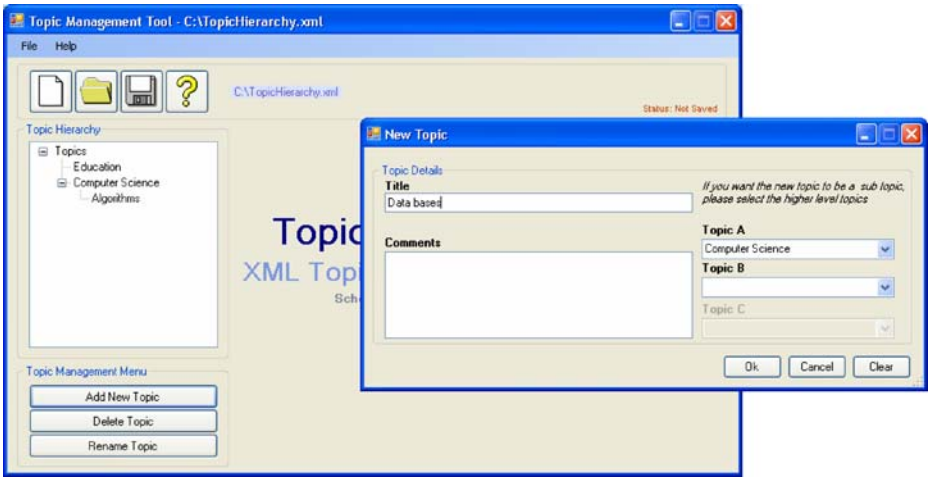

Fig. 2 The iTM topic map management tool

```
…
<topic id="TP_2">
        <baseName>
                <baseNameString>Computer Science</baseNameString>
        </baseName>
</topic>
<topic id="TP_3">
        <instanceOf>
                <topicRef xlink:href="#TP_2"/>
        </instanceOf>
        <baseName>
                <baseNameString>Algorithms</baseNameString>
        </baseName>
</topic>
…
```

Fig. 3 Excerpt from a topic map XML file

The interface of iTM (Fig. 2) and of the next tools follows the conventions of the Microsoft office tools so as to minimize the learning effort for users. Users may create, open, save or save as the topics and ask for help. These functions are available through the menu options or via the available icons. Each function is independent of the others and thus new functions, e.g. printing, can be easily added. From an engineering point of view, the testing system offers various levels of modularity, i.e. modularity at the level of tools and modularity at the level of functions.

3.1.2 Learner profile management

Learner profiles consist of a set of attributes which describe the personal data of learners (e.g. name, date of birth), their formal education (e.g. degrees), their working experiences, their goals (e.g. to prepare for a final exam), their preferences (e.g. to be presented with the questions of increased difficulty) and their evaluation data (e.g. activities performed during the test).

Learner profiles are managed through the learner profile management tool, iLM (see Fig. 5). The learner profiles management tool supports initialization of profiles, import from other sources, update and display of profiles in alternative modes. Initialization is performed either per user or in batches, based on specific learner pre-defined stereotypes as in older tools [23]. At the moment these stereotypes are beginner, intermediate and advanced and concern primarily the knowledge level of learners to specific topics. Through

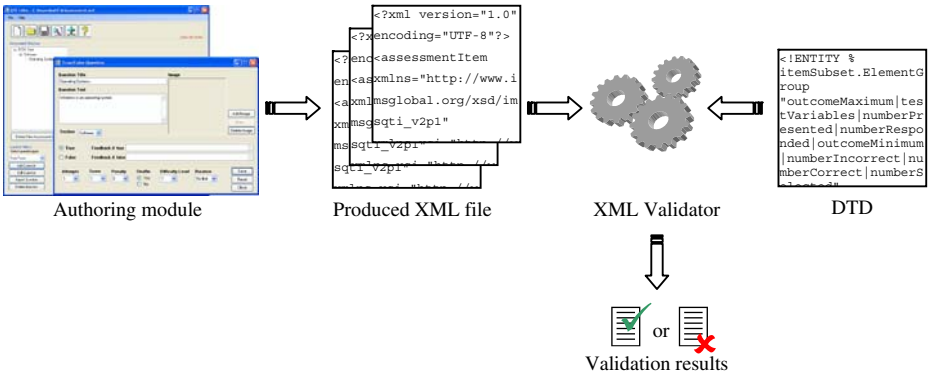

Fig. 4 Validation of the produced XML files

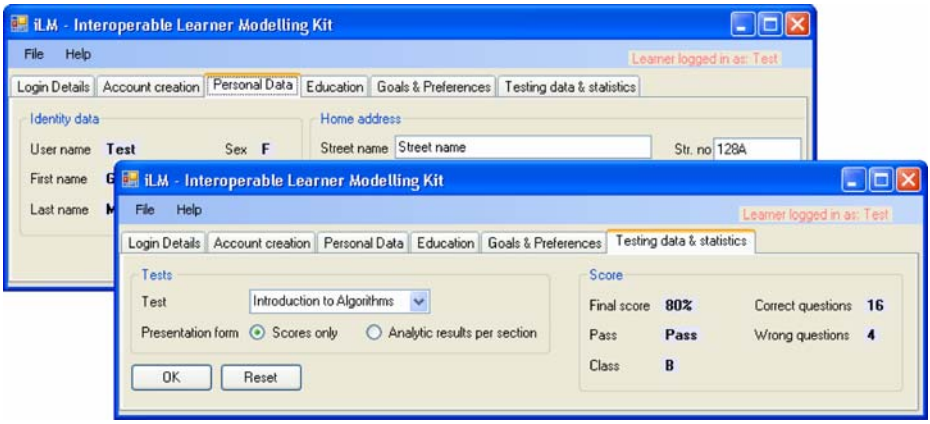

Fig. 5 The iLM, learner profile management tool

the learner management profile module both learners and educators can log in and based on their assigned rights, edit or view their profiles. Educators can additionally create or update new accounts.

In adaptive education hypermedia systems, topics are overlaid in the learner profile [7]. That is, the knowledge of the learner is associated and measured with respect to specific topics of the topic hierarchy, usually called domain model. So a learner profile should contain in addition the knowledge of each learner related to the topics of the topic network created with iTM tool. So iLM needs to be able to read the topic map XML file produced from iTM tool and to overlay this information in the user profiles.

Suitable specifications for learner profiles are IEEE PAPI (Public and Private Information - http://edutool.com/papi) and IMS LIP (Learner Information Profile). We selected IMS LIP which utilizes and expands most of the features of IEEE PAPI. IMS LIP contains a large number of attributes capable of coding all the information for learners. The high level XML attributes used to code the desired information are: <learnerinformation>, <qcl>, <competency>, <goal>, <accessibility>, <transcript> and <activity> (see Fig. 6). Using these data we code the knowledge, qualifications, achievements, goals and activities of the learners.

3.1.3 Testing item and assessment management

Testing items and assessments are coded in IMS QTI standard (Question and Test Interoperability). This specification structures material into assessments, sections, and items. An item is the formal name for a question and assessment is the terminology used for a test. The latest version of IMS QTI supports simple and composite questions and question repositories.

Figure 7 shows iQTIed which allows the creation of various question types (e.g. true/false, multiple choice, gap match, association, etc) and assessments and Fig. 8 shows an example of its output. Using the visual tool, educators are able to create questions and assessments and to associate them with specific topics and also view their assessment in a hierarchical mode. The editor allows creation, import and modification of QTI items. The output of the editor conforms to the latest QTI version 2.1. Each question item is stored in a different XML file and is referenced in the assessment's file as the specification requires.

```
…
<partname>
        <typename>
                <tysource sourcetype="imsdefault"/>
                <tyvalue>First</tyvalue>
        </typename>
        <text>Gloria</text>
</partname>
<partname>
        <typename>
                <tysource sourcetype="imsdefault"/>
                <tyvalue>Last</tyvalue>
        </typename>
        <text>MacTest</text>
</partname>
…
```

Fig. 6 Learner profile coded in IMS LIP

3.1.4 Adaptation management

Adaptive testing systems are computerized tests aiming to accurately reflect the instructional level of each student with a shorter number of queries tailored to the characteristics of the individual learner [42, 44]. The existing computerized adaptive testing (CAT) systems base their adaptation mainly on the learner's performance using various

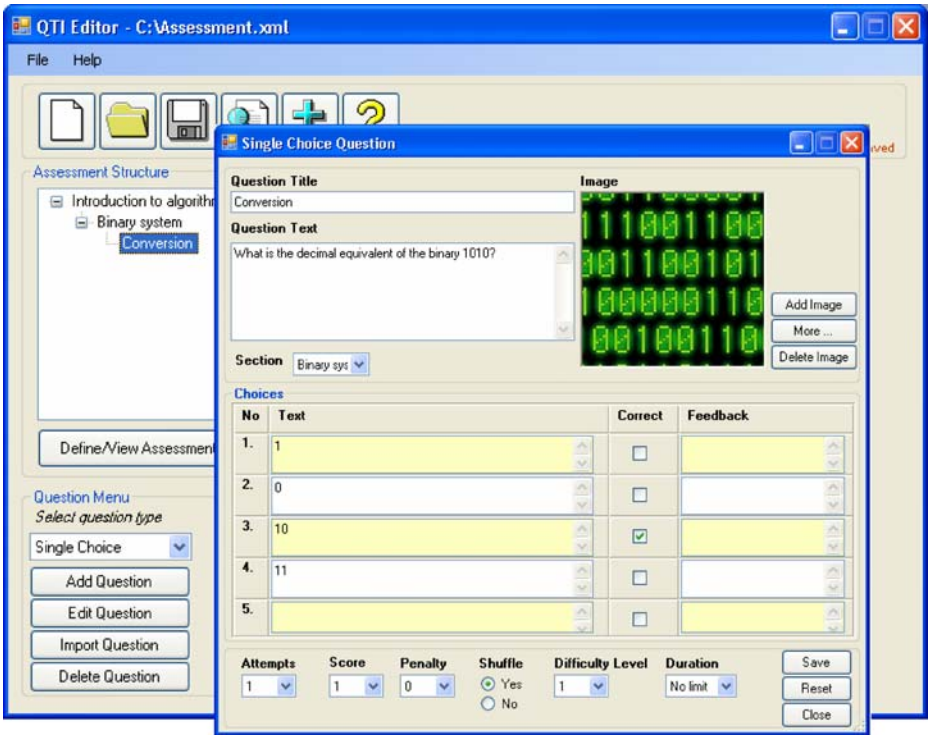

Fig. 7 The iQTIed Testing item management tool

```
…
<assessmentItem … identifier="choice" title="Conversion" adaptive="false" timeDependent="false">
    <responseDeclaration identifier="RESPONSE" cardinality="single" baseType="identifier">
        <correctResponse>
            <value>3</value>
        </correctResponse>
        <mapping lowerBound="0" upperBound="1" defaultValue="0">
            <mapEntry mapKey="3" mappedValue="1"/>
        </mapping>
    </responseDeclaration>
    <outcomeDeclaration identifier="SCORE" cardinality="single" baseType="float"/>
    <itemBody>
        <p>
            <img src="images/binary.jpg" alt=""/>
        </p>
        <choiceInteraction responseIdentifier="Item_Item1" shuffle="true" maxChoices="1">
            <prompt>What is the decimal equivalent of the binary 1010?</prompt>
            <simpleChoice identifier="1" fixed="false">1</simpleChoice>
            <simpleChoice identifier="2" fixed="false">0</simpleChoice>
            <simpleChoice identifier="3" fixed="false">10</simpleChoice>
            <simpleChoice identifier="4" fixed="false">11</simpleChoice>
        </choiceInteraction>
    </itemBody>
    <responseProcessing template="map_response"/>
</assessmentItem>
    …
```

Fig. 8 Excerpt from an IMS QTI assessment item file

statistical models. In our tool, adaptation of the testing procedure relies on the performance, the prior knowledge, the preferences and the goals of the system. Educators are able to force advancement of the procedure where a learner has passed a threshold or when a specific educational aim (e.g. answer all the high difficulty questions correctly) is met.

The adaptation management tool, called iAdaptEd (see Fig. 9), allows educators to visually customize predefined rules and apply them to certain questions. The latest version of QTI supports adaptive items and rules and hence it is used to store the data for the adaptation as well. The <responseProcessing> QTI element and its sub-elements such as <responseCondition>, <responseIf>, <and>, <not>, are used, among other QTI XML elements, to code our rules (Fig. 10). These rules are stored into different QTI XML files than the QTI questions in order to isolate the basic learning content from the adaptive procedure. IMS Simple Sequencing [19] could also be used in a future expansion of the adaptive rule subsystem. The integration of the Simple Sequencing specification will be straightforward and transparent to the rest of the system due to the modularized nature of the tool.

Each adaptation rule has an initiation point, a condition to be met and a specified action. Initiation points are set on specific questions of the assessment and testing routes change based on whether the conditions are met and actions happen.

The conditions concern:

- the performance of the learner
- the prior knowledge and formal education of the learner
- the current goals or the preferences of the learner

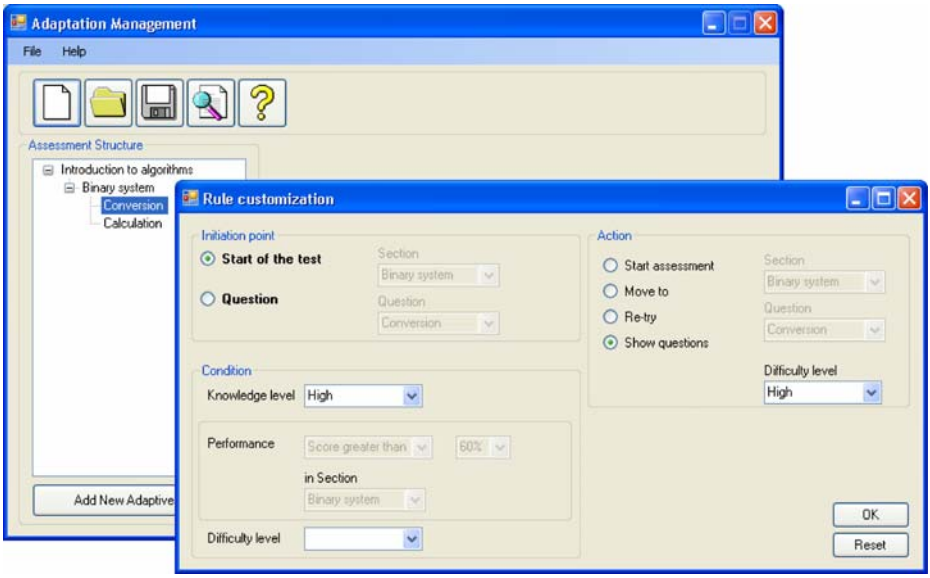

Fig. 9 The iAdaptEd adaptation management tool

The supported set of actions is:

- start assessment on question
- start assessment on section
- move to question
- move to section
- re-try specific questions
- re-try assessment
- show questions of specific difficulty level

```
…
<responseCondition>
    <responseIf>
        <match>
            <variable identifier="RESPONSE"/>
            <correct identifier="RESPONSE"/>
        </match>
        <setOutcomeValue identifier="SCORE">
            <baseValue baseType="integer">1</baseValue>
        </setOutcomeValue>
    </responseIf>
    <responseElse>
        <setOutcomeValue identifier="SCORE">
            <baseValue baseType="integer">0</baseValue>
        </setOutcomeValue>
    </responseElse>
</responseCondition>
…
```

Fig. 10 Excerpt from an IMS QTI response processing file

The described system is designed both for summative and formative evaluations [15, 36]. Formative assessment is often performed at the beginning or during a program, thus providing the opportunity for immediate evidence for student learning in a particular course or at a particular point in a program. Summative assessment is comprehensive in nature, provides accountability and is used to check the level of learning at the end of the program.

In our system, summative assessments can rely on the performance of the learner, so customization could be left to the rules set by the educators. Formative assessments aim at helping learners to appreciate their real knowledge. So the adaptation could be based on the goals of the learners. If learners feel that they are well prepared they could ask for difficult questions. Or if they do not understand some questions, they could try easier questions. In any case our tool is flexible enough to support various testing strategies and accommodate the goals of both learners and educators.

3.2 Test execution and data presentation system

The test execution component takes as input all the data categories and runs the test. Users log into the system and select one of their assigned tests (see Figure 11). Based on the adaptation rules users may be asked to either define the difficulty of the questions they wish to try, or the system may start the adaptation at a point which matches their knowledge level. The test execution system updates the learner profile data. Mainly it updates the knowledge level of a learner at specific topics and subtopics and records the activities of the learner so that both the learners and educators can have a clear image of their testing attempts.

At the end of the assessment, users may log into the Learner Management system and view statistics and some conclusions on their testing procedure (see Figure 5). At the moment the statistics are shown in textual form and concern primarily the number of correctly answered questions and the classification of the learner, the number of wrong answers, the number of questions per test, etc. Based on the abundance of testing data, it is relatively easy to produce more statistics for each learner in both graphical and textual forms.

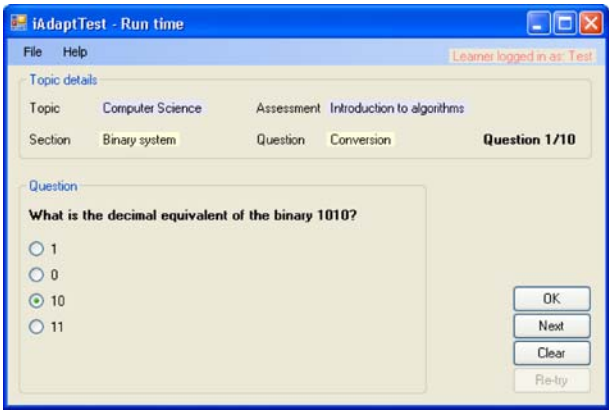

Fig. 11 The iAdaptRT run time environment of the iAdaptTest

4 Evaluation

4.1 Aims and user data

The previous section described the architecture and the design decisions of a modular, rule based adaptive system, complying with Topic Maps, IMS LIP and IMS QTI. The iAdaptTest allows educators to create assessment activities and customize them according to a number of criteria. The focus of this paper is on how the tool presented increases reusability and interoperability of the e-learning content. Therefore, although the system can be utilized for diverse pedagogical aims and its significance can be evaluated under different perspectives, the focus of the present evaluation is on how the teachers feel about such a system and whether the produced learning content is indeed shareable.

The specific aims of the evaluation were to understand if:

(i) the tool is easy to use by educators
(ii) the educators will adapt the assessment using different rules
(iii) the supported set of rules is expanded enough to support different testing strategies
(iv) the content of assessments can be shared across different e-learning applications.

With the aid of eleven educators we ran an evaluation study of the adaptive assessment tool. Four of the educators were computer science trainers, two were teachers of linguistics in high schools, and the last five were teachers of physical sciences (chemistry, physics). The evaluations were performed in May 2008 with the first group of four teachers and in December 2008 with the remaining users. First, the tool was shown to each user for 30 minutes in order to help them understand its features and options. During the demonstration of the system its basic characteristics (multi-criteria adaptation, interoperability, etc.) were explained to the users.

The educators were then asked to prepare some formative and summative assessments for their classes. Both assessment types were run at the end of the courses in which the participants taught. The educators were free to create as many sections and questions per section as they judged necessary. The aim of the formative assessments was to allow learners who understand their current knowledge to prepare for the final assessments. Further, these assessments could help educators to understand the true knowledge of their students.

4.2 Analysis of the assessments

Each teacher created one formative and one summative assessment. 22 different assessments were eventually created. As seen in Tables 1 and 3, educators formed a different number of questions and sections per assessment. Further, they utilized different strategies especially in the formative assessments (see Tables 2 and 4).

More specifically, the formative assessments were based primarily on the prior knowledge or the goals of the learners. The educators preferred to base the self evaluations on these factors to help their students adjust the assessment to their needs. Educators created different number of questions and sections (Table 1). They created 1 or 2 or 3 sections and the questions in each section ranged from 5 to 18 questions. Each section was associated to a different topic and the difficulty level of each question varied from medium to difficult.

Tables 3 and 4 concern the data from the summative assessments. More specifically they present the number of the created sections and questions per assessment and the employed rules. In this case, the majority of the educators preferred to base their assessment primarily on the performance of the learners. The number of sections and questions were different

Table 1 Data related to the formative assessments

Formative assessments				
Educator	No of questions per section			Employed rules
	Section 1	Section 2	Section 3	
1.	18	–	–	Performance of learner
2.	15	15	–	Prior knowledge of learner
3.	7	7	7	Goals of learner
4.	6	6	8	Prior knowledge of learner
5.	5	5	–	Prior knowledge of learner
6.	8	8	–	Prior knowledge of learner
7.	9	8	–	Goals of learner
8.	10	10	–	Performance of learner
9.	10	10	10	Performance of learner
10.	15	15	–	Goals of learner
11.	8	8	8	Prior knowledge of learner

from the formative assessments. Taking into account that each section is related to a different subtopic of the main topic, the increased number of sections may be due to the need for the educators to test the knowledge of their students in different topics. Two of the educators did not apply any rule because they believed that in summative assessments all students need to be examined on the same number of questions. Nevertheless, the same educators did employ some adaptive rules in formative assessments. These different pedagogical approaches of the same educator are quite rational given the fact that they

Table 2 Adaptive rules employed by the educators in the formative assessments

Educator	Adaptive rule
1.	If a learner has answered correctly the first 10 questions then move to question no 15
2.	If a learner's knowledge level on the topic is high, then show questions of high difficulty only
3.	Ask the students to decide the difficulty level of the questions they wish to try
4.	Show questions that match the difficulty level of the student
5.	If a learner's knowledge level on the topic is high, then show questions of medium or high difficulty only
6.	Show questions that match the difficulty level of the student
7.	Ask the students to decide the difficulty level of the questions they wish to try
8.	If a learner has answered correctly the first 6 questions then move to next section
9.	If a learner has answered correctly the first 10 questions then move to next section or end the assessment
10.	Ask the students to decide the difficulty level of the questions they wish to try
11.	Start the assessment on the first question of each section that matches the knowledge level of the learner

Table 3 Data related to the summative assessments

Summative assessments				
Educator	No of questions per section			Employed rules
	Section 1	Section 2	Section 3	
1.	10	10	–	Performance of learner
2.	12	12	12	Performance of learner
3.	6	6	10	Performance of learner
4.	8	8	8	Prior knowledge of learner
5.	15	15	15	Prior knowledge of learner
6.	6	6	6	Prior knowledge of learner
7.	5	5	5	No rule applied
8.	6	6	6	Performance of learner
9.	8	8	8	No rule applied
10.	10	10	10	Performance of learner
11.	8	8	8	Performance of learner

originate from different goals. Our system is flexible and can accommodate all these different approaches.

Table 2 shows the specific adaptive rules as adjusted by the educators in the formative assessments. Although some of the rules were based on the same factors (i.e. prior knowledge), the final rules were quite dissimilar reflecting the pedagogical strategies and beliefs of the educators. For example, some educators based their adaptation decisions on the prior knowledge of the learners but they adjusted the same rule in different modes. These initial statistics are indications of the flexibility of the system to accommodate various pedagogical and testing strategies.

Table 4 shows that the educators were able to customize the existing available rules. Even though most of the educators preferred to base the adaptation to the performance of the learners, the constructed rules were quite dissimilar. In CAT systems [42] the questions to be shown are decided mechanically based on a complex set of statistics and thus the educators do not have full control over the system. We also need to note that in both formative and summative assessments the participants of our study utilized only one rule across the assessments, although the tool supports the application of different rules both per section and per assessment. This could allow the construction of more complex assessments, more tailored to the needs of the learners and the preferences of the educators.

4.3 Reusable content

During this evaluation study, concerning the authoring capabilities of the tool, 309 distinct questions were constructed. These questions were true/false, single choice and multiple choice. Some of the educators reused in the summative assessments some of the questions they had prepared for the formative assessments. Each question validated against the DTD (Document Type Definition) of IMS QTI v2.1 (i.e. the latest at the time the paper is written) to ensure full compatibility with the e-learning standard.

In a recent survey investigating the issue of true conformance to the QTI learning standard we found that some of the tools which claim to be QTI compliant have only a low,

Table 4 Adaptive rules employed by the educators in the summative assessments

Educator	Adaptive rule
1.	If a learner has answered correctly the first 8 questions then move to the next section
2.	If learner's score is higher or equal to 80% in the first 9 questions of each section then move to the next section
3.	If learner's score is equal to 100% in the first 5 questions of each section then move to the next section
4.	Show questions that match the difficulty level of the student
5.	If learner's knowledge level on the topic is high, then show questions of medium or high difficulty only
6.	If learner's score is equal to 100% in the first 5 questions of each section then move to the next section
7.	No rule applied
8.	If learner's score is equal to 100% in the first 4 questions of each section then move to the next section
9.	No rule applied
10.	If learner's score is equal to 100% in the first 6 questions of each section then move to the next section
11.	If learner's score is equal to 100% in each section then move on to the 4^{th} question of the next section

selective, conformance to the standard [27]. The ability of tools to export data which validate against the DTD of a QTI version and the ability of tools to import data were tested. Most of the tools failed to fully support these services. ECQuiz (http://plone.org/products/ecquiz/) was the only tool which could import and render correctly True/False and Multiple choice questions of the latest version of QTI. Our tool supports this QTI version as well. Therefore we randomly selected 44 of the constructed questions and imported them into ECQuiz to see if they are rendered correctly. Four questions from each educator were selected and the educators were asked to import them to the new tool. 43/44 (97.93%) of these questions were imported and rendered correctly as appeared in our tool. ECQuiz added automatically a new option called "I don't know" in each question as an alternative to not selecting any of the options. However, this did not affect the available options and the correct answers.

The images of the questions created with our tool, iAdaptTest, were stored in the same directory as the QTI XML file. But they had to be manually copied to a new directory of ECQuiz to render correctly. This is why one of the questions failed to render accurately. However, even in this case the import and rendering of the QTI item's text and options were correct. When the images were manually copied into the appropriate directory everything interpreted correctly.

As a final step in this test, the participants ran the newly formed test containing the 44 questions. The test was executed and completed without any difficulty by the participants.

We also managed to import the topic hierarchy to two other tools which are based on topic maps. The topic hierarchy could be projected without problems and then based on the abilities of the new tools it could be extended or altered. Unfortunately there are no freely available learner profiling tools abiding to IMS LIP so we could not perform any tests concerning this category of data.

These tests verify the hypothesis of this paper; that is content sharing is easy when tools abide closely to specific versions of e-learning standards. The modularized architecture of

our tool supports the independent utilization of the testing data in other e-learning applications. Each module stores its data on separate files, so the produced XML can be re-used in different application without any difficulty. Even the QTI rules are stored in different files than the QTI testing data in order to improve the reusability of the testing items. Further, the modularized nature of the adaptive testing tools permits the easy integration of new modules or the independent utilization of the authoring subsystems.

4.4 Questions to the educators

Following the construction of the assessments and the interoperability test, a short focused interview was conducted separately with each participant in order to record their opinions of the system. The following paragraphs present the questions and comment on the replies of the participants.

- *Did you find the tool easy to use?*

All the participants agreed that the tool is easy to use mainly because of the unified interface through which users can activate the respective subsystem. Also the fact that all the interfaces follow the same convention and aesthetics helped them in mastering the interface very quickly.

- *Do you consider the system useful?*

We asked the participants to comment on the significance of the system. All the participants agreed that the system is indeed useful as it allows them to test the students using different methods. This is one of the most important questions of the evaluation as it helps us understand the opinion of real users. Although, the number of the participants is not as large as we would wish, the unanimous concordance of the educators is a strong indication of the system's significance.

- *What is the most important feature of the tool?*

The next question asked to each educator was to identify the most important feature of the adaptive testing tool. We provided the users with the following list of choices:

– Different testing strategies
– Ability to create shareable content
– Easy to learn and to use
– Supports different question types
– Attractive interface
– Other (please specify): …….

These choices were shown in a different order to each user to avoid order effect bias. Eight of the eleven (72.72%) educators indicated the ability to customize the assessment procedure is the most important feature of the tool. The other three, who have used some automated assessment tools in the past, believed that the ability to share the content is quite important. The other educators also considered this option as very important although they focused primarily on the pedagogical aspects of the tool. That is why they consider the flexibility in applying different strategies as the most important feature of the tool.

- *Do you think that the adaptive rule set needs expansion?*

As expected all the educators agreed that the adaptive rules need some extensions to cover other aspects of the learners, such as their cognitive ability, possible disabilities,

learning style, computer environment, etc. Also they would like the existing rules to be more flexible. However, they also pointed out that a tool which might be too complex, may be unusable in the long run as the educators may not be willing to spend too much time in mastering it. The replies and the comments of the educators to this question will be addressed in our future research plans.

- *Would you use the system in your classes?*

We asked the users if they are willing to continue using this tool in their classes. Although they were some worries about the stability of the system, since it is the result of ongoing research, all the educators were willing to use the tool in their next courses. The main reason is that they feel that with such a flexible tool they can diagnose the real knowledge of their students better, with a shorter set of questions. Further, they mentioned that they can use the material in other tools so they are very positive about utilizing the tool.

- *Do you consider content sharing as an important feature?*

The final question posed, concerned the content sharing ability of the tool. We wanted to discuss the opinions of the educator with respect to this feature of the system. Ten of the eleven (90.91%) educators considered this as an important feature, while the other one considered it as a desirable feature. As discussed earlier in the paper, most of the educators consider the flexibility in the formation of the rules as the most important characteristic of it. Nevertheless, as the complexity of the educational applications and the penetration of software tools in education increase, interoperability and shareability will be vital. At least this is the main conclusion from the literature review and from the discussion with the educators.

5 Discussion and conclusions

This paper presents a highly modularized adaptive testing tool based on distinct e-learning standards. The system comprises of an authoring subsystem and a run—time subsystem. With the authoring environment teachers can create both adaptive and non-adaptive tests. Initially a topic hierarchy can be created and extended conforming to the Topic Maps ISO standard. This topic hierarchy is used to associate questions with specific concepts and could be easily expanded to form a graph of concepts since it abides to an XML based standard.

The next module of iAdaptTest is the learner profile management tool abiding to the IMS LIP e-learning standard. Teachers can manage the creation and maintenance of user profiles with the aid of this tool creating either individualized profiles or stereotype users. The learner profiles maintained in our system consist of a set of attributes which describes the personal data of learners, their formal education, their working experiences, their goals, their preferences and their assessment data. The IMS LIP profile management module is further used to display the accumulated learner knowledge at the end of the tests. Coding the data in an XML format grants us the freedom to provide alternative representations of the learner knowledge.

The questions and assessments are authored with the aid of an easy to use graphical tool. The structure of each assessment is visually presented to the authors and the data are coded in the latest version of IMS QTI. Each question may contain multimedia features such as images and video and for each question customization of a number of QTI attributes, such as the number of attempts, the difficulty level, the minimum and maximum score are

supported. The authoring tool may also be used as a standalone application to create interoperable items and assessments.

Another subsystem is used to add adaptive rules to the formulated assessments. These rules are treated as distinct XML entities and therefore they do not affect the assessment data. These categories of data are only combined during test execution. The adaptive rules concern the performance and a number of individual learner characteristics, such as their goals and their prior knowledge.

The run time environment combines all the data categories and executes the test, and at the same time updates the user profile. At the end of the assessment learners can privately access the profiles and view their progress through the IMS LIP learner profile management tool. At the moment, statistics are textually presented but alternative presentations are simple to create due to the hypermedia nature of the system.

A prototype of our highly modularized tool has been used by a number of educators in teaching examples. The participant educators created some summative and formative assessments and they utilized adaptive rules based on different criteria. During the evaluation we were able to import a representative sample of the questions and the created Topic Map in other educational applications. This experiment showed that having data that abide fully to a distinct e-learning standard stored in different repositories makes content sharing easier and more efficient. The mixture of various e-learning standards increases the data sharing ability of the tool, since every data category could be separately utilized in other tools. The focused interview with the educators helped us to realize that the data sharing ability and the flexibility in forming alternative testing paths are the most important features of iAdaptTest. In addition, the discussion with the participants allowed us to identify the rule extensions needed and to shape our future research directions.

To sum up, the main contribution of our research is the design and implementation of a highly modularized system complying to the Topic Maps, the IMS LIP and the IMS QTI standards. The tool supports efficiently the easy and independent utilization of the e-learning data and supports the future addition of new modules. The authoring modules of the system could be used for creating standards-based learning data which could be used to form adaptive assessments or they could be used in other educational applications. New releases of the e-learning standards could be easily assimilated as only the respective module needs to be maintained. The hypermedia nature and the simplicity of the interfaces helped the educators in easily utilizing the tool. In addition, the fact that each module offers a distinct set of functionalities reduces the learning effort compared to the overloaded interfaces of non-modularized tools. In any case more tests are needed in order to fully appreciate the pedagogical value of the tool and understand its potential shortcomings. Additional operations which could increase the usability of the system, such as semantic searching in the repository of the created questions and rules are needed to support more rapid test creation through the utilization of the existing data.

References

1. Arapi P, Moumoutzis N, Mylonakis M, Theodorakis G, Christodoulakis S (2008) A Pedagogy-driven Personalization Framework to Support Automatic Construction of Adaptive Learning Experiences. In: H. Leung et al. (Eds.) Proccedings of ICWL 2007, LNCS 4823, pp. 55–65
2. Bacon RA (2003) Assessing the use of a new QTI assessment tool within Physics, Proceedings of the 7th CAA Conference, Loughborough University, Loughborough, UK. Available at: https://dspace.lboro.ac.uk/dspace/handle/2134/1902

3. Boticario J, Santos O (2007) An open IMS-based user modelling approach for developing adaptive learning management systems, Journal of Interactive Media in Education, 2, http://www-jime.open.ac.uk/2007/02/
4. Chang W-C, Hsu H-H, Smith TK, Wang C-C (2004) Enhancing SCORM metadata for assessment authoring in e-Learning. J Comput Assist Learn 20(4):305–316. doi:10.1111/j.1365-2729.2004.00091.x
5. Conlan, O., Dagger, D., & Wade, V. (2002). Towards a Standards-based Approach to e-Learning Personalization using Reusable Learning Objects. In Proceedings of E-Learn 2002, October 15 - 19, 2002, Montreal, Canada
6. Conole G, Warburton B (2005) A review of computer-assisted assessment. ALT J Res Learn Technol 13 (1):17–31
7. De Bra P, Aroyo L, Cristea A (2004) Adaptive Web-Based Educational Hypermedia, in Levene and M. Poulovassilis A. (Eds.). Web Dynamics, Adaptive to Change in Content, Size, Topology and Use. Springer, pp 387–410
8. Dichev C, Dicheva D, Aroyo L (2004) Using Topic Maps for Web-based Education. Interntl J Adv Technol Learn 1(1):1–7
9. Dolog P, Gavriloaie R, Nejdl W, Brase J (2003) Integrating adaptive hypermedia techniques and open rdf-based environments. *Proceedings of 12th International World Wide Web Conference*. Budapest, Hungary
10. Dolog, P., Henze, N., Nejdl, W., Sintek, M. (2004) The Personal Reader: Personalizing and Enriching Learning Resources using Semantic Web Technologies, *Adaptive Hypermedia and Adaptive Web-Based Systems Conference*, LNCS 3137, pp. 85–94
11. Dublin Core (2008) Dublin Core Metadata Initiative http://dublincore.org/
12. Duitama F, Defude B, Bouzeghoub A, Lecocq C (2005) A Framework for the Generation of Adaptive Courses Based on Semantic Metadata. Multimedia Tools and Applications 25(3):377–390
13. Ghali F, Cristea A (2008) Interoperability between MOT and Learning Management Systems: Converting CAF to IMS QTI and IMS CP, *Proceedings of the 5th international conference on Adaptive Hypermedia and Adaptive Web-Based Systems*. LNCS 5149:296–299
14. Giacomini Pacurar E, Trigang P, Alupoaie S (2005) A QTI editor integrated into the netUniversité web portal using IMS LD. Journal of Interactive Media in Education, Vol. 9. Available at: http://jime.open.ac.uk/2005/09/
15. Harlen W, James M (1997) Assessment and Learning: differences and relationships between formative and summative assessment. Assess Educ Princ Policy Pract 4(3):365–379
16. IEEE LOM (2002) IEEE Standard for Learning Object Metadata http://www.ieeeltsc.org/standards/1484-12-1-2002
17. IEEE PAPI (2002) Public and Private Information, jtc1sc36.org/doc/36N0186.pdf
18. IMS LD (2003) IMS Learning Design, http://www.imsglobal.org/learningdesign
19. IMS SS (2003) IMS Simple Sequencing, http://www.imsglobal.org/simplesequencing
20. IMS LIP (2005) Learner Information Package, http://www.imsglobal.org/profiles
21. IMS QTI (2006) Question and Test Interoperability, http://www.imsglobal.org/question
22. Johnson K, Hall T, O'Keeffe D (2005) Generation of Quiz Objects (QO) with a Quiz Engine Developer (QED), *Proceedings of the 2005 IEEE International Workshop on Wireless and Mobile Technologies in Education WMTE'05*. University of Tokushima, Japan, pp 120–122
23. Kay J (1990) UM: a user modelling toolkit. *2nd Intl User Modelling Workshop* (Hawaii), 11
24. Kazi SA (2004) A conceptual framework for Web-based intelligent learning environments using SCORM-2004, *Proceedings of the 4th IEEE Advanced Learning Technologies (ICALT)*, pp 12–15
25. Koper R (2003) Combining reusable learning resources and services to pedagogical purposeful units of learning. In: Littlejohn A (ed) *Reusing Online Resources*: *A Sustainable Approach to eLearning*. Kogan Page, London, UK, pp 46–59
26. Lazarinis F (2006) A hypermedia environment for managing online courseware based on adaptable templates, International Conference on Virtual Learning, 239–246
27. Lazarinis F, Green S, Pearson E (2009) Measuring the conformance of hypermedia assessment tools to QTI. Int J innov Learn 6(2):127–146
28. Lesage M, Riopel M, Raiche G, Sodoke K (2008) IMS-QTI sub-standards in computerised adaptive testing and interfacing. Interntl J Adv Media Commun 2(2):115–137
29. Martínez-Ortiz I, Moreno-Ger P, Sierra JL, Fernández-Manjón B (2006) <e-QTI>: a Reusable Assessment Engine, 5th International Conference on Web-based Learning, Penang, Malaysia. LNCS 4181:134–145
30. Mills C, Potenza M, Fremer J, Ward C (eds) (2002) Computer-based testing building the foundation for future assessment. New York, Lawrence Erlbaum Associates
31. Morimoto Y, Ueno M, Yokoyama S, Miyadera Y (2007) A SCORM-compliant Learning Management System that Enhances Learning By Managing the Learning Itself, Proceedings of the 7th IEEE International Conference on Advanced Learning Technologies (ICALT) pp 122-126

32. Musa DL, Oliveira, JPM (2004) Sharing Learner Information through a Web Services based Learning Architecture, In: WISM-CAISE 2004, Riga, 2004
33. Paramythis A, Loidl-Reisinger S (2004) Adaptive Learning Environments and e-Learning Standards. *Electronic Journal of eLearning, 2* (1), http://www.ejel.org/volume-2/vol2-issue1/issue1-art11-paramythis.pdf
34. Parmar K, Anane R, Hendley R (2007) Architecture of a SCORM-Compliant Assessment Authoring Tool, *5th International Conference on Computational Science and Applications*, pp 111–117
35. Rey-López M, Díaz-Redondo R, Fernández-Vilas A, Pazos-Arias J, López-Nores M, García-Duque J, Gil-Solla A, Ramos-Cabrer M (2008) T-MAESTRO and its authoring tool: using adaptation to integrate entertainment into personalized t-learning. Multimedia Tools and Applications 40(3):409–451
36. Sadler DR (1989) Formative assessment and the design of instructional systems. J Instr Sci 18(2):119–144
37. SCORM (2004) Sharable Content Object Reference Model, http://www.adlnet.gov/Scorm/
38. Topic Maps (2000) Topic Maps, www.topicmaps.org
39. Topic Maps (2002) *ISO/IEC 13250 Topic Maps 2n ed.*, http://www.y12.doe.gov/sgml/sc34/document/0322_files/iso13250-2nd-ed-v2.pdf
40. Torres da Silva V, Pereira de Lucena C, Fuks H (2001) ContentNet: a framework for the interoperability of educational content using standard IMS. Comput Educ 37(3–4):273–295
41. Towle B, Halm M (2005) Designing Adaptive Learning Environments with Learning Design. In: Koper R, Tattersall C (eds) Learning Design—A Handbook on Modelling and Delivering Networked Education and Training. Springer, Berlin-Heidelberg, pp 215–226
42. van der Linden WJ, Glas CAW (eds) (2000) Computerized adaptive testing: Theory and practice. Kluwer, Boston, MA
43. Van Rosmalen P, Vogten H, Van Es R, Passier H, Poelmans P, Koper R (2006) Authoring a full life cycle model in standards-based, adaptive e-learning. Educ Technol Soc 9(1):72–83
44. Wainer H, Mislevy RJ (2000) Item response theory, calibration, and estimation. In: Wainer H (ed) Computerized adaptive testing: a primer. Lawrence Erlbaum Associates, Mahwah, NJ
45. Zualkernan IA, Ghanam YA, Shoshaa MF, Kalbasi AS (2007) An Architecture for Dynamic Generation of QTI 2.1 Assessments for Mobile Devices Using Flash Lite, *Proceedings of the 7th IEEE International Conference on Advanced Learning Technologies* (ICALT), pp 194–195

Fotis Lazarinis is a part-time lecturer at a Technological Educational Institute in Greece. He has authored over 45 refereed papers in international or national conferences, journals and research handbooks. He has also published several Computer Science educational books in Greek and served as a review member for conferences and workshops.

Dr Steve Green is Principal Lecturer in Digital Media in the School of Computing at the University of Teesside, England. He is the section head for Digital Media and the Web and also the Director of the RIME research centre (Research into Interactive Multimedia Education). He researches and publishes in the areas of multimedia design, learning technologies and the use of computers with people with special needs.

Dr Elaine Pearson is Director of the Accessibility Research Centre at the University of Teesside, England. She has research interests in all aspects of computing to support the needs and preferences of disabled students. Since the award of a study abroad fellowship with the Leverhulme Trust in 2000, Elaine has been a Visiting Research Fellow at the University of New South Wales, Sydney, Australia. She has published and presented nationally and internationally on accessibility issues and is a member of the conference review committees for EDMEDIA and ALT and a referee for a number of journals including ALT-J and IJEL.