Towards the Interoperability of Online Assessment Tools

Vladimir Tomberg, Mart Laanpere
Tallinn University, Narva mnt. 25, 10120 Tallinn, Estonia
vtomberg@tlu.ee, mart.laanpere@tlu.ee

Abstract. The paper examines the opportunities for enhancing the interoperability of online assessment tools. The comparative analysis of different versions of specifications IMS Question and Test Interoperability is made; adaptation of new technologies in test systems are examined. Special focus is on the use of the Web services for quiz authoring, delivery and storage in connection with development of an open source Learning Management System IVA in Tallinn University.

Keywords. E-learning, assessment, interoperability, standards.

1. Introduction

Today it is difficult to imagine university in which forms of e-learning would not be used. Except for traditional use of electronic materials and communications, the increasing number of institutes has been implementing online assessment in educational process. The role of the online assessment in higher education is difficult to overestimate. E-assessment provides not only multiple-choice tests – it can contain surveys, writing of an essay, simulators, polls and other kinds of tasks. Development of test items can be quite work-some and expensive, this is why reuse of existing content is very desirable.

This paper discusses the opportunities for developing interoperable online assessment systems and tools using the latest Web technologies. The main objectives of our study are: (1) to provide a critical analysis of existing industry-driven specifications and their impact on the development of online assessment tools, (2) to demonstrate the new opportunities for developing next generation assessment tools with Web services and Business Process Execution Language, (3) to describe the conceptual design of a new Web-based quiz tool that is being developed in Tallinn University. We start with an overview of Question and Test Interoperability (QTI) specification and its dynamics during the last five years, then discuss the implementation of QTI for online assessment Web services as well as for the new quiz tool within our open-source learning management system IVA.

2. Question and Test Interoperability specifications

Since 1999, Instructional Management Systems (IMS) consortium (see imsglobal.org) has been developing a technical specification for Question and Test Interoperability (QTI). The IMS QTI Specification is the key specification and de facto standard for developing re-usable content for online assessment. The most important milestones in the development process of IMS QTI are versions 1 (1999), 2 (2003) and 2.1 (2006). Despite of the availability and supremacy of the latest versions, the most widely used QTI version is still 1.2.1. Now the attention of developers is turned to version 2.1 because the architecture of the online assessment tools has been revised. Changes are so significant that they demand a closer look before we proceed to the application side.

2.1. The first generation of QTI specifications

The main aim of the authors of specification was ensuring the possibility for exchanging the questions and tests between different assessment-related systems and tools. The majority of the quiz tools that were created before the appearance of specifications had their own (usually closed) format for storing the questions and test data. The QTI specification describes a data model (ASI: Assessment-Section-Item) and its bindings to XML schema for representing tests and their results in an uniform way, making the exchange and re-use of assessment-related content between different systems possible.

Figure 1 illustrates the architectural recommendations of IMS consortium for constructing an online assessment system in conformance with QTI specification ver. 1.2.
The main components of an assessment system are: quiz authoring tool, repository for quiz items and tests, assessment engine that delivers and executes the Web-based quizzes, as well as a learning management system (LMS). All these components should be able to exchange with each other the data in QTI-compliant XML format. This is a theory, but in practice there is no system in which all of communications are implemented and orchestrated in full scale as it was described above. In most of the cases, exchange of QTI-compliant data occurs between the pairs: between two repositories (most often), between an authoring tool and a repository (seldom), between a repository and LMS (extremely seldom). On of the reasons is that the majority of quiz authoring tools and learning management systems have incorporated a repository as an integrated unit, reducing the immediate needs for QTI-compliant data exchange.

At the moment of appearance of the first specification, QTI, IMS was developing five more specifications. One of them, IMS Content Packaging Specification (CP), was offered to joint usage with QTI for the contents packing. Another, IMS Learning Resource Metadata Specification was used for introduction of metadata in tests.

Advantages of the QTI specification, absence of common formats in systems and prospects of growth of test systems have led to forthcoming of the products declaring to some extent conformity to specifications. University researchers and developers of commercial assessment software were engaged in implementation of the QTI. Unfortunately, it is necessary to state that impact of this work has been quite limited. In one of the early reports [2], Whittington wrote: 'The transfer process at present is crude and requires a large degree of expert knowledge and intuition to succeed'. The study was conducted in 2001 when the systems using QTI only started to appear. At that time, only three packages of the software were checked and data between them were transferred badly. Two years later the number of the QTI supporting systems considerably increased. Gorissen [3] carried out research in which already 6 commercial products declaring the compatibility with QTI were checked. The conclusions were pessimistic: 'None of the applications tested in this quickscan have support for all options of the QTI specification'. The similar study (with similar results) was conducted by one the authors of the current paper in 2005 [4], in which five software packages (including the Learning Management System IVA) were validated for QTI compatibility.

Why the unique specification existing in the field of tests is so difficult to implement in practice? What are the basic problems at implementation of specifications? Firstly, it is observed in all researches, those specifications are very complex themselves, and their realization in the software demands greater resources. It is difficult to small universities to involve big groups of researchers and programmers into the project. Secondly, many advantages of specifications in practice unexpectedly turn to obstacles. For example, the specification offers five types of the basic Item types: Logical identifier, X-Y co-ordinate, String, Numeric, Logical groups. Combination of these types of Items with variety of types of visualization (render_slider, render_choice, render_hotspot etc.), gives a possibility to receive 180 potential combinations, the majority of which will be useless [5]. A freedom of a choice often leads to that the test created in one system is passes check on validity in another system, but cannot be visualized in it.

Complexity in realization of the tests exchange originates from a problem that the specification does not provide any control over import of tests. Usually system simply informs on an error, in any way without making comments on the reasons of its occurrence. It is necessary for an operator to search personally for the reasons, mainly manually analyzing a source code. Examining QTI 1.2 the author has already mentioned this problem [4] and suggested to create the special protocol of declaring of the system opportunities. Test systems by means of the similar protocol before the beginning of import of tests could check an opportunity of such import and analyze a difference between their opportunities.
and opportunities of another system. It would allow preventing attempt of import of the test, which initially cannot be imported because of a difference in opportunities of a system.

As the specification does not establish rigid requirements to the use of content packaging (textual tests can be transferred in the form of pure XML), images, audio, Flash and video files can be lost.

In that way, the first version of specification QTI, despite of all the advantages in real projects worked badly or did not work at all.

2.2. The second generation of QTI

It is necessary to note, that consortium IMS is very sensitive to a feedback incoming from developers. In version 2.0, also in the following version 2.1 the solutions practically to all problems of the first version have been offered.

Among the important differences between versions it is possible to note a refusing from data model ASI, instead of it now the new model Test, Section, Item is used. This model is based on interactions and includes support to dynamic materials. It has resulted, in turn, to greater positive changes in Item Content model. Now specifications define 15 types of the interaction, each has a unique name-identifier. Among them except for traditional interactions (Hotspot, Text Entry, etc.) there were also such new types, which could add the image and load a file on a server. The method of visualization now does not depend directly on a type of interaction, as it was the first QTI version. The developer has an opportunity to use a necessary method of visualization not worrying about restrictions, related to concrete types of visualization. On the one hand, such rigid approach looks a little bit limiting; on the other hand, it solves the most of problems of tests portability even between very different systems.

Realization of visualization is thoroughly revised. IMS 2.0 includes support of XHTML that enables, creating tests to be assured in adequacy of the final visualization result. The use of XHTML simplifies standard formatting of the text, RTF formatting is not used any more. Together with the use of Cascade Styles Sheets (CSS) XHTML gives authors of tests greater opportunities for the control of visualization. Use of XHTML element <object> enables easily to include in the test alternative media. At all of the advantages, implementation of the XHTML does not break validity of QTI XML files which absence often caused problems in the first version of QTI realizations.

Innovation in version 2.0 became introduction of question templates. The given technology allows generating a question in the form of a template based on which during the test clones of a question are created, with the data changing according to the prescribed algorithm. At use of templates students who execute identical tasks cannot see a right answer of their neighbors as their data will differ from their ones.

Changes have occurred in a feedback to the user. A feedback now is divided into a modal and an adaptive. As it was in the previous versions of specifications, a Modal Feedback does not depend on actions of the user. A new Adaptive Feedback allows to choose the information for the user depending on results of his attempt, for example, to ask additional questions or to change sequence of execution of the test. Thus, functionality of an original question changes depends on actions of the user [6]. Inline Feedback one and more innovation in the 2.1 specification. It allows to display a feedback to the user as separate parts during the test and not just by its end. A combination of the adaptive and built in feedback enables the authors of tests to make very complex, "intelligent" tasks forming educational process.

Version 2.0 of QTI specification dictates much more strict rules defining cooperation with specification CP. In the previous versions, there was an opportunity to place metadata at one's own choosing, both in a file of the manifest and inside of a file with content of tests. Now metadata are located only in separate IEEE LOM manifest and are coordinated with specification IMS Learning Resource Meta-data Specification. Thus, if the tests intended for transfer are accompanied by metadata they cannot be transferred unpacked. It eliminates a problem of the data loss described above. As to transfer files of media unpacked has become impossible, it is probable, that in the near future practice of an exchange unpacked QTI XML files will end.

Appearance of new versions of QTI specifications has brought out also the new approach to architecture of learning systems. The use cases (Fig. 2) describe the basic tools used during work with tests: authoring tool, as tools for creation separate an assessment item, an item bank, as a system of storage and service of assessment items, test construction tool, as the tool for assembly of tests from individual assessment item, assessment delivery system, as a system for de-
livery of the contents to the student and gathering of the information on results of tests. These tools are presented in a figure in the form of separate modules.

![Figure 2. The Role of Assessment Tests and Assessment Items. [11]](image)

As it is possible to notice, interaction of modules is precisely divided into two levels - assessment item and assessment test. For example, authoring tool has no access to the configured tests, and assessment delivery system cannot operate with separate items. Exceptions are: Test Construction Tool, which is used for assembly of tests from separate items and the module Learning System that can cooperate with all external modules at both levels. (In practice, Learning System usually simply includes functionality of all modules together.) Nevertheless, the offer to examine functional parts of system as separate modules is a serious step forward, in comparison with the first versions of specifications where the basic purpose was external interaction between monolithic systems. It is especially interesting if to consider, that for interaction of all modules one universal language QTI XML is used. Such distributed system offers higher level of interoperability. For example, assessment items can be prepared in one system, are stored in bank of other system, in the third system from them it is possible to collect tests, and the fourth system allows to deliver these tests to the user. Moreover, all these systems can theoretically be from different vendors.

Unfortunately, now similar interaction in a reality is improbable. The reason is not only that all existing systems have support the QTI. The most of systems, as a rule, have no such separate, strongly pronounced modules as Test Construction Tool, and realized in monolithic systems Item Banks not always allow the handling to the separate assessment item.

The most important in the new architectural approach is interaction Item Bank - Test Construction Tool - Assessment Delivery System. Even at absence of other components, the opportunity of such interaction theoretically allows to realize an exchange of tests between two and more systems. In practice in existing systems, transfer of tests will be realized manually under the control of the operator. Such approach suits an exchange between monolithic systems. For an exchange between separate components, more operative methods are required. This circumstance has forced developers to search for additional opportunities among technologies available at present.

3. Web Services for online assessment

The use cases examined above shows, that the second version of QTI specifications was initially developed in the context of a service-oriented paradigm [7]. The service-oriented approach allows to realize behavior of the application (LMS, VLE, etc.) using Web-services which can be used by other applications [8]. Web-services are based on several industry standards: WSDL for the description of service, UDDI for informing and the publication, SOAP for an exchange of messages.

One of the major advantages of Service-Oriented Architecture (SOA) is the modularity [7]. By low expenses, developers have an opportunity quickly to connect existing applications. At forthcoming new applications, only it is necessary to add new modules of Web-services. Such opportunity has an extra value in case of QTI specifications. The problem originates from that QTI now has not been standardized yet. Each new version of specifications expands its functionality. Thus, a cycle of development of the software product especially connected with requirements of high complexity, such type, as QTI, long enough. Developers should trace on the run changes in specifications, or they will be forced to release an obsolete product. In such situation, incremental addition of Web-services can provide support of new appearing services and tools.

Today's popularity and prevalence of Web-services is a good reason for their fast adaptation to existing applications in the field of education. Confirmation to it is a plenty of Web-services projects for communication between tools related to development of assessment tools. Among the
most known and well-documented projects it is important to point out the following ones:

1) APIS: Assessment Provision through Interoperable Segments. Provides a modular online assessment rendering engine and services, which transform IMS QTI v2.0 content into XHTML;

2) ASSIS: Assessment and Simple Sequencing Integration Services. Tools to enable a teacher to browse, search, preview and select assessment items from item banks for incorporation within content packages, which may include associated sequencing instructions;

3) Serving Maths. The project developed open source software tools to address issues around the use of mathematical expressions in assessments. One of the outputs of the project was the Remote Question Protocol, which supports remote processing of assessment items on behalf of assessment systems and therefore maximizes the range of item types and formats it is possible to deliver within a single assessment session;[1]

4) R2Q2: Rendering and Responses Processing for QTiv2 Question Types. R2Q2 is a definitive response and rendering engine for QTiv2 questions. While this only deals with an Item in QTi terms, it is essential to all processing of QTi questions; that is, it forms the core component of all future systems. Due to the design and use of internal Web services, the system facilitates future enhancement and can be changed to suit any application. [7]

Practice of the first development related to Web-services shows, that interaction between two applications in the field of the assessment is complex for organizing within the boundaries of one Web-service. The amount of transactions between tools within the limits of one activity is too great, and ways of realization can vary from system to system. In this connection the core task shares on separate tasks, according to them functionality, and then these small tasks are distributed between different Web-services. For example, thus Web-services are realized in R2Q2 project [7]. Because of this, the use of different sets of functionality in Web-services can lead to reduction in the common interoperability. Therefore, developers require standards that would describe patterns of Web-services behavior in the distributed multivendor system. Such standards give framework, reference models and ways and without them creation of compatible interfaces is impossible.

Development of standards in this area now occurs in the several organizations. IEEE offers Learning Technology Systems Architecture (LTSA) in which the architecture based on abstract components is described.

Other popular initiative is e-Framework, developed in cooperation UK's Joint Information Systems Committee (JISC) and Australia's Department of Education, Science and Training (DEST). E-Framework offers the service-oriented architecture based on opened standards in the field of education and sciences. The purpose of developers of the e-Framework is the reference models describing interaction of processes in different areas of education, precise service definitions and specifications [9]. In the field of the assessment e-Framework offers own reference model developed in the context of the FREMA project (Framework Reference Model for Assessment). The mentioned reference model describes workflows and processes in a number of domains, checked against a wide number of users and stakeholders in each domain. Also the project offers very useful use cases diagrams that in detail describe interaction of all processes and transactions.

As the toolkit intended for teamwork of Web-services from different suppliers the BPEL language (Business Process Execution Language) recently is often mentioned. As well as many other technologies (XML, WS, etc.) BPEL initially has not been developed for needs of electronic education. It appeared in 2002 as a result of combined efforts of BEA, IBM, Microsoft, SAP and Siebel. For today, the specification of language has the draft status version 2.0. The BPEL is XML based language and intended for separation of tasks in the distributed computing using Web-services. It is often named as a language for orchestrating Web services into flexible applications. In system which uses BPEL all requests of Web-services are intercepted, and then answers are sent to various points of BPEL-process. Such intervention allows to determine a policy that controls all of web-services actions - that is, access, authorization, registration, balancing of loading - and then wrap this policy into web-services [10]. For the present moment the most known e-Learning project using BPEL is Assis, subsidized JISC. Here BPEL it is used for interaction of two Web-services, QT1 and SS (Simple sequencing). As a result, developers have presented flexible system for realization of complex adaptive tests.

4. Implementing QTI interoperability
IVA is an open-source Learning Management System that has been developed since 2001 by the Centre for Educational Technology (CET) in Tallinn University, Estonia [12]. Among other typical LMS functionalities, IVA contains the tools for authoring, delivery and storage of the tests. Tests can be exported from IVA in QTI 1.2 compatible XML format. The need for upgrading to QTI 2.1 compatible data exchange has not been perceived as a critical one for the users and developers of IVA, but at some point it has to be accomplished. In the light of recent trends towards Personal Learning Environments (PLE, [13]), we have initiated a development of the next generation of our LMS called IVA 2.0 that is based on the Service Oriented Architecture. In IVA 2.0, quiz authoring, delivery and storage functionalities are used as the Web services. An original quiz item and test authoring tool is being developed by CET, quiz item rendering Web service is provided by R2Q2, SCAM is used as a repository for quizzes. This solution allows to use the same services also independently from IVA system and to share quizzes between different applications that are currently being developed in the CET (e.g., ePortfolio, the repository of interactive digital worksheets, job safety competency test system, traffic exam system, school information system, the online system for national tests of ICT skills). Our aim to create a central quiz repository that supports massive production, distribution and re-use of quiz-like digital content within the education and training system in Estonia. The work is still in progress, but the first prototypes will be introduced on the ITI conference.

5. References