Using open technical e-learning standards and service-orientation to support new forms of e-assessment

Using open technical e-learning standards and service-orientation to support new forms of e-assessment

Yongwu Miao, Colin Tattersall

Educational Technology Expertise Center,
Open University of The Netherlands, Heerlen, Netherlands
E-mail: {yongwu.miao, colin.tattersall}@ou.nl

Judith Schoonenboom

SCO-Kohnstamm Institute,
University of Amsterdam, Amsterdam, Netherlands
E-mail: J.I.Schoonenboom@uva.nl

Krassen Stevanov, Adelina Aleksieva-Petrova

Faculty of Mathematics and Informatics,
University of Sofia, Sofia, Bulgaria,
E-mail: {krassen, adelina}@fmi.uni-sofia.bg

Abstract: In this paper, we propose a new technical approach to support new forms of e-assessment. This approach is based on international e-learning standards and a service-oriented approach. Through a combined use of IMS LD and IMS QTI, new forms of assessment can be modelled as a unit of assessment, a specific unit of learning with a set of QTI documents or/and specific assessment services. The unit of assessment can be executed in any standard-compatible run-time environment. In comparison with traditional software development approaches, our approach fosters interoperability, flexibility, and seamless integration with learning activities.

Keywords: new forms of assessment; IMS LD; IMS QTI; service-oriented approach; unit of assessment.

1 Introduction

New forms of assessment, such as self- and peer assessment, 360 degree feedback, and portfolio assessment, are gaining in acceptance and popularity. Such assessment types are not just 'done to' learners but are also 'done with' and 'done by' learners (Harris and Bell 1990). By addressing complex student traits, these new forms aim to foster deep learning and the development of competences (Topping 1998; Boud, Cohen et al. 1999; Gipps 1999).

In comparison with traditional assessment, both judgment making and administrative processes are more problematic in new forms of assessment, which are process-based and

involve multiple roles and multiple persons. The difficulties and the potential for errors and omissions increases in a non-linear fashion as the number of candidates and assessors involved grows (Rosbottom 1994). As Bartram pointed out, 360-degree feedback by its very nature is an administrative nightmare to manage. People involved in the process tend to be geographically dispersed but also need close supervision in order to ensure that the ratings are carried out to schedule and that sufficient assessors are obtained for each focus of the assessment (Bartram 2005).

Many software tools such as SPARK (Freeman and McKenzie 2002) and eSPRAT (Lockyer 2003; Davies and Archer 2005) for supporting new forms of e-assessment have been developed in recent years. They are typically stand-alone and offer limited support for interoperability and reusability of assessment resources. In order to solve these problems, a standard-based approach is a better choice. The leading e-learning standard for the exchange and interoperability of assessments is IMS Question and Test Interoperability (QTI 2003). However, QTI can not be used to model process-based, multi-user assessments. IMS Learning Design (LD 2003) can be used to model learning processes with complicated process-control and multiple roles/users. However, assessment tools and strategies are not explicitly included in LD.

This paper's claim is that the combination of LD, QTI, and specific assessment services is able to model and deliver new forms of e-assessment. The main benefit of this approach is that existing specifications, tools and services can be used to model and deliver integrated learning designs with innovative assessments. Furthermore, reuse of such integrated learning designs and innovative assessments becomes possible together with their delivery in different platforms.

2 Characterising new forms of assessment

We distinguish new forms of assessment from more traditional approaches along four lines:

- Involvement of multiple roles/users. New forms of assessment are typically embedded in an educational context, require more stipulation of the processes of assessment and rely on higher levels of student involvement (Sluijsmans et al. 2004). Therefore, when modelling an innovative e-assessment process, multiple roles must be modelled.
- Variety in task types. Various types of tasks are performed in assessment processes. On the one hand, tasks are arranged for candidates to demonstrate their progress and capabilities such as answering a questionnaire, writing an article, providing a portfolio, conducting a performance, and so on. On the other hand, certain types of tasks will be performed by assessors for describing, collecting, recording, scoring, and interpreting information about students' learning. The types of tasks required to be performed depends on the nature of the trait to be assessed. In addition, simulation tools and domain-specific application tools may be used for assessing competences such as the use of concept mapping to assess knowledge structures, or the use of latent semantic analysis to interpret student essays (Pellegrino et al. 2001). Supporting new forms of assessment requires explicitly modelling various types of tasks.

- Complex process control. In innovative assessment processes, various tasks are carried out by many participants with different roles in sequence or in parallel. The termination of one task may trigger the start of another task. New forms of assessment require the modelling of complicated control-flows to coordinate various tasks performed by participants with different roles in sequence or/and in parallel.
- Exchange of information. In new forms of assessment, a large quantity of information is produced by participants in performing various tasks in different phases. The information must be transferred to the right persons at the right time. People with different roles interact with each other through the exchange of information. In order to model new forms of assessment, there is a requirement to model dataflow explicitly.

3 A technical approach to supporting new forms of e-assessment

This section presents how a combined use of LD, QTI, and specific assessment services can meet the four requirements identified above. New forms of assessment can be modelled as a unit of assessment, a specific unit of learning referring to QTI documents and/or specific assessment services.

- Supporting multi-role/user-involved assessment processes. QTI specification is concerned with individual learners. Although QTI does not prohibit usage in contexts involving other actors, it does not support explicitly defining other roles or sequencing behaviours that result from participation of other actors. LD can support a multi-role/user teaching-learning process.
- Supporting a variety of assessment tasks. The QTI can support item types including multiple choice, open-question, fill-in-blank, hotspot, match, drag&drop, and so on. It also provides sufficient flexibility to grow into advanced constructed-response items and interactive tasks we envisage as the future of assessment elaborates the assessment items in detail (Almond et al. 2001). Furthermore, it provides mechanisms to design structured assessment and control branches and calculate weighted scores. That is, all standard assessment tasks and structured assessment that form the core subset of current practice can be supported by using QTI tools. LD offers an approach to integrate application tools as services. Although only four services are specified in LD, in theory, any software tool can be integrated as an external service. Therefore, with an appropriate interface, any specific assessment tool (e.g., a portfolio editor, a concept-mapping, and a simulator) can be integrated as an external service into a unit of assessment.
- Supporting complicated control-flow. LD can support the modelling of a learning flow with complicated process controls. Activities can be arranged as a sequence or a selection structure. A set of role-parts can be performed in parallel within an act and acts within a play will be carried out in sequence. Multiple plays can be executed as concurrent threads. The termination of a task may trigger the start of another task according the definition. In addition, properties, conditions, and notifications provide more powerful mechanisms to control the process. Considering the complexities of

new forms of assessment in process control, LD has sufficient expressiveness to model the complicated control-flow in new forms of assessment.

- Supporting complicated dataflow. QTI version 2 provides mechanisms for declaring outcomes and specifies how an outcome variable can be coupled to an LD property. With the help of this mechanism, the data (e.g., an article) produced by a participant (e.g., a candidate) can be transferred to another one (e.g., an assessor). Additionally, scores given by all assessors can be processed into a collective (aggregated) result. This result can be transferred to the candidate or even can be used to control the branching.

4 Modelling and executing new forms of e-assessment in today's infrastructure

We have modelled a dozen of innovative assessment examples by using our approach. Because of the limitation in space, only one example is presented in this section in order to explain how to model and execute new forms of assessment.

Figure 1 shows a process model of a peer assessment example taken from (Orsmond 2004).

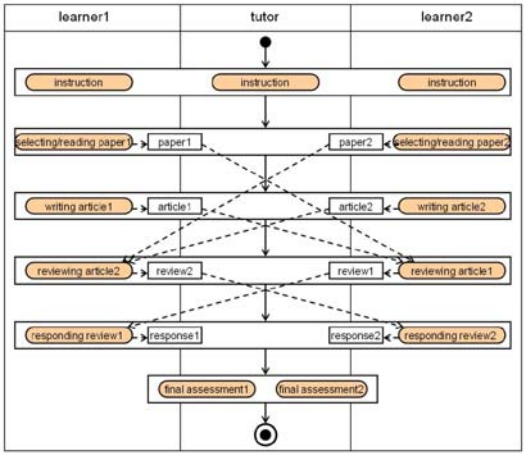

Figure 1 Process model of the peer assessment example

- Modelling roles. In this peer assessment example there are two kinds of roles: tutor and learner. In order to explicitly model the tasks of each peer student and the exchange of information between them, learner1 and learner2 are defined as two sub-roles of the learner.
- Modelling tasks. Participants with different roles are assigned to do different tasks. The tasks are modelled as learning activities (e.g., selecting/reading paper1 and responding review1) and support activities (e.g., final assessment1) in the model. Each activity has an element called activity-description, some of which (e.g., writing article1 or reviewing article2) refer to QTI documents.

- Modelling control-flow. The overall assessment process is defined as a play with six acts illustrated in the Fig. 1. Each act (represented as a box) contains more than one role-part (represented as a rounded rectangle). In the first act, the tutor teaches learners how to conduct this peer assessment and what is expected. In the second act, two peer students select a different paper respectively and read the selected papers. In the third act each student writes an article. In the fourth act students review the articles of their peers and comment on them. In the following act they response to the reviews of their peers and revise the original article if necessary. In the last act, the tutor assesses the students' work and give them scores. All acts are executed in sequence. The arrows with solid lines in Fig. 1 indicate the control-flows of the process.
- Modelling data-flow. Data are represented in LD as properties. A property can be used to record the outcome of the learner (e.g., article1 and review1) or to capture the current state of the process (e.g., are-articles-submitted). As we see in Fig. 1, data (e.g., article1, article2, review1, review2, and so on) are produced by a learner in an activity and will be used by another learner in another activity. The arrows with dash lines indicate the data-flows in the process. Viewing the value of a property is realized by using "view-property" element in a XHTML document, which is modelled as a learning resource and will be referred to by an item. The item is defined in a learning object within an environment. We define two environments for storing data regarding to the work of two learners, respectively. For example, environment1 will be associated with all activities handling article 1 such as selecting/reading paper1, writing article1, reviewing article1, responding review1, and final assessment1. Since all data concerning article1 is collected in this environment, this shared environment can be used by learner1 writing article1, by learner2 reviewing article1, and by tutor assessing learner1's work.

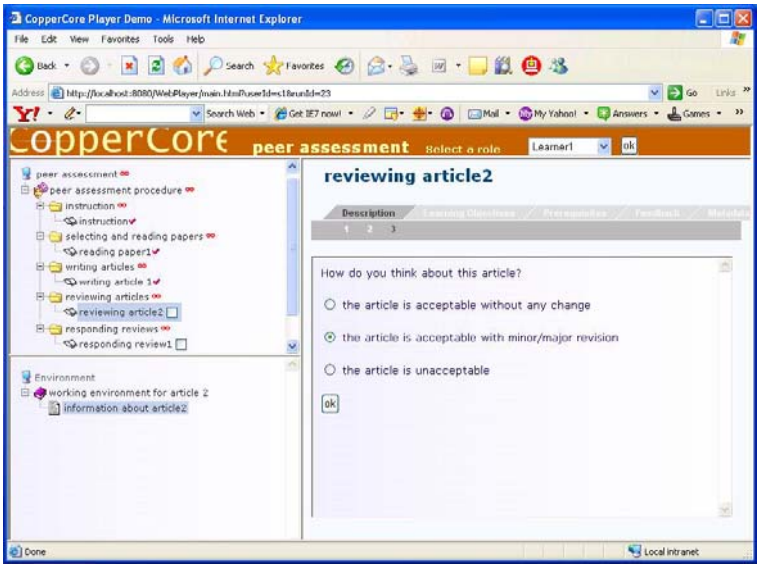

Figure 2 A screenshot of execution of peer assessment example

This process model with necessary QTI documents, wrapped as a unit of assessment, can be executed in any standard-compatible run-time environment. Figure 2 shows a screenshot of execution of the peer assessment example in an existing run-time environment including CopperCore (Vogten, Martens et al. 2006) and Apis (2004).

5 Discussion

Like the peer assessment example described above, other innovative assessment examples are modelled and tested. The test results demonstrated the feasibility of our technical approach to support new forms of e-assessment. In comparison with typical software development approaches, our approach has, at minimum, three advantages.

- Support for interoperability. Existing software tools for new forms of assessment have been developed and used as standalone application tools. They have their own data representation that is not usable by other applications. Their functions cannot be shared directly by other software. In contrast, our approach is based on international e-learning technical standards. A unit of assessment can be executed in any LD player with any integrated QTI player. The components of a unit of assessment or the unit as a whole can be stored, retrieved, and adjusted for reuse in different context and in different learning platforms.
- Support for flexibility. Each new form of assessment may vary in a number of variables. For example, in peer assessment the variables could include levels of time on task, engagement, and practice, coupled with a greater sense of accountability and responsibility (Topping, Smith et al. 2000). Software may support flexibility to a limited extent. However, once software has been developed, it is difficult to change to fit different learning contexts and specific needs. In particular, if a certain domain-specific application tool is needed as a specific assessment tool, it will be very difficult for existing assessment software tools to extend their functions. However, by adopting our approach, we can easily modify the definitions of components and their relations in a model (e.g., learning activities, assessment activities, their sequence, referred learning objects and assessment items, integrated specific assessment services, and so on). Assessment developers can customize their units of assessment with less effort and time.
- Support for seamless integration with learning processes. Existing software tools for supporting new forms of assessment can be used for formative assessment and summative assessment. However, the integration between learning activities and assessment are manually implemented. That is, the users have to manually shift learning environments or application tools for performing learning activities and for conducting assessment tasks. Our approach is based on LD that can formally describe a wide range of pedagogical approaches (Koper and Olivier 2004). If learning activities are also represented in LD, there will be a seamless integration between the learning activities and new forms of assessment, since both are be specified within the same unit of learning. For example, if an e-assessment is arranged as a formative assessment, the assessment results defined as outcome variables in QTI can be used by LD engine as properties to choose appropriate following-up activities for each user according to assessment results. The shift from generic assessment services and/or specific assessment services to learning

management systems is transparent for users when they shift from conducting assessment tasks to performing learning activities.

Although our approach has the advantages described above, The required level of technical knowledge of LD and QTI for authoring new forms of assessments is significant at the moment, because of the lack of easy to use graphical tools that support average practitioners in complex learning models.

6 Conclusions and future work

In this paper, we identify the characteristics of new forms of assessment from the perspectives of process support: involvement of multi-roles and multi-users, variety in task types, complicated control-flow and complicated data-flow. Four corresponding technical requirements were derived for supporting new forms of e-assessment. Through an analysis on the strength and weakness of LD and QTI, we found that a combination of LD, QTI, and a service-oriented approach can meet all identified requirements. We have tested this technical approach through modelling a dozen of innovative assessment examples. The successful executions of these examples in standard-compatible run-time environments demonstrate the feasibility of our technical approach. In comparison to existing software tools supporting new forms of assessment, our approach has advantages in supporting interoperability, flexibility, and a seamless integration with learning activities.

In the paper, we identified that it is too difficult for average practitioners to model new forms of e-assessment using existing tools. Our current work is directed towards the development of a high level assessment process modelling language for specifying new forms of e-assessment. A corresponding assessment authoring tool will be developed to enable average practitioners to model and customize their own assessment processes intuitively. The resulting assessment process model will be automatically transformed into corresponding LD and QTI documents, wrapped as a unit of assessment. Subsequently, these units of assessment can be executed in any integrated LD and QTI platform.

References

Almond, R. G., L. Steinberg, et al. (2001). A sample assessment using the four process framework. *CSE Report 543. Center for study of evaluation.* Los Angeles, University of California.

APIS. (2004). "Assessment Provision through Interoperable Segments." Retrieved Oct. 20, 2006, from http://www.jisc.ac.uk/whatwedo/projects/apis.aspx

Bartram, D. (2005). Testing on the Internet: Issues, Challenges and Opportunities in the Field of Occupational Assessment. D. Bartram and R. Hambleton (Eds) *Computer-Based Testing and the Internet: Issues and Advances.* Wiley.

Boud, D., R. Cohen, et al. (1999). "Peer Learning and Assessment." *Assessment and Evaluation in Higher Education* 24(4): 413-426.

Davies, H. and J. Archer (2005). "Multi-source feedback: development and practical aspects." *Clinical Teacher* 2(2): 77-81.

Freeman, M. and J. McKenzie (2002). "Implementing and evaluating SPARK, a confidential web-based template for self and peer assessment of student teamwork: benefits of evaluating across different subjects." *British Journal of Educational Technology* 33(5): 553-572.

Gipps, C. (1999). "Socio-cultural perspective on assessment." *Review of Research in Education* 24: 355-392.
Harris, D. and C. Bell (1990). *Evaluating and Assessing for Learning*. London, Kogan Page.
LD. (2003). "IMS Learning Design Specification." Retrieved Oct. 20, 2006, from http://www.imsglobal.org/learningdesign/index.html
QTII (2003). "IMS Question and Test Interoperability Specification." Retrieved Oct. 20, 2006, from http://www.imsglobal.org/question/index.html.
Lepsinger, R. and A. D. Lucia (1997). *The Art and Science of 360 Degree Feedback*.
Lockyer, J. (2003). "Multisource feedback in the assessment of physician competencies." *J Contin Educ Health Prof* 23(1): 4-12.
Koper, R. and B. Olivier (2004). "Representing the Learning Design of Units of Learning." *Journal of Educational Technology & Society* 7(3): 97-111.
Orsmond, P. (2004). Self- and Peer-Assessment: Guidance on Practice in the Biosciences. *Teaching Bioscience Enhancing Learning*. e. a. Maw.
Pellegrino, J. W., N. Chudowsky, et al. (2001). *Knowing what students know: the science and design of educational assessment*. Washington, DC, National Academy Press.
Rosbottom, J. (1994). "Software Tools for Peer Assessment." *CTI Annual Conference*, from http://www.ulster.ac.uk/cticomp/stpa.html.
Sluijsmans, D. M. A., S. Brand-Gruwel, et al. (2004). "Training teachers in peer-assessment skills: effects on performance and perceptions." *Innovations in Education & Teaching International* 41(1): 60 – 79.
TENCompetence. (2006). "TENCompetence Project Website." Retrieved November, 2006, from http://www.partners.tencompetence.org/
Topping, K. J. (1998). "Peer assessment between students in colleges and universities." *Review of Educational Research* 68: 249-276.
Topping, K. J., E. F. Smith, et al. (2000). "Formative peer assessment of academic writing between postgraduate students." *Assessment and Evaluation in Higher Education* 25(2): 150-169.
Vogten, H., H. Martens, et al. (2006). CopperCore Service Integration - Integrating IMS Learning Design and IMS Question and Test Interoperability. *The 6th IEEE International Conference on Advanced Learning Technologies*, Kerkrade, The Netherlands, IEEE Computer Society.

Acknowledgement

The work on this chapter has been sponsored by the TENCompetence Integrated Project (TENCompetence 2006) that is funded by the European Commission's 6th Framework Programme, priority IST/Technology Enhanced Learning. Contract 027087.