Towards a Standardized E-Assessment System: Motivations, Challenges and First Findings

Mohammad AL-Smadi, Christian Guetl and Denis Helic
Graz University of Technology, Graz, Austria

doi:10.3991/ijet.v4s2.919

Abstract—“Global Learning” with shared learning contents, resources, activities and goals is one of the contributions of Globalization. With the capability to use new Information and Communication Technologies (ICT) it is a bit easier to have a technology based learning systems that enable learners to share the learning resources and possibilities. As a result many Learning Management Systems (LMS) were developed with divers of platforms and approaches. Consequently, sharing learning resources and components has become a major challenge. E-assessment as a primary activity of any LMS is facing the same challenges and problems. In order to stand on this challenge people in the field of technology enhanced learning have recommended that LMS should conform to specific standards. This paper discusses this challenge, the consequences and limitations of standards in the modern learning settings. Moreover, it shows a service oriented framework for assessment which aims to make the e-assessment systems flexible and also to initiate the term of “Global Learning Assessment” with the possibility of sharing the e-assessment system components.

This paper is an extended version of the IMCL2009 paper.

Index Terms—E-assessment, Standards, Standardized e-Assessment System, Service-Oriented Framework for Assessment (SOFA).

I. INTRODUCTION

Members of our society are affected by rapid changes in every part of our modern life. Terms such as “post-industrial society”, “information society” and “knowledge society” have been used to identify and understand the extent of these changes. Knowledge has become a primary resource for production instead of capital and labor. As a result the knowledge society creates shares and uses knowledge to improve and to have a well-being of its people. Another term of “global society” with a shared knowledge is one of the aims of globalization and using new Information and Communication Technologies (ICT). Therefore, “global learning” is needed as a primary mean of delivering this shared knowledge to the society people. As a result many open-source or even commercial Learning Management Systems (LMS) were developed. The variety of the platforms and approaches used in these LMSs makes it difficult to exchange information between them. Therefore, some of them have become obsolete and dedicated for specific institutions [1]. Moreover, the e-learning content has not been carefully designed. According to [2], e-learning content should not be electronic replications of the classroom materials, rather than it should be value-added. Careful design of the learning content with integrated presentations, exercises, and valuable evaluation and feedback, as well as flexible content navigation and usable user interfaces may motivate learners for further success [2]. Learning content reusability and interoperability, learner’s information accessibility and share ability, are main matters of quality for any LMS. For instance, suppose that you are a lecturer in a university where they have a LMS and you moved to work to other university using different LMS, if the first LMS of your university is not well designed to share your learning materials you will not be able to use the same materials in the new LMS. Another example, if you are using LMS and a stand-alone e-assessment system and both of them are not interoperable then you have to register the student twice in both of the systems as well as you will have duplicate data and you are not capable to use the assessment results in the process of learning content adaptation and redesign. For avoiding such matters and problems, LMS should be designed and implemented to be standard-conform. E-assessment as an important part of any e-learning system also faces the same challenge and problem. Different standards have been used to design and develop e-assessment systems components. The multiplicity of such standards increases the difficulty of making those systems sharable and interoperable.

In order to have flexible e-learning and e-assessment systems, a set of features and requirements have been identified. One of these requirements is standards conformation. According to [3], standards foster e-learning and e-assessment systems to insure seven abilities:

- Interoperability.
- Reusability.
- Manageability.
- Accessibility.
- Durability.
- Scalability.
- Affordability.

Interoperability is defined by [4], as the ability of different systems to share information and services in a common file format. Reusability refers to the ability of using the learning content by different tools and platforms. Manageability is how much the system is able to keep track on the learning experience and activities, rather than the ability of tracking how learning objects are created, stored, assembled and delivered to users. Accessibility is the ability of customize, access and deliver learning contents from anywhere and anytime. Durability means that the learning content does not need any redesign or redevelopment even with new versions of the system. Scalability refers to the ability of the system to grow from small to large. Affordability Is the system affordable?
This paper discusses three main research questions. What is a standardized e-assessment system?, why e-assessment systems must be standard-conformant?, and where we are in our research towards a flexible e-assessment system with regards to standards?. To this end, the rest of this paper is organized as follows: Standards organizations and types are discussed in section 2. Section 3 shows a set of applications scenarios for e-assessment systems. In section 4, we have discussed how to make an e-assessment system standard-conform. The problems and challenges of designing a standardized e-assessment system are identified in section 5. Section 6 stresses on the importance of having a service oriented architecture for e-assessment systems in order to be flexible and standard-conform. Conclusions and outlook is the content of section 7.

II. STANDARDS IN A MODERN LEARNING SETTINGS

The learning process has changed from being repetitive to a new form of learning based on learner self-regulation and skills improvement [5]. The learning theories have also changed from being associative and behavioral to be more cognitive and constructive, where the measurement have evolved from being scientific measurement (separated from the instruction activity) to have a new culture of assessment (where measurement and instruction have been integrated) [6]. Fig. 1, shows the learning process cycle in a LMS. Clear learning objectives and goals are identified at the beginning of any learning process. Based on these objectives, the learning content is carefully selected and designed (Curriculum design). The content is then delivered using appropriate media (Curriculum delivery). Based on adequate assessment activities, the learning content is evaluated against the pre-defined objectives from the first step as well as an adaptation process may occur to enhance the content and insure the achievement of these objectives (Curriculum evaluation). Standards and specifications are needed in every step of this learning cycle. Designing the learning content to be standard-conform makes it reusable and interoperable, as well as durable with easier adaptations and enhancements. The process of publishing educational standards starts by defining a set of specifications that describes some e-learning topics such as learning objects metadata, Learner/educator information, content sequencing and services delivery. This step is done by many organizations and consortia like Dublin Core (DC) [7], The Instructional Management System Global Learning Consortium (IMS GLC) [8], The Aviation Industry CBT (Computer Based Training) Committee (AICC) [9], The Alliance of Remote Instructional Authoring and Distribution Networks for Europe (ARIADNE) [10] and the EU-funded PROMoting Multimedia access to Education and Training in Europe (PROMETEUS) [11]. Specifications are then tested by organizations such as Advanced Distributed Learning (ADL) [12] to be tested specifications such as, ADL Sharable Courseware Object Reference Model (SCORM) [13]. The tested specifications are forwarded then to a standard committee as IEEE Learning Technology Standardization Committee (IEEE LTSC) [14]. At the end standards are approved by official standards organizations as ISO and ANSI to be official standards. Standards vary according to their approval and use. There are four types of standards based on their approval [15]:

- **Official Standard**: a set of definitions, requirements, formats and design guidelines for e-learning systems or their components that a recognized standards organization has documented or approved. e.g. IEEE LTSC (Learning Technology Standardization Committee), ISO/IEC JTC1 (Joint Technical Committee)[16].
- **de facto Standard**: the same as the official one, but accepted only by the community and industry.
- **Specification**: the same issues as the official standards, but less evolved; usually developed and promoted by organizations or consortia of partners from academia, industry and educational institutions. e.g. IMS Global Learning Consortium, PAPI Learner (Public and Private Information)[17].
- **Reference Model**: an adapted and reduced version of a combination of standards and specifications focusing on architectural aspects of an e-learning system, definitions of parts of the system and their interactions. e.g. LTSA (Learning Technology Systems Architecture)[18], SCORM (Sharable Courseware Object Reference Model).

In the e-learning domain, standards can be classified according to their applications into the following [19]:

- **Metadata Standards**: a set of standards used to describe Learning objects’ (LO) attributes, Such as the authors, title and languages. This description can be published with the LOs to facilitate their search and retrieval. such as, IEEE Learning Object Metadata (LOM) [20], IMS Meta-data [21].
- **Packaging Standards**: describes the assembly of LOs and other complex learning units (e.g. online courses) from various texts, media files and other resources. Such assembly can be stored in a Learning Object Repository (LOR) and imported in a Learning Management Systems (LMS), such as, IMS Content Packaging and IMS Learning Design [22].
- **Learner Information Standards**: Formulates the description of the learner information and used to exchange this information between several systems, rather than its use in users modeling and personalization such as, IMS LIP (Learner Information Package) [23] and PAPI Learner (Public and Private Information).
• **Question and Test Standards**: Special types of standards which are used in the assessment systems to represent questions and tests. IMS QTI (Question and test Interoperability) [24] is an example of such standards.

• **Communication Standards**: specify the users’ access to the LMS content, assessments, collaboration tasks and services communication, such as , IEEE LTSA (Learning Technology Systems Architecture).

• **Quality Standards**: specify the pedagogical, technical, design and accessibility perspectives for the LOs’ quality.

• **Semantic Standards**: specify how we can organize content and refer to it in the semantic web.

III. APPLICATION SCENARIOS FOR E-ASSESSMENT

In order to identify the main requirements of e-assessment systems, as well as to figure the limitations of the available standards we will outline a set of application scenarios for e-assessment in modern learning settings.

WebSys is a software company that requires any job applicant to have a specific certificate related to their system. They are looking for a tool that can be engaged to their system with the ability to prepare tests to evaluate the new applicants. In order to handle this need and prepare factual knowledge questions based on the selected content, the e-assessment tool must have a modular design that facilitates the process of integration with the current system. Also the tool should support a flexible number of standards to facilitate the engagement process especially the ones used in the current system.

Ali is a lecturer in a university who teaches Management Information Systems for the students of the second year in the college of Management and Administration. His didactic objectives include the level of understanding of the factual knowledge by his students through a continuous assessment. To do that, he decided to use an e-assessment tool to deliver the tests and to analyze the results through a set of continuous feedback during the course. The e-assessment tool should have flexible and user-friendly interfaces to help him to generate his tests and deliver them to his students. As well as helping him to automatically generate the tests based on the selected contents and to assess the results. Furthermore, the tool should be designed to analyze the answers of the students and provide a feedback which makes it useful for him to conduct continues assessment during his courses.

Sara is a lecturer and teaches Algebra to undergraduate students. One of her didactic objectives is to use computers to assess and assist students during her courses. She believes that when her students practice Algebra on computers and do more and more on-the-fly generated exercises they can easily pass the course. In this situation, the e-assessment system should provide her with flexible and easy to use interfaces to design here algebraic questions and save them in a database. Then, the tool itself can generate a set of exercises to the students and assess their answers based on the answers had been prepared by Sara before, or based on the algebraic engine that the tool should have. Moreover, the tool must provide a feedback to the student about her/his metacognitive knowledge and an appropriate feedback about the progress of the same student during the same course.

Jake is a teacher in a high school and he is interested in applying a set of online rubrics to assess the students’ results according to a specific criteria. Regards online rubrics, the e-assessment tool should be flexible to help him to design a set of rubrics to automatically grade the students’ results based on these rubrics.

IV. STANDARDIZED E-ASSESSMENT SYSTEM

A standard-conformant e-assessment system is the system that their components are designed and implemented according to specific standards. As depicted in Fig. 2, the conceptual e-assessment system has three main parts. The first one is the core e-assessment system which should be flexible to work as a stand-alone system or to be part of any other system. The second part is the interface which is responsible for the external communication between the core system and the other external ones. The interface should support as many as possible of different standards which keeps the core system flexible and interoperable. The last part is the system users and the systems which could be e-learning systems or e-assessment systems.

In order to have a flexible e-assessment system, two levels of standardization have been defined. The first level is the Internal one, where the core e-assessment system’ components should be conformant to specific types of standards and/or specifications. The External level is related to the ability of this system to interact and exchange components with other systems. This level of standardization takes place in the interface where, modules to support different types of standards and/or specifications are implemented. The availability of these two levels guarantees the overall e-assessment system to be standard-conform. Consequently, it will be flexible and interoperable.

A. The Internal Level

Based on the application scenarios in the last section, e-assessment systems should have the following features: (a) **flexible design** to be used as a stand-alone service or to be easily integrated in existing systems. (b) **User-friendly interfaces** for both students and educators where a user interaction and online submission of solution and evaluation can be done. (c) Assessment environment for **various learning and assessment settings** which supports guided as well as self-directed learning. (d) **Management and (semi)automatic support** over the entire assessment lifecycle (exercises creation, storage and compilation for assessments, as well as assessment performance, grading and feedback provision). (e) **Rubrics design and implementa-**
tion interfaces to allow the educators to design their own rubrics based on learning objectives to assess learners’ performance against a set of criteria. (f) Support of various educational objectives and subjects by using various tools sets which for example enables automatic exercise generation or selection, automatic grading and feedback provision. (g) Results analysis and feedback provision (immediately or timely) of the current state of user knowledge and metacognitive skills for both educators and learners and also for adapting course activities and learning contents based on users’ models. (h) Standard-conform information and services to be easily sharable, reusable and exchangeable. This will include the tests’ questions, answers and students’ results, rather than any other required services. And finally, (i) Security and privacy where a secure logon of users based on pre-defined levels of access, and also users’ authentication based on machine (domain users) or by usernames/passwords. For further information and for the conceptual architecture of this system you can refer to [25].

Fig. 3 shows some of e-assessment system components and the possible standard(s) and/or specifications that can be used for representing them. The Test Preparation Unit is responsible for the purposes of tests Authoring and Delivery. A specification such as IMS QTI is used by this unit during the test authoring. In cases of having learning objects related to the test we may use the IEEE LOM standards. The tests can be analyzed by the use of Test Analysis Unit which is based on the same type of specifications to provide a feedback (timely or immediate) to the system users (individuals or organizations). The system users are managed by the User Unit which is a standard-conform to provide some services as user personalization and modeling. Standards such as PAPI Learner or IMS LIP can be used.

B. The External Level

The external level is represented by the functionalities of the interface that is underpinned by a set of available standards and/or specifications. The interface is responsible for the external communication between the e-assessment system (the internal level) and the other related systems. Via this interface information such as questions/exercises and answers, users’ information, list of enrolled students, courses information and learning objectives can be shared with other systems and tools. The more standards this interface supports the much more flexible our e-assessment system will be. As shown in Fig. 4, different examples of possible standards and specifications that the interface should be flexible to support.

V. PROBLEMS AND CHALLENGES

This section discusses the problems and challenges for designing a standardized e-assessment system based on the previously discussed conceptual e-assessment system and the application scenarios in section 3.

Based on the scenarios discussed in section 3 we will show some recommendations and limitations on the available standards. In the scenario of the Company, the e-assessment system should be flexible to work as a standalone system or to be engaged with other systems such as the case in this scenario.

To make this applicable the e-assessment system must have a modular design so that some modules can be integrated with other systems such as the system in this company. In this situation the problem of standards appears where the e-assessment system must support the standards used in the other system. Therefore, we recommend that the e-assessment system should support as much as possible of the available standards.

The second scenario is a traditional one where the e-assessment system is applied as a standalone system to deliver and assess the students’ tests and provides feedback. The limitation of standards appears again in the third scenario where a mathematical representation of the question (symbolic representation) is needed. For example, when the student is going to solve an equation we need some symbolic representation for the solution. Furthermore, a standard such as IMS QTI do not have the ability to represent the solution as a set of symbolic representation of the equations using XML. Therefore, no reference answer is available to automatically assess the student result and provide him a valuable feedback. One of the other limitations of IMS QTI specification is rubrics representation. The problem appears in the fourth scenario where online rubrics are needed to assess the students answers based on a specific criteria.

One of the most important problems and challenges of designing a standard-conform system is the so-called impedance mismatch between the features offered by the standard and the ones needed in a particular application...
for interfaces that facilitates the creation and retrieval of and interoperable. It also provides a common definition for interfaces that facilitates the creation and retrieval of tests and results [27]. Even though the IMS QTI has these features it still has some difficulties in the application domain (such as, foreign languages teaching). One of these difficulties is that the IMS QTI is designed to formulate general types of questions and does not take into consideration some specific questions and test types for a particular domain [28]. Crossword puzzles which are used in the domain of foreign language teaching are an example of those not supported question types by the QTI [26].

According to [29] the QTI standard are not related to didactical issues and tries to be didactically neutral as possible. Another example is what authors of [30] have noted about the IEEE LOM (Learning Object metadata). They noted that IEEE LOM from a perspective of metadata don’t provide enough information to support the learning process. According to [19] some developers find parts of IEEE LOM too restrictive or imprecise. And they also argue that the amount of metadata is not enough to facilitate the search and retrieval of the LOs.

Another major challenge is the problem of selecting the most appropriate standard in cases of having different types of standards for the same aspect of the Learning Management System (LMS) [19]. For example IEEE PAPI Learner and IMS Learner Information Package (LIP) both of them are related to the issue of learner modeling. Even though they look similar but there are a lot of differences in the way how they model the learner. Therefore, the developer should have a good understanding of the current available standards and the main requirements that helps him to choose the most appropriate standard.

VI. SERVICE-ORIENTED FRAMEWORK FOR ASSESSMENT (SOFA)

A framework represents a rich vocabulary that is used to support people in the domain as well as software developers to overcome the problems encountered through the software development. A framework is also used to create a shared language that will describe the problems and their solutions. In the domain of e-learning and e-assessment, frameworks are not used to build generic LMS/e-assessment systems. Rather than, they encourage “coherent diversity” where common service definitions are provided and used to achieve the diverse goals of the organizations. Therefore, the organizations unique infrastructures will stay coherent and consistent with respect to each other. A framework also supports organizations to develop service-oriented architectures by identifying the main services that these organizations may need to develop their applications. The main aim of a framework is to have the ability to identify the services as well as to assign one or more open standards and specifications for each service. [27].

According to [27], a service-oriented framework consists of a set of services, each of which represents a pattern to solve a particular problem. Services can be represented in different ways, as a web service, as an API, as well as a manual business process. The process of defining the framework services is called “factoring services”, where an ongoing process is done by the organizations to identify a service to solve a specific problem and then group the related services (e.g. domain related services, problem related,…etc) in a framework for further enhancement and development. The more fine-grained the services the more flexible and supportive the framework. There is no current “best practice” for factoring services, but the fine-grained they are the more useful they will be. Moreover, there is no concern about how the service is implemented as the concern of the service functionality and behavior towards the other services. A service-oriented framework may support e-assessment systems to easily share and exchange test between each others. Services for tests, items, results, users information…etc, can be easily implemented in the system and they are flexible to be used by other authorized services or systems. For example, students that are registered for a specific test can only attend the e-learning course in other system and vice-versa.

Service-oriented architectures allow the development of modular and flexible systems [28], where the components of the system are flexible to be added, replaced or removed. As well as, new systems can be composed from a collection of suitable services. Based on what we have discussed earlier and a step towards the flexible e-assessment system design and implementation, a Service-Oriented Framework for Assessment (SOFA) has been identified. Fig. 5 shows the SOFA layers and services.

SOFA has five abstraction layers as follows:

- **Users and Systems** represent the external possible users, tools, and systems that may interact with the e-assessment system. Such as, assessment systems and LMS as well as any other authoring tools.
- **Interface** as discussed earlier, the interface is used for the external communications between the e-assessment system and the other external systems, users, and tools. The interface layer should be underpinned with a set of specifications and standards in order to facilitate the integration and communication between the core e-assessment system and the external user agents.
- **Assessment Services** represent the fundamental services for any e-assessment system. The services in this layer are used to perform the main functionality of the assessment process from authoring the items until exchanging them. Special interfaces are used to make the interaction between these services and the assessment portal and users. For which specifications and standards of the internal level of the e-assessment system are used.
Towards a Standardized E-Assessment System: Motivations, Challenges and First Findings

- **Common Services** a lower level of services those are not assessment-specific such as authorization and authentication.
- **Infrastructure** represents the internal networks, storage and processing capabilities that the e-assessment system requires.

SOFA has two layers of services in the internal level: assessment services and common services. The assessment services in SOFA have been identified based on FREMA (Framework REference Model for Assessment) processes concept map [32]. All of the services in this group are assessment services and work together in order to support the assessment process. The group of Common Services is a set of services that may be found in any assessment system or any other system such as e-learning systems. The services are organized in a set of layers based on the IMS GLC Abstract Framework (IAF) [33], which consists of four main layers, the “Application Layer”, the “Application Layer Services”, the “Common services” and the “Infrastructure Layer”. The assessment services are described as follows:

- **Author** this service is responsible for creating the tests, items, templates and items’ pools. This service should be standard-conform in order to have an interoperable components. An example of possible standards and specifications is IMS QTI.
- **Manage users** this service maintains the secure logins and passwords to the e-assessment system. This process is handled by identifying the possible groups of the users as well as the rules for them. Possible specifications and standards can be IMS LIP or PAPI.
- **Schedule** sets up the test where the required hardware and software systems are identified as well as the candidate users of the assessment. This service is usually used by the timetabler.
- **Pre-delivery** handles the necessary operations after authoring and scheduling of a test and before delivering this test.
- **Deliver** presents the items to the candidate. According to FREMA definitions, this service involves the following processes: (1) deciding next item to be presented (2) retrieving the item file and its related resource files (3) displaying the item, which may require specialized software (4) reading in (and perhaps validating) the candidate’s response (and perhaps confidence level).
- **Mark** assigns a mark or a score to the candidate response.
- **Grade** assigns the final mark to a grad, for example the mark 90 is assigned to a grade “A”.
- **Analyze** The candidate responses as well as their marks and grades are analyzed for providing feedback as well as for further moderation and enhancements.
- **Feedback** displays useful information for the candidates during/after the assessment process.
- **Certify** this service is responsible for the different processes of candidate performance recording ranging from the paper certificates to the electronic certificates in e-portfolios.
- **Moderate** checks the assessment process is satisfactory or not.
- **Process data** produces useful information after the test is finished. This information is usually used by the stakeholders and decision makers.
- **Process appeals** allow the candidates to appeal against their final grades.
- **Quality assure** assures the quality of the assessment processes ranging from the fairness of an item to the satisfactoriness of the assessment processes to the institution goals and objectives.
- **Exchange** responsible for the exchange and purchase of the tests and items.

VII. Conclusions and Outlook

This paper has discussed three main research questions: “What is a standardized e-assessment system?”, “Why e-assessment systems must be standard-conform?”, and “where we are in our research towards a flexible e-assessment system with regards to standards?”. A standard-conform (standardized) e-assessment system is the system that their components are designed and implemented according to specific standards. In order to be more clearly about this question, we have distinguished between two levels of standardization in the e-assessment system. The internal level and the external one and we have shown in some detail how and where standards can be used in both levels. Standards-conformance is the way of how to ensure that our e-assessment system will be flexible, interoperable, reusable, manageable, accessible, durable, scalable, and affordable. Several organizations and consortia are working on e-learning and e-assessment standards in particular. The diversity of these standards has made some challenges and problems to the people in the field of designing and implementing e-assessment systems. The lack between the features offered by the standard and the ones needed in a particular application domain is one of these challenges. Another major challenge is the problem of selecting the most appropriate standard in cases of having different types of standards for the same aspect of the LMS e.g. IEEE PAPI Learner and IMS Learner Information Package.

A Service-Oriented Framework for Assessment (SOFA) has been identified in order to have a flexible standard-conform e-assessment system. SOFA consists of five layers, users and systems layer, an interface, assessment services, common services, and infrastructure. The services should be designed and implemented to be standard-conform. Once the system is based on fine-grained standard-conform services it will be more flexible and worthy.

In the near future, an in-deep requirements analysis and evaluation of the available standards will be conducted. The evaluated standards will be used in the services design and implementation of SOFA. A cross-domain reference model will be indentified. A first phase of a prototype implementation will be started soon after that.

References


II. MOTIVATIONS AND CHALLENGES

A. Motivations


