Using Wiki for Rapid Authoring of Huge Numbers of E-Assessments

Ilire Hasani-Mavriqi
Institute for Information Systems and Computer Media
University of Technology Graz, Austria
ihasani@iicm.edu

Christoph Portsch
Institute for Information Systems and Computer Media
University of Technology Graz, Austria
portsch@student.tugraz.at

Christoph Trattner
Institute for Information Systems and Computer Media
University of Technology Graz, Austria
ctrattner@iicm.edu

Brigita Kacjan
Institute for German studies, Faculty of Arts
University of Maribor, Slovenia
brigita.kacjan@uni-mb.si

Denis Helic
Institute for Knowledge Management
University of Technology Graz, Austria
dhelic@tugraz.at

Hermann Maurer
Institute for Information Systems and Computer Media
University of Technology Graz, Austria
hmaurer@iicm.edu

Abstract: Nowadays, Question and Test Interoperability (QTI) is a de facto standard for management of so called E-Assessments. It is our experience from numerous technology-supported education projects that QTI supports interoperability and exchange of assessments but completely neglects the authoring process of these assessments. In particular, the task of authoring of a large number of assessments for teachers without technical background is very time and resources consuming one. Therefore, we developed a novel format that focuses on supporting teachers in rapid authoring of huge number of E-Assessments and implemented it as a part of an educational Wiki system. In this paper we present this new format together with its implementation and the first evaluation results of the implemented tool.

1. Introduction

The upswing of Web-based educational tools in recent years has been in particular visible, among others, in the field of didactical methods such as assessments. The tremendous growth of new tools, formats, and content in this area resulted in a completely new research field so-called E-Assessment. Nowadays, it is a common understanding in the research community that E-Assessments include apart from traditional multiple choice questions or free text exercises also technologically advanced tools and formats such as multimedia, interactive simulations, or educational games. According to Hambleton (Hambleton 2004) the biggest challenge in the years to come will be to properly manage the complete life-cycle of E-Assessments starting with creation phase and then moving on to

adoption, evaluation, and maintenance phases. In this respect it is of primary importance to properly handle interoperability, reusability, and exchange of E-Assessments by means of technology.

The de facto standard for E-Assessments nowadays is "Question and Test Interoperability" (QTI) published by IMS Global Learning Consortium (IMS QTI 2010). The goals of the standard are to provide a common format for different tools and systems to guarantee the interoperability between systems and to support exchange of E-Assessments between different tools. QTI describes an E-Assessment model for representing questions (assessment items), tests (assessments), and question and test results. The major promise of QTI is that by introduction of common format developers can concentrate on developing innovative tools, whereas teachers can focus on defining new and groundbreaking methods of how to apply those tools in an online environment. No one would need to think about the technological infrastructure anymore – QTI provides such an infrastructure. Moreover, another useful property of QTI is its general design, i.e. QTI is not tailored to a single field but it is general enough to cover multiple educational fields such as mathematics, physics, or language teaching.

Although, QTI holds the most of its promises in respect to interoperability, exchangeability, or assessment management it is our experience from previous E-Assessment projects that the standard exhibits serious problems in respect to creation or authoring of such assessments. Some of these problems include but are not limited to:

- Authoring tools that implement the QTI standard are too complex and not usable enough to be operated by a teacher without sufficient knowledge of technology. For example, the majority of tools reflect only the standard and therefore the teachers need to deal with technical terms which they usually do not understand. However, this is not the problem of implementation of the tools but one of the standard itself – the standard is technically too complex.

- From the technical point of view the standard is purely elaborated - the standard mixes data management with data presentation features. For example, in order to define a "hotspot" question the assessment item needs to define also space coordinates where the question will appear on the user screen. This is clearly a mixture of data management and presentation and leads to maintenance problems but also to serious authoring problems. An author needs to know the exact position of the question on the user screen although the question can be displayed on a wide range of screen resolutions and even on a wide range of completely different devices.

- Although the generic approach of the standard has a number of advantages in respect to interoperability it also has, at least in certain educational areas, a number of disadvantages. In the area of language learning the majority of assessments are heavily text-based. In certain cases these text-oriented assessments are combined with other assessment types to reflect a particular didactical method. For example, a common assessment in language learning is to recognize a word or a word phrase that is nor correctly written. This can be easily solved with a "hotspot" where students need to click on the part of the text that is written incorrectly. However, in language learning students are often asked to make corrections in such incorrect places. Thus, after students click on an incorrect part they need to obtain a text field where they type the correct answer. Such a combination of assessments is not possible with QTI but for language teachers it is an indispensable tool in their teaching practices.

In this paper we present a novel E-Assessment format developed specifically to support authoring process. The format is focused on development of E-Assessments in language teaching field but can be used in other educational fields as well. Moreover, the format can be easily extended to support specific requirements of another educational field. Last but not least, the format is not developed as a replacement of the QTI standard but only as a format to support easy creation of potentially huge number of assessments by a non technical personal. All general assessments can be mapped to QTI format. The specific assessments developed for a particular educational field can be mapped to a general QTI assessment item but without extra features present in that specific assessment.

The rest of the paper is organized as follows. In the 2nd Section we introduce shortly the language teaching project for which the format has been developed. In the 3rd Section we present the overall design of the format, whereas in the 4th Section a concrete implementation of the format within a Wiki system is presented. The 5th Section provides an overview of the results of the first system evaluation. Finally, we conclude the paper with some directions for the future work.

2. Sprichwort Project: Didactical and Technical Requirements

The Sprichwort project deals with teaching proverbial in foreign languages. The project proverbial database contains around 1500 proverbial in 5 different European languages: German, Hungarian, Check, Slovak, and Slovenian. Apart from the proverbial database the project's own Wiki site will contain at the end of the project

around 1000 assessment items combined into more than a hundred assessments. The basic goal of this didactical unit is self-assessment and self-checking of knowledge of foreign language students. Another overall goal of the project is to involve interested student, teachers, researches, and proverbial experts in working with the system resulting in a sustainable European proverbial community.

At the beginning of the project for the didactical component of the project two sets of the requirements have been identified: didactical and technical requirements.

The didactical requirements concentrate on certain aspects of assessment in autonomous/independent learning in a foreign language classroom: (a) the assessment has to be appealing by means of interactive tasks; (b) it should be independent of traditional time frames as lessons; (c) in lack of a language teacher being present, the assessments have to be easily understandable; (d) the tests as one of the assessment formats include various test types depending on the evaluated subject matter and the underlying exercise typology; (e) a self-evaluation tool is available including the possibility of immediate feedback. In addition to the immediate feedback of one`s own success, the learners` community offers the opportunity to get into contact with other learners and experts in order to discuss certain specific matters.

The technical requirements can be summarized as follows: (a) the authoring component of the system has to be easy-to-use even for users with a non-technical background; (b) the authoring component needs to support authors in rapid creation and modification of large number of assessments; (c) the learning curve for the authoring component must not be steep. Apart from these main requirements another set of technical requirements dealing mainly with interoperability, exchangeability, and reusability of assessments has been identified. In principle, assessments need to be supported by means of widely adopted Web-based standards.

3. Conceptual and Implementation Design of Assessment Component

The basic idea of the assessment component is to abstract the underlying data, functional, and presentational model from the users. To achieve this abstraction a layered abstraction design has been adopted. In such a design an abstraction layer is put on the top of a more basic abstraction layer hiding in this way the basic layer from the layers above.

Conceptual design of the Sprichwort Assessment Component is comprised of two abstraction layers. The first layer defines the assessments, their presentation, and their functionality in terms of client-side, i.e. browser technologies such as HTML, JavaScript, and CSS (see List. 1).

```
<div      dojoType="dispue.HotspotMarking"
controller="c1"
     params="{ 'params': [{
     'content':'Berge',
     'isCorrect':true,
     'hint':'Das passt hier her!'},
     { 'content':'Wolkenkratzer',
     'isCorrect':false,
     'hint':'Das stimmt nicht ganz so!'}
     ]}"
  isRandomized=true>
```

Listing 1: Abstracting Assessments in HTML, JavaScript, and CSS

The second layer abstracts the first one by introducing Wiki syntax to hide the low-level elements from the above layers. A Wiki engine which operated on the data from the second layer is responsible for mapping documents written in the Wiki syntax onto the combination of HTML, JavaScript, and CSS. The second layer also introduces special assessment syntax on top of the Wiki syntax. As such this syntax is a domain-specific syntax tailored to the didactical needs in the form of assessments (see List. 2).

```
Ein alt bekanntes Sprichwort ist:
 [{SpHotspotKennzeichnung
|
 inhalt='Eigener Herd ist goldes wert.'
korrekt=true
|
inhalt='Eigener Ofen ist silber wert.'
korrekt=false
|}]
```

Listing 2: Abstracting Assessments with Wiki Syntax

This component of the second layer is supported by means of a number of Wiki-plugins that transform the assessments into a combination of Wiki syntax and HTML/JavaScript/CSS code. The Wiki syntax is in the next step transformed by the Wiki engine onto final HTML code. The process view of this design is depicted in the (Fig. 1).
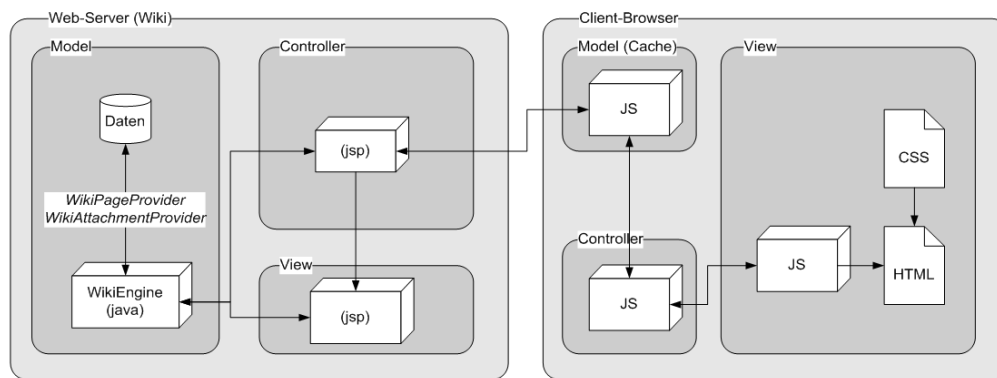
Figure 1: Process View on Design of Assessment Component

4. Implementation of Assessments: Sprichwort-Platform

As already mentioned in the previous sections one of the main goals of the Sprichwort project was to facilitate the creation and authoring process of E-Assessments in the field of language teaching. This goal is accomplished by implementing a web-based authoring tool which is easy-to-use even by teachers and instructors with no technical background. The authoring tool provides an editor and a simple syntax for creating and editing assessment items known as exercises and collections of assessment items known as tests. Furthermore, the assessment authoring process is characterized as a collaborative activity with a many-to-many communication between authors. Certainly, the most appropriate solution that satisfies this purpose is a Wiki system which also builds the basis of the Sprichwort-Plattform. As a result of an evaluation process of different Wikis (Portsch 2010), the JSPWiki (JSPWiki 2010) was chosen for the implementation of the web-based authoring tool. JSPWiki presents an open-source Wiki software which is build based on Java, Servlets, and Java Server Pages and is released under the Apache Public License. JSPWiki follows the concept of the Model-View-Controller (MVC) design pattern as depicted in the (Fig. 1). One of the main advantages of the JSPWiki is its extendibility through custom Wiki-plugins, filters, and dynamic styles. This feature enhances the E-assessment authoring process by providing means for authors to include available Wiki-plugins and styles directly in their assessments.

The E-Assessments developed under the scope of our project require a certain level of user interaction which is best implemented by applying JavaScript. Thus, one of the most widely used JavaScript libraries - Dojo Toolkit (Dojo 2010) is used on the client side of the system. Dojo Toolkit enables the AJAX communication, dynamically content changes, and encapsulates the complexity of cross-browser implementations. Moreover, Dojo Toolkit provides a considerable number of widgets and also supports the implementation of custom widgets. The drag and drop functionality is identified as an important feature for different types of E-Assessments in the field of language teaching. This functionality is also fully supported by DOJO Toolkit. The main types of assessment items defined in

the context of language teaching are (Hricko and Howell 2005): hotspot, multiple choice, drag and drop, free text, and memory. The (List. 3) represents a complete example of a hotspot exercise implemented with the DOJO Toolkit.

```
<div dojoType="dispue.AssessmentController" id="c1"></div>
<p><b>Aufgabe: </b>Markieren sie die Fehler in den Sprichw&ouml;rtern. <br>
<hr>
 <div class="content">
 Die Wasserballer schafften mit […] im Besprechungszimmer Sprüche wie diesen  aufhängen: » Wenn es eine
  <div dojoType="dispue.HotspotMarking" controller="c1"
          params="{ 'params': [{
          'content':'n',
          'isCorrect':true,
          'hint':'Hmm. Das passt doch oder?'},
          { 'content':'m',
          'isCorrect':false,
          'hint':'Komisch oder?'}
          ]}"
 isRandomized=true></div>&nbspGlauben gibt, der
 <div dojoType="dispue.HotspotMarking" controller="c1"
          params="{ 'params': [{
          'content':'Berge',
          'isCorrect':true,
          'hint':'Das passt hier her!'},
          { 'content':'Wolkenkratzer',
          'isCorrect':false,
          'hint':'Das stimmt nicht ganz so!'}
          ]}"
 isRandomized=true>
  </div>&nbspversetzen kann, dann ist es der Glaube an die eigene Stärke.
 </div>
```

Listing 3: A complete example of an assessment item implemented with DOJO

Referring to the abstraction levels introduced in the section 3, the example in (List 3.) corresponds to the first abstraction layer. It is understandable that for authors with lack of knowledge of HTML and JavaScript it is not trivial to create and edit such E-Assessments, because it involves HTML and JavaScript configurations.

Thus, a second abstraction layer was introduced which encapsulates the complexity of the DOJO layer and increases the efficiency of the authoring process by providing authors with a simple Wiki syntax.

Depending on the configuration, a Wikipage can contain an exercise, a collection of exercises resulting in tests, or a self-evaluation questionnaire. Wiki-plugins and filters are used to include DOJO widgets in the Wikipages. The following example shows the Wiki syntax of a hotspot marking exercise (List. 4) and its visualization (Fig. 2).

```
1.[{SpHotspotKennzeichnung
   |
   inhalt='Dass Kinder das kritischste Publikum sind'
   hotspot=false | }],  […]
   [{SpHotspotKennzeichnung
   |
   inhalt='Ende gut, alles gut'
   hotspot=true | }]  […]
 2. Dem Staat war Philips Leben 3500 Pfund wert.  […] Der Vater sagt: "Es heißt,
    [{SpHotspotKennzeichnung
    |
    inhalt='die Zeit heilt Wunden'
    hotspot=true | }]  […]
```

Listing 4: Wiki syntax of a hotspot marking exercise

Figure 2: Visualization of the hotspot exercise from (List. 4)

Drag and drop exercises represent an important part of E-Assessments in the SprichWort-Plattform. They serve as basis for different types of assessment items such as matching source text with target pictures or vice versa, matching source text with target text, or forming a proverb from a bunch of text pieces known as a salad exercise. The example in (List. 5) and the (Fig. 3) should illustrate the implementation of a drag and drop exercises.

```
[{SpDndSalat
|
inhalt=('Der Glaube';'versetzt';'Berge.')
|
inhalt=('Die';'Hoffnung';'stirbt';'zuletzt.')
|
inhalt=('Noch';'ist';'nicht';'aller';'Tage';'Abend.')
|
inhalt=('Eine Schwalbe';'macht';'noch';'keinen Sommer.')
|
inhalt=('Wer sucht';'der';'findet.')|}]
```

Listing 5: Wiki syntax of a drag and drop exercise

Figure 3: Visualization of a salad exercise

5. First Evaluation Results

In a preliminary evaluation all three target groups were asked to answer some questions in order to evaluate the achieved goals of the project. The most prominent results are presented in this short survey.

The first target group was the group of the developers, who were asked to comment on the usability of the authoring components. The answers to this basic question sketched and therefore marked the starting point for the further development of the authoring tools, as the authors of the assessments had only little or almost no technical background. So it was necessary to organize a workshop in order to enable them to use the authoring components. After the training the prepared authoring tools were easy to use and made it possible to create all of the required assessment components. From a didactical viewpoint the matter i.e. the project required a typology of theme-based and appropriate tasks and exercises, which was constructed at the beginning of the project (Kacjan et al. 2009) and constituted the didactical base of the assessments and assessment components.

The production of the content-based feedback was a quite bigger challenge than the technical aspect of the feedback component. It has by default no evident limits, therefore the feedback had to be reduced to a manageable extent. The content of the assessment component self-evaluation sheet was designed at a relatively high level of abstraction and according to this fact, the necessary feedback had to be more abstract as well, which can be regarded as a weak point until the problem will be solved.

36 teachers of German in the participating countries formed the second target group and answered questions concerning the usability and intelligibility of the presented assessments. The results showed that approximately 78% of the teachers were sure that the presented kinds of assessments were productive or very productive for learners. The same percentage of teachers stated that the assessments were appealing or very appealing. In both cases the missing 21 percent were due to missing answers and not to the answers less or not at all appealing or productive. 57% of the teachers were satisfied with the technical design of the assessments, while about 14% described the technical design to be not quite OK. This was mainly due to the fact, that some problems appeared when using the browser Internet Explorer and not Mozilla Firefox. But this problem probably will be solved in the future.

The third target group consisted of 161 learners of German as a foreign language spread over the participating countries and beyond. They stated mostly (up to 89%) that the assessments were appealing or very appealing, 2% marked it as not really appealing, while 8% of the questioned learners did not answer this question. The assessments seemed helpful or very helpful to around 80% of the learners, while 7.5% described it as not or not at all helpful. Similar results were obtained concerning the question about the technical design of the assessments, approximately three quarters of the learners called it good or very good, 11% of them described the design as not really good, mostly due to problems with the server or too many similar assessments, regarding their comments at this question.

From the didactical viewpoint, we can conclude, that the identified didactical and technical requirements have been met to the widest possible extent.

6. Conclusion

This paper introduced a novel E-Assessment format for a web-based authoring tool in the field of language teaching. This format should enable teachers and instructors with no technical background to efficiently create and edit a huge number of E-Assessments. In this way they can concentrate on the content of the assessment items and do not need to experience any steep learning curves by using the authoring tool. The implementation of this novel format intends to overcome the drawbacks of the QTI format regarding the authoring of questions and tests. Furthermore, the E-Assessment format can be easily extended to support specific requirements of another educational field.

The E-Assessment authoring tool was implemented as a combination of a Wiki framework and a JavaScript library, which ensure the simplicity and the collaborative components of the system. In order to encapsulate the complexity of the underlying data, functional, and presentational model from the users, a layered abstraction design has been adopted.

The first results of the performed evaluation study showed that the Sprichwort authoring tool is mostly welcomed by users and it satisfies the purposes of the project.

References

Dojo, (2010). *Dojo Toolkit*. http://www.dojotoolkit.org/.

Hambleton R.K. (2004). Theory, Methods, and Practices in Testing for the 21st Century. In: *Psicothema* 16 (4), 696–701.

Hricko M. and Howell S. L. (2005). *Online Assessment and Measurement: Foundations and Challenges.* Hershey, PA, USA: Information Science Publishing.

IMS QTI (2010). *IMS Global Learning Consortium Question & Test Interoperability Specification (QTI)*. http://www.imsglobal.org/question/.

JSPWiki (2010). *JSPWiki*, http://www.jspwiki.org/

Kacjan B., Chovaniaková D., Kozáková V., Kispál T. (2009). *Übungstypologie*. Maribor, http://www.sprichwort-plattform.org/sp/Ergebnisse.

Kacjan B., Chovaniaková D., Kozáková V., Kispál T. (2009) *Mustertests*. Maribor, http://www.sprichwort-plattform.org/sp/Ergebnisse.

Kacjan B., Chovaniaková D., Kozáková V., Kispál T. (2009). *Selbstbewertungsbögen*. Maribor, http://www.sprichwort-plattform.org/sp/Ergebnisse.

Milligan, C. (2003). *Question and Test Interoperability (QTI): Extending the specification for Mathematics and Numerical Disciplines,* Maths CAA Series, The Maths, Stats & OR Network, University of Birmingham, UK, Nov. 2003.

Portsch, C. (2010). *Wiki-based Assessment System for the Creation of Interactive Language Exercises*. Graz University of Technology, Austria.

Smythe, C., & Roberts, P. (2000). An Overview of the IMS Question & Test Interoperability Specification. *Computer Aided Assessment (CAA'2000)*. Leicestershire, UK, 2000.

Roberts T.S. (2006). *Self, Peer, and Group Assessment in E-Learning*. Hershey, PA, USA: Information Science Publishing

Acknowledgments