An Enhanced Approach to Semantic Markup of VLEs Content Based on Schema.org

Abdulaziz Aldaej and Paul Krause

University of Surrey, Guildford, UK
{a.aldaej,P.krause}@surrey.ac.uk

Abstract. In June 2011, the major search engines (Bing, Google, Yahoo! and Yandex) announced the new innovation of schema.org. However, there is a clear need to integrate Virtual Learning Environments (VLEs) with the wider Web and maintain its learning contents freely open in order to support the sharing and reuse of learning resources. This paper focuses on our proposal to schema.org by proposing new vocabularies to describe VLEs content with rich semantic information due the lack in specific vocabularies for VLEs in the current schema. In order to test this proposal, we have prototyped a simple VLE that was proposed in another paper [1], which makes use of the Semantic Content Management System (SCMS) Drupal to provide a more open, social and semantic structured learning environment. This proposal is discussed here in order to elicit a broad review from the community.

Keywords. Learning Management System, Virtual Learning Environment, Drupal, Ontology, RDFa, Semantic Web, schema.org.

1 Introduction

In recent years, we have increasingly seen advanced web technologies such as the semantic web or web of data being used to facilitate reusing and sharing of data on the web [2]. Very recently, there has been an obvious focus on the use of semantic markup especially after the innovation of schema.org, which is a joint effort between the major search engines to have global and official vocabularies that can be understood by the big search engines. The importance and success of schema.org is clearly visible now, as it has been used widely in many applications and research projects in the past few years, as reported in the statistics in [3]. A very recent research is the UKOER project; they add semantic markup to their education resources using schema.org vocabularies in order to enhance the discoverability of education resources [4]. Also, Rosati and Mayernik have used schema.org to markup their webpages to increase the discoverability and connectivity of their resources in the web [5]. However, there is a lack of VLE vocabulary's support in schema.org to describe VLEs or online courses content with rich semantic data due to schema.org being a new innovation and still evolving. A very recent support to describe educational resources with

rich semantic data has been proposed by the Learning Resource Metadata Initiative (LRMI), which was just added officially to schema.org in April 2013 [6].

Of course, VLEs and other online courses will benefit from these learning resource vocabularies that have included to schema.org. However, there is still a lack of learning terms to cover all VLEs structure vocabularies and hierarchies such as courses, sessions and assignments according to our experience during the implementation of our demonstration VLE. Therefore, we will benefit from the recent and existing vocabularies to propose new vocabularies that are unavailable now with the current version based on VLEs' need, as schema.org is still evolving and is willing to receive new proposals from the community. This proposal aims to extend to the previous work that has been included in the schema by LRMI in order to provide an enhanced approach to describe learning contents with rich semantic data in a VLE context. Furthermore, this proposal can be also appropriate for other open online courses such as MOOCs to benefit from this proposal to structure their courses' content with rich semantic information. Consequently, this will allow course instructors and students to have more accurate and meaningful learning contents when they search within the major search engines. This will also increase the visibility of VLEs' content within the big search engines by providing a machine-interpretable semantics to course data.

To our knowledge, all the "conventional" VLEs lack enhanced semantic features and support [7]. Consequently, a SCMS was selected in this paper to deploy and support the development of different parts of e-Learning services for higher education institutes. Currently only Drupal among the most popular CMSs has advance support for semantic features built into the core of its latest version. Very recently, schema.org has started to support RDFa 1.1, in particular RDFa Lite, additionally to its preference for Microdata after the emerging of this version of RDFa as it has the simplicity of Microdata and the high demand from the community. This is very important to us, as we use Drupal 7 to implement our VLE and its core design built on RDF.

This paper is an extension of an earlier paper that proposed our demonstration VLE that makes use of web 2.0 and web 3.0 tools [1]. However, the current paper focuses on our proposal to LRMI and schema.org in order to elicit a review by the community before submitting it officially. We will now discuss our proposal specification below.

2 Our proposal to Schema.org

This proposal suggests new vocabularies including six new types and 12 new properties to be embedded with schema.org to structure VLEs content with rich semantic information. Fig. 1 shows the proposed classes and properties, how they related to the existing schema.org hierarchy and the recent support for educational resources by LRMI. This proposal will rely on the current expressive types CreativeWork and Action, as they have most of the properties to describe VLEs. It is suggested by the schema.org extension mechanism to use existing and relevant vocabularies in the schema in order to create new classes or properties. There are many extended types to CreativeWork and Action; some of them do not have any new properties. An example is the CreativeWork/Photograph type; it does not have any new property but simply

puts CreativeWork in a photographic context. Therefore, the main reason of this proposal is to add new types and limited properties, which do not exist in the current schema to put CreativeWork and Action in a VLE context. However, search engines will ignore these vocabularies until they become included officially in schema.org [5]. We will use our demonstration VLE in order to implement and test this proposal.

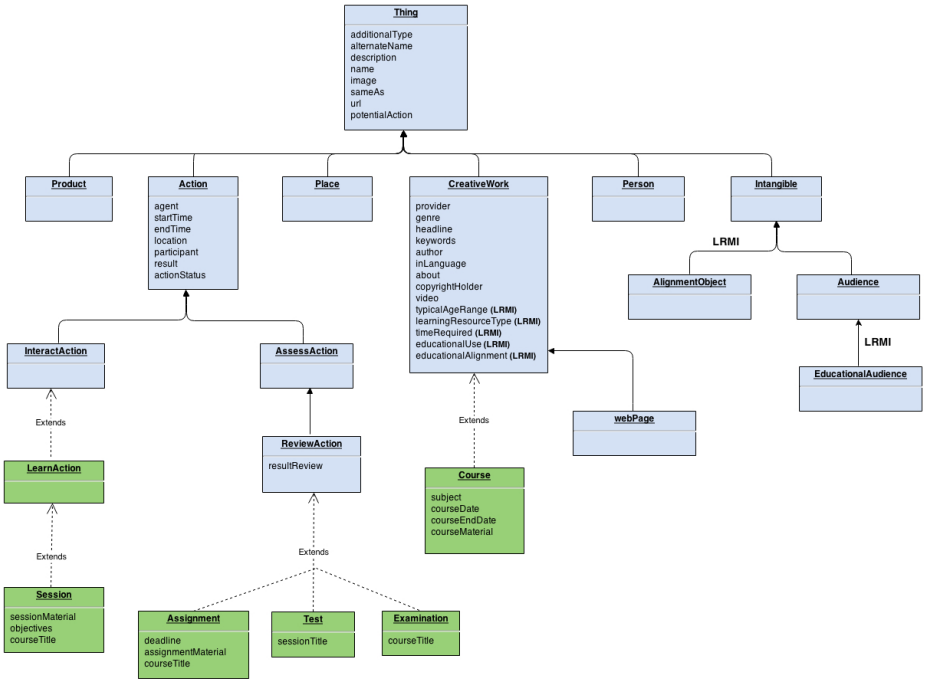

Fig. 1. The proposed classes and properties (green boxes) linked to schema.org hierarchy

In addition, our VLE structure consists of four main classes: Course, Session, Assignment and Forum. However, this proposal does no rely only on our VLE structure, it benefits also from the existing educational vocabularies such as AIISO and TEACH in order to have a comprehensive semantic structure for VLEs and other online courses [8], [9]. Regarding to the Forum class, it has been automatically marked up by default using SIOC and DC ontologies as this VLE was built on the SCMS Drupal 7 and we will not include any further details here as this paper focuses mainly on the new proposal. Furthermore, we have marked up our demonstration VLE homepage semantically using the existing type WebPage that is an extension of CreativeWork. Full details of our proposal for the new types and properties are described below.

2.1 Course

This new type is extended from the existing class CreativeWork to put this type in a Course context. CreativeWork has most properties that are needed to describe Course

fields with rich semantic information. We have used the following existing properties: name, description, image, about, provider, genre, inLanguage, typicalAgeRange, headline and author. Full details about these properties can be found in schema.org/CreativeWork. We propose new properties, which are shown in Table 1 for the fields where there are no relevant vocabularies available in schema.org. Each new property is extended from an existing one, which is a bit relevant to its context, as the schema.org extension mechanism has suggested. For example, subject is extended from about. courseMaterial is extended from additionalType. Also, courseDate and courseEndDate are extended from datePublished. This new type Course and its new property (subject) have already used in AIISO and TEACH. However, we use subject here as property not as class like in AIISO.

Table 1. New Course type properties Specification

Property	Expected Type	Description
subject	Text	The field of the course; e.g., 'Computer Programming'.
courseMaterial	URL	Any resource or file for the course.
courseDate	Date	The start date of the course.
courseEndDate	Date	The end date of the course.

2.2 Session

This new type is extended from the existing class InteractAction that is a subtype of Action to put this type in a Session context. However, we propose a Session as subtype of our proposed type LearnAction in order to describe InteractAction in a LearnAction context as this existing type has several specific types such as RegisterAction and JoinAction. Action has most properties that are needed to describe Session fields. We have used the following existing properties: name, description, image, agent, startTime, endTime, participant and location to describe session fields in rich semantic information. Full details about these properties can be found in schema.org/Action. However, the use of agent property here is different with the use of author property in Course. author is used to describe the course instructor in a Course. On the other hand, agent is used here to describe the lecturer of the session, as there could be different lecturers for each session within the course. We also propose three new properties to the Session type that are not available in the current schema and it would be valuable to add them to schema.org and LRMI specifications. Table 2 shows the three new properties with a short description to each property. sessionMaterial is extended from the existing property additionalType. courseTitle is extended from citation to link the session to its course. Also, objectives is extended from headline. Furthermore, Session type is not available in AIISO and TEACH like Course and Assignment. It has been used as a Lecture class in TEACH. However, we use Session here instead of Lecture, as Session has a much broader meaning and could be used to describe lecture, seminar, lab class and tutorial. We could extend the Session type with these subtypes but we would like to keep this proposal as simple as we can.

Table 2. New Session type properties Specification

Property	Expected Type	Description
sessionMaterial	URL	The session files such as handouts and presentation slides.
courseTitle	Course or Text	A citation or reference to the name of the course.
objectives	Text	The objectives or headlines of the session.

2.3 Assessment

Assessment is a broad expression in educational context. There is existing type ReviewAction, which is an extended type of AssessAction that could be used to describe Assessment. However, there are no available specific types of AssessAction and ReviewAction in order to describe Assessment in an educational context. Therefore, we propose three new specific types of Assessment (Assignment, Examination and Test) that are extended types from the existing type ReviewAction as shown in Fig. 1. Thus, all these three new subtypes can inherit and use the properties from ReviewAction and its root types. Full details of these new types are described below.

Assignment. We propose three new properties to this new type as shown in Table 3. deadline is extended from the existing property endTime. assignmentMaterial is extended also from the additionalType property. Also, courseTitle is extended from citation to link the assignment to its course. These three new properties are really important to be included to the Assignment, as most assignments have these fields. Furthermore, this new type Assignment and its new properties (deadline and courseTitle) are already available in TEACH specification.

Table 3. New Assignment type properties Specification

Property	Expected Type	Description
deadline	Date	The deadline of the assignment.
assignmentMaterial	URL	Any resource or file for the assignment.
courseTitle	Course or Text	A citation or reference to the name of the course.

Examination. This type of assessment can be described as a formal test or summative after a series of sessions, lectures and lessons or after a period of time. We propose one new property to this new type as shown in Fig. 1. courseTitle is extended from the existing property citation in order to link the examination to its course.

Test. Test is checking learning after a specific session, lecture or lab class. We propose one new property to this type as shown in Fig. 1. sessionTitle is extended from citation in order to link the test to its session or lecture. Therefore, this is the key reason of proposing a separate Test type and not includes it with the Examination type,

as test usually conducts after specific session and should be linked to this session. For example, test can be seen obviously in lab classes, as usually there is a test after each lab class especially in programming labs. There is also further type of assessment in educational context “Quiz”, which can be listed under the Test type. Quiz is a shorter version of test on a specific topic and could be described as a Test in this proposal.

3 Conclusion

In conclusion, we have shown in this paper our proposed specification for schema.org and LRMI to structure VLEs content in rich semantic information. We have described our proposal in this paper in order to have it reviewed by the community before submitting it officially. We have implemented and tested this proposal using Drupal and the structured data testing tool which is offered by Google. Full details about the implementation and testing of this proposal will be discussed in a subsequent paper that is in preparation. As we have mentioned, our VLE will benefit from the advantages of semantic web to provide a more open learning environment in a machine-understandable context. Basically, it embeds a semantic markup description within HTML webpages and uses the broad and supported vocabularies “schema.org”. Finally, we try to maintain this work as a contribution to the technology-enhanced learning community’s main focus in the past few years of sharing and reusing educational resources, as described in [10].

References

1. Aldaej, Abdulaziz, Paul Krause, and Abbas Strømmen-Bakhtiar. "E-Learning evolution: Next Steps Semantic web and foundations of E-learning." In World Conference on E-Learning in Corporate, Government, Healthcare, and Higher Education, vol. 2012, no. 1. 2012.
2. Bizer, Christian, Tom Heath, and Tim Berners-Lee. "Linked data-the story so far." International Journal on Semantic Web and Information Systems (IJSWIS) 5, no. 3 (2009): 1-22.
3. Mika, Peter, and Tim Potter. "Metadata Statistics for a Large Web Corpus." LDOW 937 (2012).
4. Hawksey, M., Barker, P., and Campbell, L.M. New Approaches to Describing and Discovering Open Educational Resources. In Proceedings of OER13: Creating a Virtuous Circle. Nottingham, England (2013).
5. Rosati, A., Mayernik, M. Facilitating Data Discovery by Connecting Related Resources. Semantic Web Journal (2013).
6. http://www.lrmi.net/about
7. Bratsas, C., Bamidis, P., Dimou, A., & Antoniou, I. Semantic CMS and Wikis as Platforms for Linked Learning. Linked Learning 2012, Lyon, France.
8. http://vocab.org/aiiso/schema
9. http://linkedscience.org/teach/ns/
10. Dietze, Stefan, Salvador Sanchez-Alonso, Hannes Ebner, Hong Qing Yu, Daniela Giordano, Ivana Marenzi, and Bernardo Pereira Nunes. "Interlinking educational resources and the web of data: A survey of challenges and approaches." *Program: electronic library and information systems* 47, no. 1 (2013): 60-91.