

# Linking a Community Platform to the Linked Open Data Cloud

Enayat Rajabi  
Universidad de Alcalá  
Computer Science Department  
28805 Alcalá de Henares, Spain  
enayat.rajabi@uah.es

Ivana Marenzi  
University of Hannover  
Appelstrasse, 4  
30167 Hannover, Germany  
marenzi@L3S.de

## ABSTRACT

Linked Data promises access to a vast amount of resources for learners and teachers. Various research projects have focused on providing educational resources as Linked Data. In many of these projects the focus has been on interoperability of metadata and on linking them into the linked data cloud. In this paper we focus on the community aspect. We start from the observation that sharing data is most valuable within communities of practice with common interests and goals, and community members are interested in suitable resources to be used in specific learning scenarios. The community of practice we are focusing on is an English language teaching and learning community, which we have been supporting through the LearnWeb2.0 platform for the last two years. We analyse the requirements of this specific community as a basis to enrich the current collected materials with open educational resources taken from the Linked Data Cloud. To this aim, we performed an interlinking approach in order to enrich the learning resources exposed as RDF (Resource Description Framework) in the LearnWeb2.0 platform with additional information taken from the Web.

## Categories and Subject Descriptors

H.3.5 Online Information Services

## Keywords

Linked Data, metadata, learning object, language teaching community, teacher training, LearnWeb2.0, interlinking.

## 1. INTRODUCTION

Even though nowadays the Web is a huge source of information, it is not easy to find resources suitable for learning: it is difficult to define whether a source of information is reliable, or whether a resource (e.g. a YouTube video) has pedagogical value in a specific learning context. One difficulty is also to find resources relevant for educational purposes in general, because this is too broad as a goal. Teacher networks and professional online communities offer a valuable opportunity to collaboratively collect, evaluate and share learning materials, as well as to critically reflect and discuss on the pedagogical value of such materials, based upon their practical experience.

Our ultimate goal is to detect user requirements to find appropriate solutions for improving discovery of appropriate learning resources. In this paper, we address in particular the needs of the YELL (*Young English Language Learners*) professional community of teachers [3]. We propose an interlinking approach to match the

existing resources collected by the YELL community on the LearnWeb2.0 platform with DBpedia categories in order to enrich them with additional information taken from the Web.

### 1.1 LearnWeb2.0 for YELL

YELL is a lifelong teacher education project that brings together trainees, teachers, trainers and experts from different schools and backgrounds. The overall aim is to promote collaborative peer-learning and peer-teaching professional improvement for teachers of English as a foreign language across school levels. The collaborative searching and sharing platform LearnWeb2.0 [4], customised for YELL from January 2012 onwards, is the solution that allows the community to expand professionally bridging the gap between university teacher education and school teaching and everyday practice [3].

Since its creation on the 24th of January 2012, the YELL community has created 13 groups and collected 462 resources on LearnWeb2.0. At the time of writing, the categorization of resources in groups is based on the requirements expressed by the various users who contributed to its development, and can be categorised into two wide-ranging sections:

1. User-generated resources: materials produced and/or collected by the users for teaching/learning purposes and shared with the community of peers and educators.
2. Resources generated by others and uploaded by YELL users.

YELL users also added comments to the resources explaining why they are relevant, how they can be used in the classroom, or for what kind of students these resources are more relevant.

### 1.2 Evaluation of the current situation

A case study carried out in 2013 [3] qualitatively analyses the perceptions users have of the YELL platform as professional virtual environment to share materials and find resources. All respondents clearly state that YELL is relevant to them for the teaching materials that they can use for their own teaching in schools. Still, something is missing: in particular there is the need of an easier way to enrich the current resources collected manually with additional OERs from the Web.

At present the developers of the LearnWeb2.0 platform, in collaboration with the YELL teachers, are working towards the enrichment of LearnWeb2.0 resources, with OERs and with contents from the educational Linked Data (LD) cloud.

## 2. METHODOLOGY

Dealing with a wide range of educational resources encourages learners and researchers towards linking and sharing learning resources. Within the YELL community for example, teachers and students collect and share materials suitable for nursery, primary and secondary school learning scenarios.

Copyright is held by the International World Wide Web Conference Committee (IW3C2). IW3C2 reserves the right to provide a hyperlink to the author's site if the Material is used in electronic media.

WWW 2015 Companion, May 18–22, 2015, Florence, Italy.

ACM 978-1-4503-3473-0/15/05.

<http://dx.doi.org/10.1145/2740908.2741742>

Linking online learning courses in different digital repositories to useful knowledge on the Web improves resource seeking [8]. This allows current isolated data platforms to evolve towards an open educational space including resources independent of their geographic and system boundaries. In the following, we outline our approach to link LearnWeb2.0 metadata to the LD cloud by exposing its collection according to the LD principles [2].

## 2.1 Exposure of LearnWeb2.0 metadata

Nowadays, many educational institutions (e.g., the University of Muenster, the Open University and the University of Southampton) exploit their educational resources as Linked Data. One of the practical approaches to represent various kinds of data as LD is mapping the collection in RDF and making it available for interlinking. In our approach, we apply a set of applications for exposing the LearnWeb2.0 metadata as RDF and link them to DBpedia.

### Analyzing the metadata

At the time of this research, LearnWeb2.0 includes 12,150 resources collected by users in two different ways:

- Retrieving the resources by searching them over the Web
- Uploading the actual resources manually and creating the metadata (including additional information such as title, type, description and so on)

The learning resources are mostly PDF documents, PowerPoint presentations, images and videos that have been retrieved from different sources (e.g., YouTube, Flickr, SlideShare), or created by LearnWeb2.0 users. When a resource is uploaded, a new instance of the resource is created and specific additional information (i.e. metadata) can be associated to each instance. All metadata are stored in a MySQL database and maintained in the LearnWeb2.0 file system. In our study, we compared the metadata collected in LearnWeb2.0 with the Dublin Core (DC) metadata standard and we realized that most of them can be aligned to the DC schema (Table 1). Furthermore, we added new elements (e.g., learnWeb:Resource) and shaped all of them as one ontology<sup>1</sup>.

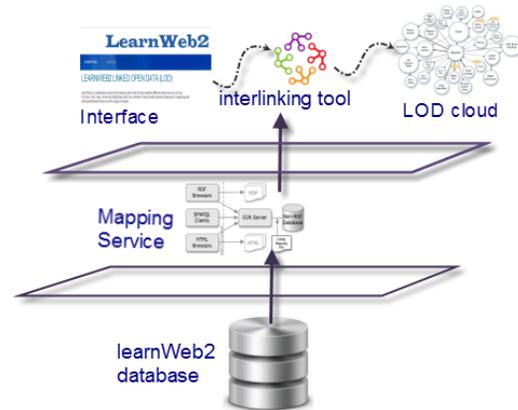
**Table 1. Mapping LearnWeb2.0 metadata to Dublin Core**

LearnWeb2.0 metadata element	DC element
Type of the resource	dcterms:Format
Resource title	dcterms:title
Resource description	dcterms:description
User who published the resource	dcterms:publisher
Resource URL	dcterms:identifier
Resource category	dcterms:subject

### Exposing the metadata in RDF

After mapping the metadata to the DC schema, we applied the D2RQ mapping service [1] to expose the learning resources metadata according to the RDF standard model for data interchange on the Web. Although according to a recent research [8] this approach is not scalable, it can be utilized when the number of resources is not huge as in the LearnWeb2.0 platform. Figure 1

depicts the Linked Data architecture that we developed<sup>2</sup> for exposing the LearnWeb2.0 metadata as Linked Data. This exposure will be completed by adding links to other datasets later.



**Figure 1. Linked Data exposure of LearnWeb2.0 metadata**

### Interlinking to LOD datasets on the Web of Data

Several approaches for interlinking data in the context of Linked Open Data (LOD) have been implemented. In a comparative study [7], several interlinking tools have been discussed from different aspects. According to findings of this study, Silk [9] and LIMES [5], were identified as two most practical tools on the Web of Data that can be used for interlinking purposes. We applied LIMES, which is also a time-efficient approach [5], to examine the interlinking between LearnWeb2.0 resources and the DBpedia dataset as the most important dataset on the Web of Data [6]. As the subject of LearnWeb2.0 resources were provided in English, we performed the interlinking task in English language. The categories in the DBpedia dataset fit well with LearnWeb2.0, as it includes a subject field for each resource object which can be linked to DBpedia as well. Using LIMES, 494 records were matched to the DBpedia categories with more than 98% similarity (according to the LIMES algorithm) and 1,579 records were identified as similar terms with more than 50% similarity. In the following we concentrate on the 494 records that we considered as exact matches. By going through several resulting categories and given that the target dataset provides general information about each one, a user can find additional details about the resource (e.g., its classification in various languages and references to other subjects related to the topic). Furthermore, users can access other related resources that are listed in the same category in the DBpedia dataset.

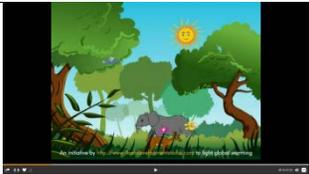
In order to understand the actual relevance of such results, a sample of 300 links over the 494 results was manually evaluated. Three researchers (including one of the authors) opened each resource in LearnWeb2.0 and compared them with the associated DBpedia category reported by LIMES so as to verify that the interlinking is really relevant. 259 over 300 resources were evaluated as relevant by all three evaluators, and 10 over 300 resources were evaluated as relevant by two evaluators. In total 269 matches were considered relevant by at least two raters, resulting in 89% of relevance. In order to statistically assess the reliability of agreement between the three raters, we calculated the Fleiss' kappa number. Since the Fleiss' kappa can be used only with binary or nominal-scale ratings, we created a matrix of the evaluation results where each evaluator indicated whether a match was relevant or not. We obtained an

<sup>1</sup><http://learnweb.l3s.uni-hannover.de/linkedata/learnWeb.owl>

<sup>2</sup>Exposure is available at <http://learnweb.l3s.uni-hannover.de/linkedata/>

agreement of 0.83. Table 2 shows a sample of these well matched results. Also, 31 matches over the 300 investigated were considered not relevant. For example, 20 resources were associated by LIMES to numerical DBpedia categories.

**Table 2. Sample of results associated to DBpedia categories**

Links	Resource
<p>DBpedia category: <a href="#">Global_warming</a></p> <p>LearnWeb: <a href="http://learnweb.l3s.uni-hannover.de/lw/resource.jsf?resource_id=6043">http://learnweb.l3s.uni-hannover.de/lw/resource.jsf?resource_id=6043</a></p> <p>Original resource: <a href="http://www.slideshare.net/Flashdevelopmentstudio/global-warming-4758959">http://www.slideshare.net/Flashdevelopmentstudio/global-warming-4758959</a></p>	 <p>Video appreciated by school teachers in Slideshare: “I am a teacher for high school students. Your material is just what I need for my class.”</p>
<p>DBpedia category: <a href="#">KID_games</a></p> <p>LearnWeb: <a href="http://learnweb.l3s.uni-hannover.de/lw/resource.jsf?resource_id=75032">http://learnweb.l3s.uni-hannover.de/lw/resource.jsf?resource_id=75032</a></p> <p>Original resource: <a href="http://learnenglishkids.britishcouncil.org/en/kids-games">http://learnenglishkids.britishcouncil.org/en/kids-games</a></p>	 <p>Games for children in various school levels: teens and kids. This resource matches the search requirements of the YELL teachers.</p>

### 3. CONCLUSION & FUTURE WORK

Linking educational resources to e-learning repositories on the Web of Data has the potential to make enriched content accessible for learners and teachers. In this paper, we applied a Linked Data approach for connecting the LearnWeb2.0 learning platform, to LOD datasets. We started from specific user requirements collected from the YELL community of teachers, highlighting the need for accessing more enriched online resources to support their courses. To address these specific requirements we conducted a feasibility study to investigate a preliminary mapping approach for exposing learning materials as Linked Data and linking them to Linked datasets. The current outcomes of the study show that the LearnWeb2.0 platform, and the resources collected by the YELL community as a specific scenario, can be linked to other datasets on the Web of Data.

From a technical point of view, a positive outcome of our preliminary interlinking work is the uniform categorization of LearnWeb2.0 resources. LearnWeb2.0 allows users to access and collect materials from various datasets and web sources, each of them having a different way to categorize their resources.

Matching LearnWeb2.0 resources with DBpedia categories provides a more coherent categorization. Furthermore, it allows enriching the LearnWeb2.0 resources with related DBpedia occurrences in various languages that is relevant in language learning scenarios such as the YELL community.

In the next step, after gathering more data about teachers information needs, we plan to interlink LearnWeb2.0 with other datasets as for example the Open Discovery Space portal<sup>3</sup> and to enrich the LearnWeb2.0 resources with additional resources taken from the Linked Data cloud which are relevant specifically for the YELL teachers.

### 4. ACKNOWLEDGMENTS

The work presented in this paper has been funded in part by the Open Discovery Space project (Grant Agreement: 297229) under the ICT Policy Support Programme CIP-ICT-PSP.2011.2.4-e-learning, and by the LinkedUp project (Grant Agreement: 317620) under the FP7 Programme of the European Commission.

### 5. REFERENCES

- [1] Bizer, C. 2004. D2RQ - treating non-RDF databases as virtual RDF graphs. In Proceedings of the 3rd International Semantic Web Conference (ISWC2004).
- [2] Bizer, C., Heath, T., Berners-Lee, T. 2009. Linked Data - The Story So Far. *Int. J. Semantic Web Inf. Syst.* 5, 1–22.
- [3] Bortoluzzi, M., and Marenzi, I. 2013. YELLing for Partnership: A social platform for sharing practice and reflection in teacher training for language learning. In Riem, A., Dolce, M.R., Mercanti, S. and Colomba, C. (Eds). *The Tapestry of the Creative Word in Anglophone Literatures*, 249-262. Udine: Forum.
- [4] Marenzi, I., and Zerr, S. 2012. Multiliteracies and active learning in CLIL: the development of LearnWeb2.0. *IEEE Transactions on Learning Technologies (TLT)*. Los Alamitos, CA, USA, 5(4), 336-348.
- [5] Ngonga, A., Sören, A. 2011. LIMES - A Time-Efficient Approach for Large-Scale Link Discovery on the Web of Data. Presented at the IJCAI.
- [6] Rajabi, E., Sanchez-Alonso, S., Sicilia, M.-A. 2013. Analyzing Broken Links on the Web of Data: an Experiment with DBpedia. *J. Am. Soc. Inf. Sci. Technol. JASIST* 65.
- [7] Rajabi E., Sicilia M. A., Sanchez-Alonso S. 2014. An empirical study on the evaluation of interlinking tools on the Web of Data, *J. of Information Science* 2014; 40: 637-648.
- [8] Rajabi, E., Sicilia, M., Sanchez-Alonso, S. 2015. Interlinking Educational Resources to Web of Data through IEEE LOM. *Computer Science and Information Systems*, 12, 1, 233–255.
- [9] Volz, J., Bizer, C., Berlin, F.U., Gaedke, M., Kobilarov, G. 2009. n.d. Silk. – A Link Discovery Framework for the Web of Data, 2nd Workshop about Linked Data on the Web (LDOW2009), Madrid, Spain.

<sup>3</sup><http://portal.opendiscoveryspace.eu/>