
ENRICHING THE E-LEARNING CONTENTS USING INTERLINKING

ENAYAT RAJABI, SALVADOR SANCHEZ-ALONSO
Information Engineering Research Unit, Computer Science Department,
University of Alcalá, 28805, Spain
{enayat.rajabi, salvador.sanchez}@uah.es

Abstract: *Interlinking e-Learning resources on the Web enriches the contents and help learners to reach more qualified resources. Due to the acceptance of the Linked Data approach by many e-Learning repositories, creating links between e-Learning contents and useful information on the Web of Data can lead to the enrichment of the contents in the repositories. This paper focuses on the text enrichment aspect of interlinking, where we discuss how the descriptions of learning objects as part of the metadata can be enriched when they are connected to the LOD datasets. A case study on Open Discovery Space (as an e-Learning repository) is also presented in this research.*

Keywords: *E-Learning, Linked Data, Interlinking, Text enrichment.*

al [2] investigated several solutions aimed at linking educational resources by exposing them as Linked Data

1. INTRODUCTION

The emergence of the Linked Data approach [1] in recent years has facilitated the discoverability and availability of e-Learning contents on the Web of Data [2]. Moreover, a learning resource is enriched when it is connected to useful and related information on the Web. To take an example, when a researcher explores a learning portal to discover a specific resource, she might find a course on the subject of her interest. She becomes more precise on one of the courses' description, but there exists several specific terms that she does not have any clue about. As the resources in the portal were previously interlinked with DBpedia¹, she finds more explanations about the terms including their different translations or can be redirected to the related resources. As the example above shows, creating links between e.g., descriptions of a resource in a learning repository to related datasets on the Web, enriches, defines and to some extent explains the text itself. Moreover, it improves the level of knowledge of learners, as they discover more digital objects in different formats (e.g., pictures, video) in various domains when they explore a learning portal.

This brief paper focuses on the enrichment aspect of interlinking. Here, we will show that how we can enrich a text by linking its terms to the LOD datasets. The rest of this paper is structured as follows. In Section 2, we will briefly describe the related works in the interlinking context and exposing e-Learning resources as Linked Open Data. Section 3 presents the proposed approach along with a case study. Finally, conclusion is provided in Section 4.

2. RELATED WORKS

Interlinking different kinds of data on the Web has been recently addressed by several studies [2] [3] [4]. Dietze et

¹ <http://www.dbpedia.org>

using interlinking techniques. Another approach, also proposed by Dietze et al [3], proposed linking learning resources to exploit the wealth of existing technology-enhanced learning (TEL) data on the Web of Data in the context of the mEducator project [5]. A study by Fernandez et al [6] suggests linking educational resources across universities using the Linked Data approach by focusing on exploiting information of lectures in video format produced by several educational institutions according to some vocabularies, e.g. FOAF. In another empirical study, Rajabi et al. [4] applied two matching techniques to interlink a semi-structure learning collection to 2 datasets and discussed the generated results in details. Several projects such as LinkedUp [7] and Linked Universities [8] have been devoted to reuse and share e-learning resources based upon the Linked Data principles. Notably, LinkedUp aims at advancing the exploitation of the vast amounts of public data available on the Web in educational institutions and organizations.

In the case of linking tools and applications, a framework for data interlinking was proposed by Scharffe and Euzenat [9]. In this effort, several linking tools were discussed and compared to be applied in interconnecting data. Datalift [10], another project in this area, described the data linking task within the project deliverables and analyzed 11 linking tools to select the most appropriate software.

Given its short length, this paper focuses on one of the advantages of interlinking only: text enrichment. Although there exist a lot of interlinking tools for finding similarities between datasets on the Web, parsing and enriching a large metadata description of learning objects by making use interlinking has not been largely explored.

3. APPROACH

Interlinking e-Learning contents can be taken place in various contexts with different approaches. One of the solutions to carry out the interlinking among educational repositories implies the use of interlinking tools to connect some parts of an e-Learning resource to useful information and datasets on the Web. Enriching the

learning object description and the rest of its metadata elements with some external links can help students and educators to extend their knowledge about the context. To identify an appropriate element in the metadata of learning objects, we investigated several schemas in the e-Learning context and realized that the description of an e-Learning resource is one of the main specifications of each metadata schema, as it describes the resource itself. The content of this description is usually represented under the form of descriptive text strings in different languages and includes a detail explanation about the resource using specific terms. Enriching descriptions with new links to references on the Web also improves the quality of the content. Appendix shows the proposed approach for text enrichment using interlinking. The process in the workflow shows how after a text extraction step, an interlinking process is carried out. In particular, each term of the text is tested to be interconnected to the Web of Data. As a result, a list of similarities for each text is presented to the repository owner. The content owner reviews the list of results for each resource and decides either to accept or to reject each individual term. The accepted concepts are later imported to the actual repository so that the final enriched text is presented to the end users. The acceptance of each term may refer to its importance in the context. In particular, one term may be approved because it is very scientific or difficult for learners to understand the meaning in the context. In this case, providing a link to a dataset e.g., DBpedia describes it in detail.

To implement this approach as a case study, we selected several random e-Learning resources from Open Discovery Space (ODS)², as a collaborative and multilingual open learning infrastructure designed to boost demand for Europe-wide e-Learning Resources. At the time of research, ODS includes nearly 700K educational resources collected from 24 digital repositories in e-Learning context. ODS has currently released a sample of its e-Learning resources as Linked Open Data temporarily³. This endpoint allows other educational datasets to query the ODS learning contents. As a consequence, the description of 20 resources in English language were retrieved using SPARQL⁴ queries. As part of the interlinking, a JAVA program read each individual e-Learning resource description and split it to several string tokens in order to check them against the DBpedia dataset. If the term was available in DBpedia, we stored the term along with the corresponding URL as an output. The following table illustrates a sample of output in which the user rejected some results due to their simplicity.

² <http://www.opendiscoveryspace.eu>

³ <http://83.212.86.12:3030/>

⁴ <http://www.w3.org/TR/rdf-sparql-query/>

Table 1: A sample of results

Term	Link	Accepted
an	http://www.dbpedia.org/page/An	No
insurance	http://www.dbpedia.org/page/Insurance	Yes
market	http://www.dbpedia.org/page/Market	Yes
with	http://www.dbpedia.org/page/With	No
asymmetric	http://www.dbpedia.org/page/Asymmetric	Yes

Figure 2 depicts the percentage of linking between the descriptions of 20 selected learning resources and DBpedia. The X-axis illustrates the resource number whose description was selected for our evaluation, while the Y-axis shows the percentage of terms linked to DBpedia. As an example, resource 1 had 70 words of which 51 (nearly 73%) were linked to DBpedia. As it can be seen, around 77% of the terms in average were linked to DBpedia pages before the user review.

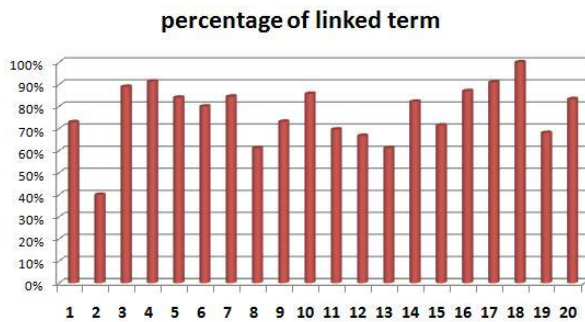


Image 2: percentage of interlinking for 20 selected texts

5. CONCLUSION

This paper concisely reported on a research about interlinking educational resources to the Web of Data. We believe that enriching e-Learning metadata elements, particularly the description element studied in this paper, helps learners and teachers to reach more useful information on the Web. Beyond that, the repository owners can connect their content to more than one datasets at the same time depending on the context of learning resources and their complexity, particularly when they include scientific terms and vocabularies.

ACKNOWLEDGEMENT

The work presented in this paper has been part-funded by the European Commission under the ICT Policy Support Programme CIP-ICT-PSP.2011.2.4-e-learning with project No. 297229 “Open Discovery Space (ODS)”.

REFERENCES

[1] C. Bizer, T. Heath, and T. Berners-Lee, “Linked Data - The Story So Far,” *International Journal on Semantic Web and Information Systems*, vol. 5, no.

3, pp. 1–22, 33 2009.
 [2] S. Dietze, S. Sanchez-Alonso, H. Ebner, H. Q. Yu, D. Giordano, I. Marenzi, and B. P. Nunes, “Interlinking educational resources and the web of

- data: A survey of challenges and approaches,” *Program: electronic library and information systems*, vol. 47, no. 1, pp. 60–91, Feb. 2013.
- [3] S. Dietze, H. Q. Yu, D. Giordano, E. Kaldoudi, N. Dovrolis, and D. Taibi, “Linked education: interlinking educational resources and the Web of data,” in *Proceedings of the 27th Annual ACM Symposium on Applied Computing*, New York, NY, USA, 2012, pp. 366–371.
- [4] E. Rajabi, M.-A. Sicilia, and S. Sanchez-Alonso, “Interlinking Educational Data: an Experiment with GLOBE Resources,” presented at the First International Conference on Technological Ecosystem for Enhancing Multiculturality, Salamanca, Spain, 2013.
- [5] E. Mitsopoulou, L. Woodham, C. Balasubramaniam, T. Poulton, A. Protopsaltis, and S. Dietze, “mEducator: multi type content repurposing and sharing in medical education,” *The Academy Subject Centre for Medicine, Dentistry and Veterinary Medicine Newsletter 01*, vol. 22,, pp. 26–28, Jul. 2010.
- [6] M. Fernandez, M. d’ Aquin, and E. Motta, “Linking data across universities: an integrated video lectures dataset,” Berlin, Heidelberg, 2011, pp. 49–64.
- [7] “LinkedUp: Linking Web Data for Education,” 13-May-2013. [Online]. Available: <http://linkedup-project.eu/>. [Accessed: 13-May-2013].
- [8] “Linked Universities :: Home.” [Online]. Available: <http://linkeduniversities.org/lu/>. [Accessed: 13-May-2013].
- [9] F. Scharffe and J. Euzenat, “MeLinDa: an interlinking framework for the web of data,” arXiv e-print 1107.4502, Jul. 2011.
- [10] “The Datalift project, a catalyser for the Web of data.” [Online]. Available: <http://datalift.org/en>. [Accessed: 24-Nov-2013].

Appendix: Proposed approach for the enrichment of the contents of eLearning repositories

