

# Querying Repositories of OER Descriptions: The Challenge of Educational Metadata Schemas Diversity

Hiba Hajri<sup>(✉)</sup>, Yolaine Bourda, and Fabrice Popineau

LRI, CentraleSupélec, bat. 650 (PCRI), 91405 Orsay, France  
{hiba.hajri,yolaine.bourda,fabrice.popineau}@centralesupelec.fr

**Abstract.** Personalized applications like recommendation systems or open adaptive hypermedia require to facilitate reuse and interoperability of educational resources. Thus we have to solve the problem of discovering educational datasets and the problem of the fragmented landscape of competing metadata schemas. In this paper, we compare a subset of metadata schemas used by educational datasets obeying the Linked Open Data principles and we propose mappings between them.

**Keywords:** Linked open data · Educational datasets · Metadata schemas

## 1 Introduction

The efficiency of personalized applications of Technology-Enhanced Learning (TEL) depends in particular on the quality of the metadata describing the education open resources (OER) [6]. In this article we are interested in personalization based on OER and especially on those supporting interoperability [3] with descriptions following the Linked Open Data principles (LOD) [1].

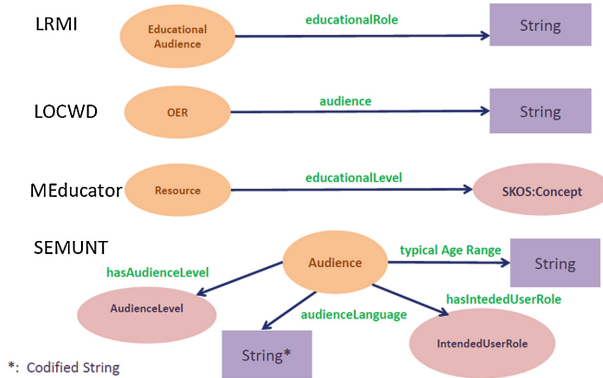
We start by searching for repositories accessible on the Web, which contain descriptions of OER and which provide SPARQL endpoints [10] for data access. Therefore we look for the *Linked Education Catalog*<sup>1</sup> provided by the Linked Up Project<sup>2</sup>. It is a catalog of data sets relevant to educational applications with SPARQL endpoint for each one and also a common SPARQL endpoint that provides access to VOID<sup>3</sup> descriptions of currently included datasets [2]. There are also other repositories in the literature not referenced in the Linked Up Project like SemUNT [7] which is a project of French digital thematic universities.

Once we have found accessible repositories and in order to be able to interrogate them with one SPARQL query for a given question (e.g. finding exercises about calculus for undergraduate students), we have to study closely the distinct metadata schemas used. We start with comparing a subset of metadata schemas.

<sup>1</sup> <http://data.linkededucation.org/linkededup/catalog/>.

<sup>2</sup> <http://linkedup-project.eu/>.

<sup>3</sup> <http://www.w3.org/TR/void/>.



**Fig. 1.** Elements describing the audience

Actually the schemas specific to education and more precisely OER are either the LOM [8] and its applications profiles or extensions of generic schemas such as Dublin Core (DC)<sup>4</sup> and FOAF<sup>5</sup>.

In our comparison, we focus on a subset of schemas which have complete RDF or OWL bindings and which are specific for OER description. We choose firstly Learning Resource initiative(LRMI)<sup>6</sup>: a schema based on DC. Secondly, Linked OpenCourseWare Data (LOCWD) [9]: it describes OER in the OpenCourseWare context. Thirdly mEducator [5]: a LOM application profile for the health sector. And the last one SemUNT [7]: a schema defined from an application profile of the LOM for French UNT.

The principle of our comparison is to take all educational aspects and to study how they are described in each one of these schemas. An example of educational aspects is the description of the audience (for whom is the OER?). In Fig. 1, we detail the elements used by each schema to specify this aspect.

## 2 Definition of Mappings and SPARQL Queries

The study of schemas enabled us to define mappings or SPARQL queries, as the case may be, based on our comparison results. The problem of schema mapping has been pointed out as one of the big challenges when interlinking educational resources and LOD. Many approaches have been proposed. Fully automated approaches are more scalable than manual approaches, but they perform lower in terms of precision [4]. Thus we choose to use a manual approach in our work because personalized applications need to be based on precision and also because we don't have a large number of classes and properties to deal with. To this end we define the LOOM: *Learning Object Ontology of Mapping* devoted to the specification of mappings. We note that the defined mappings have, as unique

<sup>4</sup> <http://dublincore.org/documents/dces/>.

<sup>5</sup> <http://xmlns.com/foaf/spec/>.

<sup>6</sup> <http://dublincore.org/dcx/lrmi-terms/2015-02-09/>.

purpose, to allow writing queries for repositories based on diverse educational metadata schemas. Consequently when one of the used metadata schemas is changed, we have just to change the mapping defined in our ontology.

Our LOOM ontology provides the following classes and properties:

**loom:Learning Resource** a class representing educational resources;

**loom:Audience** a class representing the audience of the learning resource;

**loom:educationalRole** a property used to describe the role of the audience intended by the resource (e.g. Teacher, Author, Learner, etc.). Here is its definition:

```
loom:educationalRole    owl:datatypeProperty;
                        rdfs:domain loom:LearningResource.
```

**loom:educationalLevel** a property used to describe the level of the audience intended by the resource (e.g. Master, License, etc.). Here is its definition:

```
loom:educationalLevel  owl:datatypeProperty;
                        rdfs:domain loom:Audience.
```

The next step is to define mappings with the elements of the educational metadata schemas cited in our comparison. First, we define mappings between the classes representing the resource:

```
loom:Learning Resource  owl:EquivalentClasses  mEducator:Resource
loom:Learning Resource  owl:EquivalentClasses  semunt:LearningResource
loom:Learning Resource  owl:EquivalentClasses  locwd:OER
loom:Learning Resource  owl:EquivalentClasses  lrmi:CreativeWork
```

Secondly we define mappings between the classes representing the audience:

```
loom:Audience  owl:EquivalentClasses  lrmi:Educational Audience
loom:Audience  owl:EquivalentClasses  semunt:Audience
```

Thirdly, we define mappings between properties describing the audience level:

```
loom:educationalLevel  owl:equivalentProperty  semunt:hasAudienceLevel
loom:educationalLevel  owl:equivalentProperty  mEducator:educationallevel
```

The property *educationalLevel* of *mEducator* is a property of the class *Resource* while *hasAudienceLevel* of *SemUNT* is a property of the class *Audience*. Thus we propose the following SPARQL query to select the resources and the educational levels of SEMUNT and mEducator.

```
SELECT ?lr ?el
WHERE {
    ?lr a loom:LearningResource;
        hasAudience?/loom:educationalLevel ?el}
```

Concerning the properties dealing with the audience role, we propose the following mappings:

```
loom:educationalRole owl:equivalentProperty lrmi:educationlRole
loom:educationalRole owl:equivalentProperty semunt:hasIntendedUserRole
loom:educationalRole owl:equivalentProperty locwd:audience
```

While, on the one hand, *educationalRole* of LRMI and *hasIntendedUserRole* of SemUNT are properties of the classes representing the audience, on the other hand, *audience* of LOCWD is a property of the class *OER*. So we propose the following SPARQL query.

```
SELECT ?lr ?er
WHERE {
    ?lr a loom:LearningResource;
        hasAudience?/loom:educationalRole ?er}
```

### 3 Future Work

This work is only the beginning and there is still a range of shortcomings that need to be addressed. It is important to continue the study with other used metadata schemas and to compare them on the basis of all educational aspects. Another future work is to use the results of this comparison in order to complete our LOOM ontology. As longcomings, we have to benefit from all these outcomes to integrate a personalized application to a platform of MOOC or LMS such as the OpenEdx platform.

### References

1. Bizer, C., Heath, T., Berners-Lee, T.: Linked data: principles and state of the art. In: World Wide Web Conference, pp. 1–40 (2008)
2. d’Aquin, M., Adamou, A., Dietze, S.: Assessing the educational linked data landscape. In: Proceedings of the 5th Annual ACM Web Science Conference, pp. 43–46. ACM (2013)
3. Dietze, S., Drachsler, H., Giordano, D.: A survey on linked data and the social web as facilitators for TEL recommender systems. In: Manouselis, N., Drachsler, H., Verbert, K., Santos, O.C. (eds.) Recommender Systems for Technology Enhanced Learning, pp. 47–75. Springer, New York (2014)
4. Dietze, S., Yu, H.Q., Giordano, D., Kaldoudi, E., Dovrolis, N., Taibi, D.: Linked education: interlinking educational resources and the web of data. In: Proceedings of the 27th Annual ACM Symposium on Applied Computing, pp. 366–371. ACM (2012)
5. Evangelia, M., Taibi, D., Giordano, D., Dietze, S., Yu, H.Q., Bamidis, P., Bratsas, C., Woodham, L.: Connecting medical educational resources to the linked data cloud: the mEducator RDF schema, store and API (2011)
6. Hylén, J.: Open educational resources: opportunities and challenges. In: Proceedings of Open Education, pp. 49–63 (2006)

7. Isaac, Y., Bourda, Y., Grandbastien, M.: Semunit-french unt and linked data. In: LiLe-2012 at WWW-2012, vol. 840, pp. 1–6. CEUR workshop proceedings (2012)
8. Nilsson, M., Palmér, M., Brase, J.: The LOM RDF binding: principles and implementation. In: Proceedings of the Third Annual ARIADNE Conference, Leuven, Belgium (2003)
9. Piedra, N., Tovar, E., Colomo-Palacios, R., Lopez-Vargas, J., Chicaiza, A.J.: Consuming and producing linked open data: the case of opencourseware. *Progr. Electron. Libr. Inf. Syst.* **48**(1), 16–40 (2014)
10. Quilitz, B., Leser, U.: Querying distributed RDF data sources with SPARQL. In: Bechhofer, S., Hauswirth, M., Hoffmann, J., Koubarakis, M. (eds.) ESWC 2008. LNCS, vol. 5021, pp. 524–538. Springer, Heidelberg (2008)