

In Search of Reusable Educational Resources in the Web

Steinberger, Claudia^a

^aDepartment of Applied Informatics, Universität Klagenfurt, Austria.

Abstract

Nowadays there is a high demand from teachers to precisely find online learning resources that are free from copyright restrictions or publicly licensed to use, adapt and redistribute in their own courses. This paper investigates the state of the art to support teachers in this search process. Repository based strategies for dissemination of educational resources are discussed and critiqued and the added value of a semantic web approach is shown. The ontology schema.org and its suitability for semantic annotation of educational resources is introduced. Current ways and weaknesses to discover educational resources based on appropriate semantic data are presented. The possibility to use the wisdom of the crowd of learners and teachers defining semantic knowledge about used learning resources is addressed. For demonstration purposes within all sections the course subject 'Semantic SEO', dealt in the course 'SEO – Search Engine Optimization' held by the author in 2016, is used.

Keywords: *semantic web; educational resources; schema.org; LRMI; semantic SEO.*

1. Introduction

To improve student learning in their courses teachers more and more frequently use a special type of blended learning and flip their classrooms. A flipped classroom shifts instruction to a learner-centered model abroad the classroom whereas during class time topics are explored in greater depth and meaningful learning opportunities are created (Tucker, 2012). For self-study outside the classroom teachers deliver educational resources to their students online. Scholarly articles, digital documents, blogpostings, video-objects, audio-objects, books and other learning resources are offered to the students mostly via a website or a learning management platform. During class they use learning resources like case studies, assessments, work group instructions or role playing games too.

However, learning resource creation is very time consuming and appropriate high quality educational resources have been already created by others in many cases. So reuse of existing materials as they are, adapted or incorporated into own materials is highly desirable for teachers designing or updating their courses. Thus teachers search on the Web for appropriate and open to use educational resources to integrate them into their individual courses.

Web search engines make access easy but the vastness of material available makes it very challenging for a teacher to find educational resources to address their specific student needs (Yu L., 2014, p 505). Let us say we are looking for high quality text suited to introduce students into the subject of 'semantic SEO'. Google lists more than 500.000 results for 'semantic SEO'. There are too many irrelevant results and the process of browsing through thousands of potential hits to find the ones that meet the specific needs are very time consuming. Also Google Scholar results for 'semantic SEO' are not well suited too because resulting scholarly articles are usually too high sophisticated for introduction purposes.

This means that there is a high demand for teachers finding reusable educational resources more precisely. This paper demonstrates capabilities and investigates further enhancements to improve teachers search performance especially using possibilities of the Semantic Web.

Section 2 of this paper investigates and criticizes present technology strategies for dissemination of content as educational resources and shows the added value of a Semantic Web approach. Section 3 focuses on *schema.org* ontology to semantically describe learning resources and also shows how to disseminate corresponding structured data. Section 4 focuses on semantic discovery of educational resources and shows how to use customer defined search engines for this purpose. Section 5 summarizes results and identifies discovered weaknesses. For demonstration purposes within all sections the course subject 'Semantic SEO' dealt in the course 'SEO – Search engine optimization' held by the author in 2016, is used.

2. Strategies for Educational Resource Semantic Description and Dissemination

(Barker & Campbell, 2016) illustrate a range of technical approaches employed to disseminate educational resources. Present technology strategies include institutional repositories and websites, subject specific repositories, sites for sharing specific types of content (such as video, images, ebooks), general global repositories and also services that aggregate content from a range of collections.

As examples (U-Now), (MIT OCW), (OpenSpire), (BBC) or (OpenLearn) *are institutional repositories*. *Subject specific repositories* and aggregators like (Humbox), (Kritikos) or (CORE-Materials) are generally designed to support subject discipline communities across multiple institutions. The materials come from a variety of sources mostly associated with UK higher education, some industry, third sector and overseas organizations. They host particular domain specific resource types and use specialized resource descriptions vocabularies. Some repositories have means of syndicating information about their resources to aggregators, but the emphasis placed on syndication varies.

Content type specific repositories such as YouTube, iTunesU, SlideShare, Scholar, Flickr and expert's blogs are currently the most popular and successful repositories of learning materials. These platforms each focus on a single media type like video, audio, presentations, images or texts and tend to make resources available for all to view. Due to their popularity and ubiquity, these sites set user expectations for the dissemination and delivery of learning resources on the web and are more sustainable than the education sector services and institutional repositories mentioned before.

Global repositories and aggregators like (MERLOT), (Solvonauts) or (OER Commons) are not limited by subject or resource type and include links to tens of thousands of peer reviewed educational resources. Their geographic scope is global, however there is a preponderance of material from the US and UK. There have been many benefits but there still exists a significant barrier to finding educational resources. Teacher's awareness of educational resource repositories is still limited. They favor web search engines but a central search across several repositories has not been available yet. If a teacher is aware of a repository, most metadata he can find there is about the content itself and not about its educational use and quality (e.g. see <http://bit.ly/2j7f7Ki>). Essential learning resource criteria like level of quality, actuality, rating value, level of complexity, learning time or intended audience cannot be considered to find appropriate results (Yu L., 2014, p.507f). Thus, current repositories are falling short of meeting user's expectations in terms of adequate support for finding appropriate content (Dichev, C., Dicheva, D. , 2012).

As a result no appropriate educational resource could be found for 'semantic SEO' neither in MERLOT nor Solvonauts or OER Commons. Browsing Scholar and expert's blogs some

suited learning resources for ‘semantic SEO’ could be found after a while of search. However, results were found more by accident than by structured search.

The basis for successful resource discovery and retrieval are common vocabularies for meta descriptions that meet user’s expectations and widespread popular tools considering these descriptions. The number of formal meta data standards have emerged over the last decade which attempt to address the issue of educational resource description. A comprehensive description and analysis of learning resource metadata standards is presented in (Dietze, et al., 2013). There are two broad strategies behind learning resource metadata (Barker & Campbell, 2016, p. 67):

- The “traditional” approach of creating catalog records which separate the metadata from the resource, creating a self-contained stand-alone metadata record that fully describes the resource. As we outlined above repositories using this approach did not really take hold.
- Augmenting web resources with semantic information to assist web search engines and other services the discovery and optimal presentation of learning resources based on their meta data. The *schema.org* initiative has been viewed as a signal of mainstream support for the idea of the semantic web (Yu L. , 2014, p. 475 ff) and we will discuss its strengths and weaknesses for educational resource dissemination and discovery in more detail in section 3.

3. Schema.org and Educational Resources

Schema.org is a joint effort by Google, Bing, Yandex, and Yahoo! launched in 2011 providing a common vocabulary for describing a wide variety of entities which can be found in the Web. At this point in time the *schema.org* contains more than 580 classes to describe the most popular types of web content. The goal of *schema.org* has been to let content publishers embed common machine readable information into their HTML pages in form of microcode, RDFa or JSON-LD. This makes web search engines semantically understand the content and therefore better search results are achievable (Mika, 2015).

In the past years the *schema.org* effort proved to be a success. Publishers have a standard vocabulary now to semantically annotate the same kind of information and tools have been developed to support the annotation process (e.g. www.schemaApp.com). Validators improved: Googles structured data testing tool (<https://search.google.com/structured-data/testing-tool>) is offered by Google to support authors in metadata tagging their content; Content management tools like Wordpress and Drupal extended to automatically produce *schema.org* markup; Semantic search engine optimization became topical.

Since the initial effort in 2011 the *schema.org* vocabulary kept evolving. The Learning Resource Metadata Initiative (LRMI) is a collaborative initiative that has been working since June 2011 to make it easier for teachers and learners to find educational materials using major search engines and specialized resource discovery services (Barker & Campbell, 2014). In 2013 LRMI added missing classes and properties to the core of *schema.org* that make the discovery of learning resources easier (Barker & Campbell, 2014b) now.

Figure 1 shows an abstracted excerpt of *schema.org* vocabulary focusing on LRMI classes and properties. Different kinds of *CreativeWork* in the Web, e.g. *Articles*, *Books* or also *Websites* or *MediaObjects* can be seen and tagged as educational resources with properties like their intended educational use, age range or language. *BlogPosting* and *ScholarlyArticle* specialize *Article*. A *CreativeWork* can be addressed to a certain *Audience*, which can also be an *EducationalAudience* of a special type and role. A *Review* is also a *CreativeWork* about another *CreativeWork* describing and rating it. Via its *educationalAlignment* it is possible to assign a *CreativeWork* to an *AlignmentObject* within an intended *educationalFramework*. As an example figure 2 shows the JSON-JD representation of a *Review* markup of a *BlogPosting* item (see website source code of <http://bit.ly/2pNFGn1>).

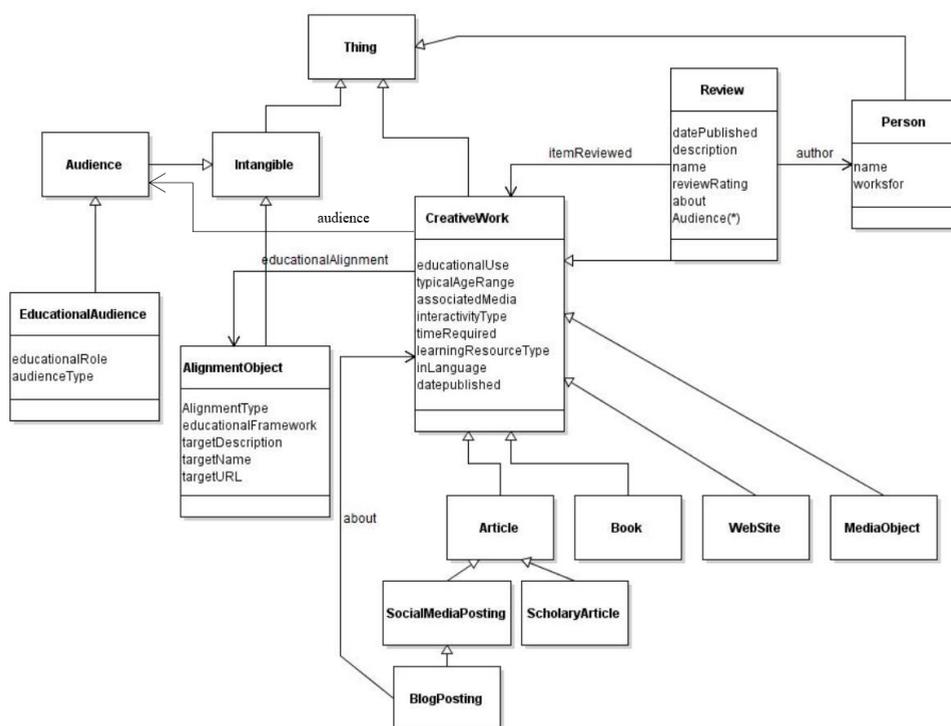


Figure 1 – LRMI vocabulary excerpt of *schema.org*.

```
{ "@context": "http://schema.org",
  "@type": "review",
  "author": { "@type": "Person",
    "name": "Claudia Steinberger"},
  "publisher": { "type": "Organization",
    "name": "AAU-AE"},
  "datePublished": "2017-01-01",
  "audience": { "@type": "EducationalAudience",
    "educationalRole": "teacher",
    "educationalRole": "student",
    "audienceType": "eBusinessInterested"},
  "description": "Really good blogposting to introduce semantic SEO",
  "name": "Introduction into Semantic SEO",
  "image": "http://positiononly.com/blog/wp-content/uploads/2016/02/ecommerce-seo-tactics.png",
  "reviewRating": {
    "@type": "Rating",
    "bestRating": "5",
    "ratingValue": "4",
    "worstRating": "1"},
  "about": "Semantic SEO",
  "learningResourceType": ["introduction", "summary"],
  "educationalUse": "selfstudy",
  "itemReviewed": { "@type": "BlogPosting",
    "@id": "http://positiononly.com/blog/seo/semantic-seo",
    "learningResourceType": ["explanation", "summary"],
    "timeRequired": "P15M",
    "author": "Rohan Ayyar",
    "datePublished": "2016-02-11",
    "headline": "Semantic SEO: Improve your E-commerce Cross-Selling Tactics"},
  "educationalAlignment": { "@type": "AlignmentObject",
    "alignmentType": "teaches",
    "educationalFramework": "Curriculum Information Management",
    "targetName": "Information Management1",
    "targetUrl": { "type": "URL",
      "@id": "http://www.uni-klu.ac.at/rechtabt/downloads/mb120b7_11_12.pdf"
    }
  }
}
```

Figure 2 – JSON-LD structured data about reviewing a Blog posting

One problem became evident to the author when tagging educational resources: most of LRMI *schema.org* properties have text as expected type but no common possible values are defined. This makes a consistent common markup difficult despite of a common vocabulary and subsequent retrieval across different providers almost impossible.

4. Semantic Discovery of Educational Resources

Some holders of educational resources like (BBC) have started to use *schema.org* to markup their materials (Mikroyannidi, Liu, & Lee, 2016). Also some repositories like (MERLOT) have started to add also *schema.org* metadata to their learning resources. Looking at semantic scholar (<https://www.semanticscholar.org>) it becomes evident that *schema.org* is present there too. Search results are tagged as ScholarlyArticles but LRMS-classes and properties are used there insufficiently or not at all.

Schema.org has been applied widely in the last years mainly for semantic search engine optimization purposes. At the moment structured data embedded on webpages are used from search machines automatically but limited. For instance Google Search only uses

special *schema.org* classes (e.g. People, Product, Recipe, Event) to generate Rich Snippets in their search results. This form of use also influences the way and willingness of content authors to tag their contents. For example structured data describing ‘educationalUse’, ‘learningResourceType’, ‘timeRequired’, ‘educationalUse’ or ‘educationalAlignments’ are ignored by Google Search at the moment. To manually query the semantic web Google Custom Search Engines (CSE) can be used. It is very simple to filter on *schema.org* classes in a CSE but to query properties complex queries are necessary. The CSE query language for *schema.org* is not well documented at the moment and complex to use for teachers. Semantic scholar tackles this problem offering a simple user interface to query ScholarlyArticle properties which should be done for Google Search too. It also does not become clear which *schema.org* properties are indexed by Google at all.

5. Summary and Outlook

The common vocabulary to describe educational resources is available in the form of *schema.org*, tools have been developed to support and test the annotation process. But still the following weaknesses could be identified:

- (1) Content providers are willing to tag their content with structured data that bring a clear benefit to them. At the moment this is only the case for search engine optimization purposes for special classes and properties (e.g. Products, Events, Persons and Organizations) but not for the LRMI extension of *schema.org*.
- (2) Possible values for LRMI *schema.org* properties with text as expected type have to be defined commonly to enable search across different providers.
- (3) Not only content authors but also content users (teachers, students) have to be motivated to tag their experiences using educational resources. As an example Figure 2 shows structured data that is added to a teachers Blogposting recommending an educational resource (see <http://bit.ly/2js9jd8>). This knowledge of the crowd could be searched for by others too to find appropriate learning resources for their own courses.
- (4) All *schema.org* classes and properties should be indexed by the major search engines and easy to handle tools have to be developed or made available by search machine providers to individually query all classes and properties.

We can expect research and development within this fields during the next years to support teachers to precisely find appropriate educational resources.

References

- Barker, P., & Campbell, L. M. (2010, 04 29). Metadata for Learning Materials. *Cognition and Learning*, pp. 225-243.
- Barker, P., & Campbell, L. M. (2014). *LRMI Implementation: Overview, Issues and Experienc*. CETIS. Retrieved from <http://publications.cetis.ac.uk/2014/1007>
- Barker, P., & Campbell, L. M. (2014b). Learning Resource Metadata Initiative: using schema.org to describe open educational resources. *Proceedings of OpenCourseWare Consortium Global*.
- Barker, P., & Campbell, L. M. (2016). Technology Strategies for Open Educational Resource Dissemination. In P. Blessinger, & B. T. J., *Open Education: International Perspectives in Higher Education* (pp. 51-70). doi:DOI: 10.11647/OBP.0103.03
- CORE-Materials. (n.d.). Retrieved from <http://core.materials.ac.uk/>
- Dietze, S., Sanchez-Alonso, S., Ebner, H., Qing Yu, H., Giordano, D., Marenzi, I., & Pereira Nunes, B. (2013). Interlinking educational resources and the web of data: A survey of challenges and approaches. *Program*, 47(1), 60-91.
- Humbox. (n.d.). Retrieved from <http://humbox.ac.uk/>
- Kritikos. (n.d.). Retrieved from <https://kritikos.liv.ac.uk/>
- MERLOT. (n.d.). Retrieved from <https://www.merlot.org/merlot/index.htm>
- Mika, P. (2015, 6 23). On Schema.org and Why It Matters for the Web. (IEEE, Ed.) *IEEE Internet Computing*, 19(4), 52-55. doi:10.1109/MIC.2015.81
- Mikroyannidi, E., Liu, D., & Lee, R. (2016). Use of Semantic Web Technologies in the Architecture of the BBC Education Online Pages. In M. d. Dmitry Mouromtsev, *Open Data for Education* (pp. 67-85). Springer International Publishing. doi:10.1007/978-3-319-30493-9_4
- MIT OCW. (n.d.). Retrieved 01 19, 2017, from <https://ocw.mit.edu/>
- OER Commons. (n.d.). Retrieved from <https://www.oercommons.org/>
- OpenLearn. (n.d.). Retrieved 1 19, 2017, from <http://www.open.ac.uk/about/open-educational-resources/openlearn>
- OpenSpires. (n.d.). Retrieved 1 19, 2017, from <http://openspires.oucs.ox.ac.uk/>
- Solvonauts. (n.d.). Retrieved from <https://solvonauts.org/>
- Tucker, B. (2012). The flipped classroom. *Education next*.
- U-Now. (n.d.). Retrieved 01 19, 2017, from <http://unow.nottingham.ac.uk/>
- Wiley, D., Bliss, T. J., & McEwen, M. (2014). Open Educational Resources: A Review of the Literature,. In M. D. J. Michael Spector, *Handbook of research on educational communications and technology* (pp. 781-789). Springer.
- Yu, L. (2014). *A Developer's Guide to the Semantic Web*. Springer-Verlag Berlin Heidelberg. doi:10.1007/978-3-662-43796-4
- Yu, L. (2014). *A Developer's Guide to the Semantic Web*. Springer-Verlag Berlin Heidelberg. doi:10.1007/978-3-662-43796-4