Impact Factor (SJIF): 4.542

International Journal of Advance Research in Engineering, Science & Technology

e-ISSN: 2393-9877, p-ISSN: 2394-2444

Volume 4, Issue 12, December-2017

Survey On Utilizing linked data from massive open online courses

Miss. Vrundan V. Parode[1] Prof. P. L. Ramteke[2]

[1]*Department of computer Science & Engineering H.V.P.M College of Engineering, Amravati*
[2]*Department of computer Science & Engineering ,H.V.P.M College of Engineering, Amravati*

Abstract- ***Linked Data is a trend on the web with companies promoting their own means of marking up data semantically, publishing and connecting data on Web. Despite the increasing popularity of Linked Data, there are a limited number of applications that implement and take advantage of its effectiveness, particularly in the domain of education. In this project we use a semantic technologies to create a semantic data model for educational data, and data about Massive Open Online Courses and publishing this data as linked data on the Web Data from various MOOC providers is integrated and published as Linked Data. present web doorway called MOOC Link that utilizes the data to discover. The term Linked Data refers to a set of best practices for publishing and connecting structured data on the Web***

Keywords- ***ontology engineering, semantic application; linked data; education;***

I. INTRODUCTION

Linked Data involves web to create typed links between different sources of data. Source data may vary in location, size, subject-matter and how congruously it is structured. Links produced from this data create equality and define properties to make data easier to read for machines. Linked Data is relatively unexplored in the domain of education. Linked Up Challenge have surfaced to encourage innovative applications focused on open educational data Massive Open Online Courses or MOOCs are online courses accessible to anyone on the web. many of institutions have joined in an archive to make education more accessible by synchronizing with MOOC providers such as Coursera and NPTEL Delivering course content through lecture videos as well as readings and problem sets, MOOCs encourage interactivity between professors and students around the world by way of discussion forums and graded assessments. for most information associated with their course catalog. This includes properties such as a courses title, instructor, and syllabus details. Although Coursera's course catalog data is easily accessible as JSON, there is no option to retrieve and use it in a Linked Data format such as the RDF .there is no Linked Data available for MOOCs or for ontology that represent properties unlike to MOOCs In order to incorporate MOOC data into the Linked Data cloud as well as demonstrate the potential of Linked Data when applied to education, we propose to (i) developed or extend an RDF ontology which denotes properties of MOOC .(ii)use ontology that generate Linked Data from different MOOC providers

What is Linked Data?

Linked Data is nothing but using the Web to create links between different sources of data. These may be as diverse as databases maintained by two organisations in different geographical locations, or simply heterogeneous systems within one organisation that, historically, have not easily interoperated at the data level. Technically, Linked Data refers to data published on the Web in such a way that it is machine-readable, its meaning is explicitly defined, it is linked to other external data sets, and can in turn be linked to from external data sets.

While the primary units of the hypertext Web are HTML (Hyper Text Markup Language) documents connected by untapped hyperlinks, Linked Data relies on documents containing data in RDF (Resource Description Framework) format (Klyne and Carroll, 2004). However, rather than simply connecting these documents, Linked Data uses RDF to make typed statements that link arbitrary things in the world. The result, which we will refer to as the Web of Data, may more accurately be described as a web of things in the world, described by data on the Web. Berners-Lee (2006) outlined a set of 'rules' for publishing data on the Web in a way that all published data becomes part of a single global data space.

All Rights Reserved, @IJAREST-2017

II. BACKGROUND

A. Resource Description Framework: The most notable model for Linked Data is the Resource Description Framework (RDF), which encodes data as subject, predicate, object triples. The subject and object of a triple are both Uniform Resource Identifiers (URIs), while the predicate specifies how the subject and object are related, also using a URI. For the purposes of this paper, Linked Data is presented as RDF/XML, an XML syntax for RDF, which IS also used in the implementation of our application.

B. Simple Protocol and RDF: Query Language Simple Protocol and RDF Query Language or SPARQL is an RDF query language that allows users to retrieve and manipulate data stored as RDF . SP ARQL is used as our application's query language in order to retrieve data to populate web pages with course information. A SP ARQL endpoint for our data can be accessed at http://sebk.me:3030/sparq J.tp I.

C. Linked Data Principles: Tim Berners-Lee published a set of rules for publishing data on the Web so that data becomes part of a global space in which every resource is connected. The rules are as follows:

- Use URIs as names for things
- Use HTTP URIs so that people can look up those names
- When someone looks up a URI, provide useful information (RDF, SPARQL)
- Include links to other URIs, so that they can discover more things These principles provide a basis for contributing to a Linked Data cloud in which a variety of datasets from different fields of human knowledge are interconnected. Our project aims to abide by these principles.

D. .Linked Education Data: Many models have been devised for structuring educational data, among the most popular are the IEEE Learning Object Metadata (LOM) specification and Sharable Content Object Reference model (SCORM). LOM is encoded in XML and includes nine categories with sub-elements that hold data. An RDF binding for LOM exists, however development is halted at the time of this paper's writing . SCORM is an extensive technical standard, typically encoded in XML, that defines how educational content should be packaged, how it is delivered, and how learners navigate between different parts of an online course . An RDF binding of SCORM has yet to be developed. Rather than using the defunct LOM binding or creating a new binding of SCORM to RDF, we chose to extend a vocabulary meant for Linked Data, Schema.org, for which an RDF mapping exists , to include properties unique to open courseware. In 2013, the Learning Resource Metadata Initiative (LRMI) specification was incorporated into Schema.org's vocabulary for tagging educational content . The properties added in this adoption introduced fields for online course details including the type of learning resource, time required, and so on. While there is significant overlap between LRMI's additions to Schema.org, Schema.org's Creative Work properties and MOOC course details like those provided in Coursera's API, several crucial missing data fields such as syllabus details, course difficulty, and predicates linking courses to other objects, make it necessary to extend the vocabulary for MOOC data.

E. Building Ontologies: After determining that there was no educational resource ontology that denoted every property needed to create linked MOOC data, we chose to extend Schema.org's ontology. In order to support uniformity in our data, we use RDF/XML from ontology creation to final data generation. Schema.rdfs.org hosts an RDF/XML version of Schema.org's ontology, which we imported into Stanford Protege to extend with additional types and properties.

All Rights Reserved, @IJAREST-2017

II. .IMPLEMENTATION

1. Ontology Schema: As mentioned in the previous section on our data model, Schema.org is organized as a hierarchy of types. Much like object-oriented programming, types inherit properties. Figure 1 displays relevant properties from Creative Work and other Schema.org types although our new types, Course, Session and Category, inherit more properties that are not shown. Sebastian's personal domain, sebk.me, is used as a temporary namespace for the ontology

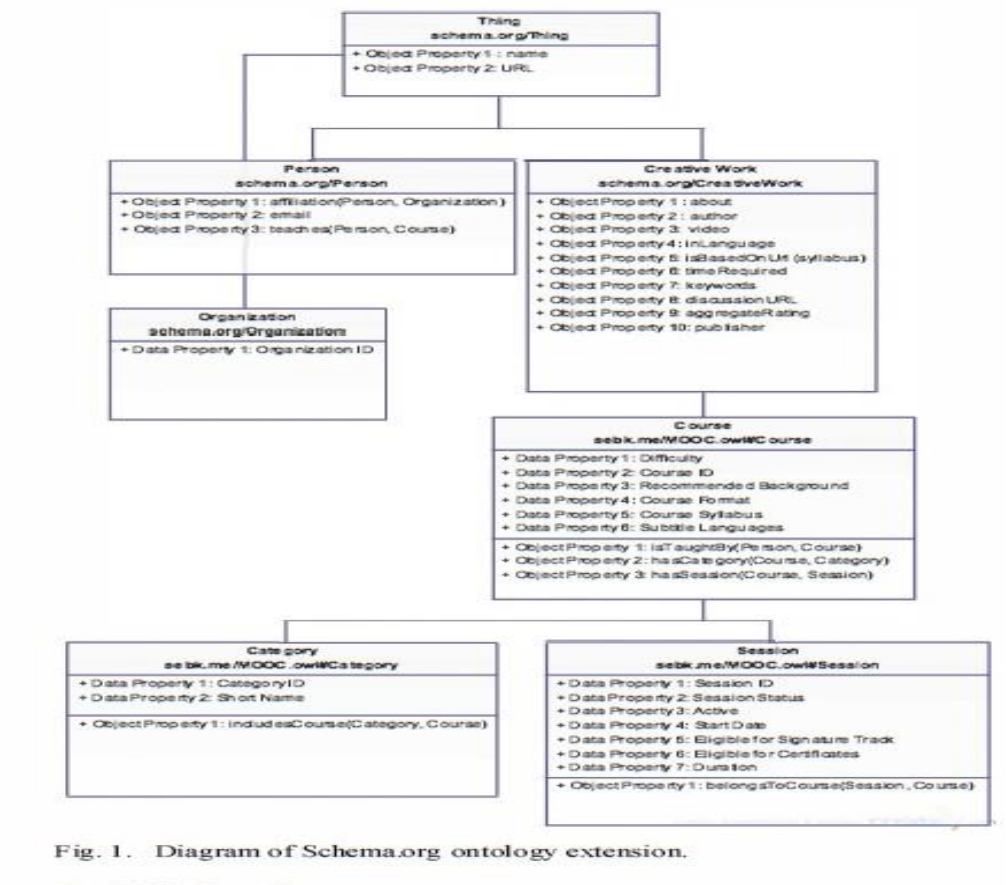

Fig. 1. Diagram of Schema.org ontology extension.

2. RDF Data: The first section of our RDF (Fig. 3) invokes the namespaces associated with (line 2) our ontology, (line 3) RDF, (line 4) OWL, (line 5) XMLSchema, (line 6) Schema.org and (line 7) RDFSchema. These namespaces are used as prefixes to abbreviate URIs throughout the data.

IV. CONCLUSION

We have presented an extension of Schema.org's Linked Data vocabulary, which incorporates types and properties found in online courses. The extended ontology allows for assignment of data properties relevant to MOOCs as well as object properties which link MOOC sessions, course details, categories and instructors together. Also presented is our approach to collecting and generating Linked Data from three MOOC providers: Coursera,. Using Coursera's AP, we collect MOOC data in JSON and convert it to RDF with Apache Jena. We describe a prototype implementation of MOOCLink, a web application which utilizes the Linked MOOC Data to allow users to discover and compare similar online courses.

V. FUTURE SCOPE

Future work will focus on: incorporating demographic data, reviews, developing an item pipeline for our crawlers, automating website updates, enabling user profiles, course tracks, natural language processing of syllabi and summaries for more robust data and search.

All Rights Reserved, @IJAREST-2017

VI. ACKNOWLEDGEMENT

I would like to thank all the individuals who played defining role in shaping this paper. Without their constant Coordination, guidance and reviewing, this task could not be completed alone. I avail this opportunity to express my deep sense of gratitude and whole hearted thanks to my guide Prof. P.L. Ramteke for giving his valuable guidance, inspiration and encouragement to embark this paper

REFERENCES

[1] Bizer, C., Heath, T., & Berners-Lee, T. (2009). Linked data-the story so far. International journal on semantic web and information systems, 5(3), 1-22. Chicago
[2] Berners-Lee, T., Hendler, J., & Lassila, O. (200 I). The semantic web. Scientific american, 284(5), 28-37.
[3] Dietze, S., Yu, H. Q., Giordano, D., Kaldoudi, E., Dovrolis, N., & Taibi, D. (2012, March). 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.
[4] Linked Up Challenge. (n.d.). LinkedUp Challenge. Retrieved July 22, 2014, from http://linkedup-challenge.org/
[5] Pappano, L. (2012). The Year of the MOOe. The New York Times, 2(12),2012. Chicago
[6] Coursera Catalog API. (n.d.). - Cow'sera Technology. Retrieved July 22, 2014, from https:lltech.coursera.org/app-platform/catalog/
[7] Klyne, G., & Carroll, J. J. (2006). Resource description framework (RDF): Concepts and abstract syntax

All Rights Reserved, @IJAREST-2017