

Practical uses of location and event data as Linked Open University Data

Barnabás Szász^a, Rita Fleiner^b, András Micsik^c

^aUniversity of Debrecen
bszasz@gmail.com

^bÓbuda University
fleiner.rita@nik.uni-obuda.hu

^cMTA SZTAKI
micsik@sztaki.mta.hu

Abstract

Linked Data is a practical application of the Semantic Web technologies for connecting data worldwide. The Open Data pursuit has achieved remarkable progress in Europe as well, and studies have shown that it has a positive impact on the quality of education at university level too. Publishing information about university or college course, their corresponding places and related events, such as exams in Linked Data format allows the event information to be aggregated, filtered and delivered to potential participants: students and lecturers via multiple channels and devices. In the previously published version of our Linked Open University Data model organisational hierarchy, staff and publication data were modelled as Linked Data. With this model one can describe faculties, professors, their teaching and research activities, etc. as Linked Data giving a skeleton of human organisation of a university. In this paper we extend the model with temporal and location data including course schedule, event data and indoor location data. To this end we present the state of the current semantic representation formats for course data, event data and indoor location data and describe our approach for creating a unified schema for Linked Open University Data. We have reviewed several overlapping RDF schema and vocabulary, their potential combinations and usability in practice. Finally a formal representation of the model is presented as an OWL 2 ontology. We also provide some use cases where this machine readable data demonstrates its benefits, as well as some RDF data excerpt based on our model and some example SPARQL query

illustrating the possible use of the data.

Keywords: Linked Data, Open Data, Semantic Web, Linked Universities

1. Introduction

Since the definition of Linked Open Data (LOD) in 2006, LOD datasets are appearing all around the world rapidly, expanding the way that we access data. Different institutes and data owners are figuring out new ways to export their data in LOD triples, integrate external linked datasets into their collections, and develop new interfaces. In our previous work [1] we described five types of data within the framework of our Linked Open University Data model: Person, Publication, Courses, Organization and Role. With this model one can describe faculties, professors, their research outcomes and course data. The aim of the current paper is to extend this model with temporal and location data including course schedule, event data and indoor location data and to describe our approach for creating a unified schema for Linked Open University Data. The concept of LOD was defined already in many papers (e.g., [2] [3]), in section 2 we provide a short review of it and it's benefit for universities. Section 3 provides an overview of the extended Linked Open University Data model, the used ontologies and RDF schemas. In section 4 we present the additional new classes for describing events with their temporal and spatial information and indoor location data and in section 5 we show some example data and SPARQL queries illustrating the possible use of the data.

2. Linked Open University Data

Linked Data isn't just about putting data on the web. It is about providing links between entities, so that the web of data can be explored. It comes with the promise that one can find related data following the links provided with the data. The standard data model for Linked Open Data is the Resource Description Framework (RDF). In RDF data is structured in triples in the form of subject, predicate and object, which is called a statement. The predicate specifies how the subject and object are related. The subject and the predicate are both URIs and the object is a URI or a string literal. SPARQL is an RDF query language, designed to retrieve and manipulate data stored in RDF format. The Linked Data concept - invented by Tim Berners-Lee in 2006 - is based on the following four principles [4]:

1. Use URIs as names for things
2. Use HTTP URIs, so that people can look up those names
3. When someone looks up a URI, provide useful information, using the standards (RDF, SPARQL)
4. Include links to other URIs, so that they can discover more things

The 4th recommendation ensures the links within the different datasets and provides options for schema and vocabulary reuse too. A crucial concept in Linked Open Data is ontology. Ontology formally represents the set of concepts within a domain, it describes the classes of the entities, their properties and the relationships among the classes. A similar concept to ontology is vocabulary. Vocabularies define the concepts, their relationships and constraints that are used to describe and represent an area of concern. In order to create a well-designed Linked Open Dataset, its data model has to be specified, which is composed of the following main tasks. First the available task specific vocabularies and ontologies have to be chosen, they will be specified under the used namespaces. Then the used classes and their relationship have to be specified and finally the possible properties of the entities for each class have to be collected. If it is possible, the property should be a member of an existing vocabulary or ontology. If a necessary property does not exist yet, then it should be created for the given data model. For each property, its meaning, its URI, its domain and range have to be specified.

The use of Linked Data is growing and appearing in more and more domains, including the use cases of education, specifically within the domain of universities. Constructing Linked Data requires well designed vocabularies, schemas or even ontologies, this was also our experience during the selection process of reusable schemas and ontologies. See section 3 for the selected ontologies and RDF schemas. Linked Universities [5] and Linked Education [6] are two European initiatives to enable education with the power of Linked Data. The Óbuda University established a project for publishing university related data as Linked Open University Data, thus joining to the Linked Universities alliance. The expected results include higher quality education as well as improved satisfaction among students and lecturers. The University of Southampton already published more than 40 linked datasets and there are 18 applications built on top of this open data. The benefits also include easy data aggregation opportunities; RDF enables third-party datasets, like bus schedules or geo information to integrate with university related data: one application aligns bus schedule with course schedules to provide a tool for students. To develop applications consuming LOD one needs to make sure that the structure of the data is well designed, clearly defined and published. The Semantic Web provides a multi-layer method for formalization of conceptualization. RDF Schema on the first level provides tools to describe classes of entities and their relationships. OWL (Web Ontology Language) builds on top of RDFS and provides tools to formalize ontological relationships and constraints. While RDFS is essential for LOD development, OWL adds further reasoning capabilities on top of the dataset and thus it is preferred.

3. The Linked Open University Ontology

In the process of creating the schema for the Linked Open Dataset it is advisable to reuse as much as possible of the available ontologies or vocabularies. In the first step suitable vocabularies or ontologies have to be looked for in order to reuse

them. If there is no suitable vocabulary term for a specific purpose, then it should be created from scratch. The following ontologies and RDF schemas were chosen to be reused and extended in our model. The AIISO¹, AIISO-Roles², Participation³ and Teach Ontologies⁴ provide a schema to describe academic concepts, like lectures and lecturers, academic roles and participation in these roles. FOAF (Friend of a Friend)⁵ provides a basic vocabulary to describe personal entities and their attributes. The purpose of Dublin Core⁶ is to be a common vocabulary describing documents and their attributes. The Geo Ontology's⁷ role is the ability to

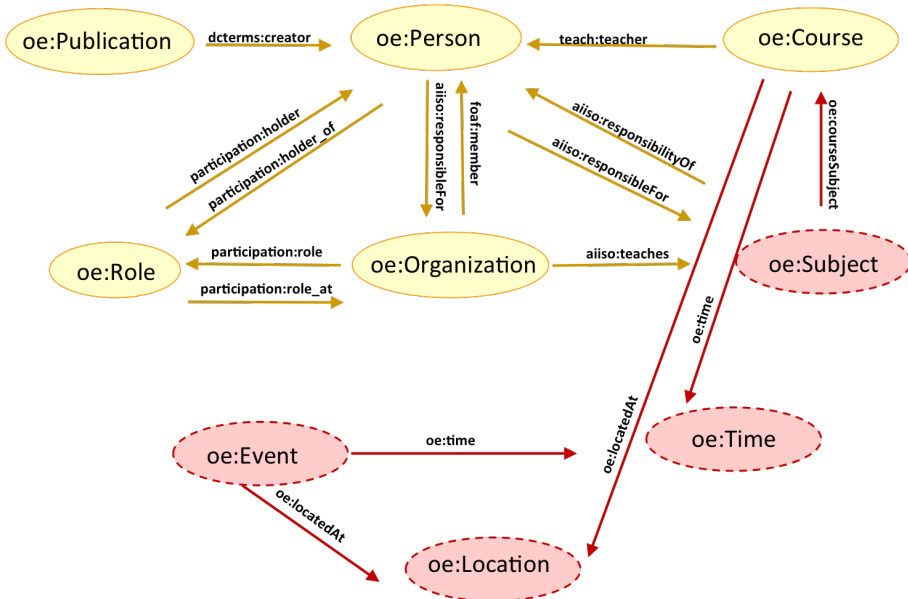


Figure 1: Extended Linked Open University Data model

geo-code our Place entities, while vcard⁸ provides attributes to describe addresses. The Event Ontology⁹ is a high level, simple ontology for describing events in the temporal and spatial space. The Time Ontology¹⁰ provides high level concepts to formally describe time instants and intervals and their relations, like overlap-

¹<http://purl.org/vocab/aiiso/schema-20080925.rdf#>

²<http://purl.org/vocab/aiiso-roles/schema#>

³<http://purl.org/vocab/participation/schema#>

⁴<http://linkedscience.org/teach/ns#>

⁵<http://xmlns.com/foaf/0.1/>

⁶<http://purl.org/dc/terms/>

⁷http://www.w3.org/2003/01/geo/wgs84_pos#

⁸<http://www.w3.org/2006/vcard/ns#>

⁹<http://purl.org/NET/c4dm/event.owl#>

¹⁰<http://www.w3.org/2006/time#>

ping events or chronological order. The Temporal Aggregates extension¹¹ adds the ability to define recurring events. The location and event description ability is an extension of our previous work; the corresponding classes are marked with red and dashed outline on Figure 1.

4. Ontologies selected for re-use

Indoor navigation is a very important research topic today. Potential use cases could include augmented reality (AR) and mapping for the visualization of the navigation [7]. Ontology can be used to describe the indoor space and the potential routes within the space. Multiple ontologies were realized for indoor navigation: OntoNav is a semantic indoor navigation system and an ontological framework of handling routing requests [8]. ONALIN [9] provides routing for individuals with various needs and preferences, it takes the ADA (American disability act) standards, among other requirements, into consideration. As neither of these location ontologies were available online at the time of our work, parts of ONALIN were implemented as the location module of the ontology. The objective of modelling

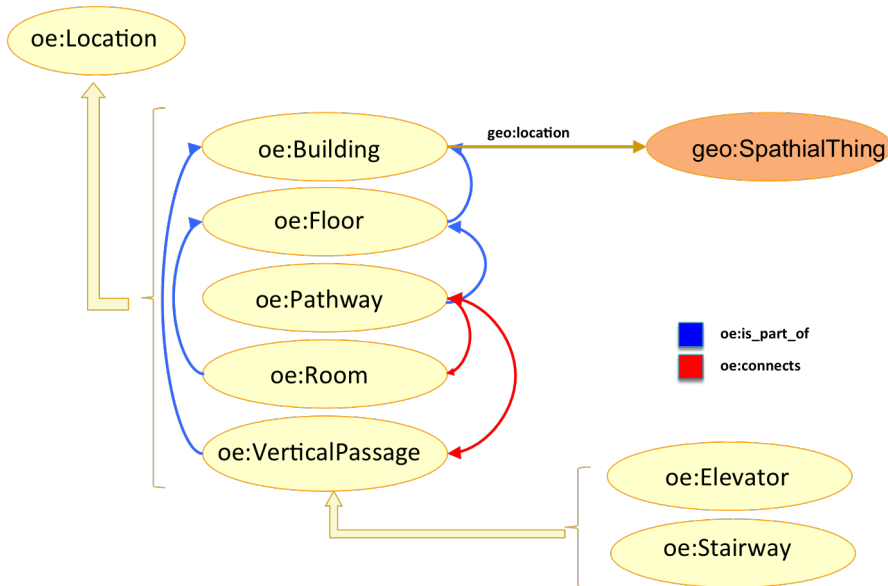


Figure 2: Concepts for specifying locations in the ontology

indoor locations is to provide location description for events and to help the navigation to there. The focus was on modelling the relevant concepts: Buildings,

¹¹<http://ontology.ihmc.us/temporalAggregates.owl#>

Rooms, Floors within a building, Hallways, Stairways and Elevators. Each Building has a number of Floors and Floors contain Rooms connected by each other or by Hallways. The transitive property used to describe this part-of relationship is *isPartOf*. Vertical connections are provided by Stairways and Elevators. If there is a walkable connection between any Room, Hallway or Vertical Passage, it is modelled with the transitive and symmetric *connects* property. To describe location of buildings, one can either define the address with the vcard vocabulary or the geo-location with the Geo Ontology.

We have requirements to describe temporal entities: events and assignments, repeating lectures and their durations. For this purpose one can choose from multiple ontologies developed for modelling temporal entities. Since our aim is to provide an OWL 2 DL compliant model, we were only considering those with such a representation. The candidates were the Event Ontology, schema.org, ical, Time Ontology, Temporal Aggregates Ontology and the Timeline Ontology. Finally, we have chosen the W3C time ontology, which is general enough to formulate a wide range of temporal descriptions, with the temporal aggregates extension we can describe a series of repeating events, such as every Tuesday from 10 to 12.

5. Usage example

The following excerpt describes a course on Databases, with its course code, lecturer, location and time schedule. Further details of the course are given as a Subject entity. The location of the course is described as being on the second floor of the main building, with address also given. There is also possibility to link buildings to GeoNames and exact geocoordinates.

```
:AB0_LA_05
  rdf:type          oe:Course ;
  oe:courseCode    "AB0_LA_05"@hu ;
  oe:courseName    "Adatbázisok"@hu ;
  oe:courseSubject :NAIAB0SAND ;
  oe:courseLocation :PC_Labor_210 ;
  oe:courseTeacher :Rita_Fleiner ;
  oe:courseType    :lab ;
  oe:courseTime    :example_lecture_everyTue_10am ;
  oe:courseTerm    :semester-2014-15-1 .

:Rita_Fleiner
  rdf:type          oe:Person ;
  foaf:familyName  "Fleiner" ;
  foaf:firstName   "Rita" ;
  foaf:mbox        <mailto:fleiner.rita@nik.uni-obuda.hu>
  .

:AdatbázisRole1
```

```
    rdf:type          oe:Role ;
    part:holder       :Rita_Fleiner .

: AII
  rdf:type          oe:Organization ;
  foaf:name         "Alkalmazott Informatikai Intézet"@hu ;
  foaf:member       :Rita_Fleiner .

: AddressOE
  rdf:type          vcard:Work ;
  vcard:postal-code "1034" ;
  vcard:street-address "Bécsi út 96/B" ;
  vcard:hasLocality  :Budapest ;
  vcard:hasCountryName :Hungary .

: NAIAB0SAND
  rdf:type          oe:Subject ;
  oe:subjectTotalHours "28e-28gy"@hu ;
  oe:subjectName      "Adatbázisok"@hu ;
  oe:subjectCode      "NAIAB0SAND" ;
  oe:subjectWeekHours "2e-2gy"@hu ;
  oe:subjectCredit    "4" ;
  oe:subjectDescription "Kezdő Adatbázis-kezelés. SQL nyelv"@hu ;
  foaf:homepage       "http://nik.uni-obuda.hu/to/tanterv/1097" ;
  oe:subjectOrganization :AII ;
  oe:subjectMajor      :ComputerEngineer ;
  oe:subjectTime       :FullTime ;
  oe:subjectLanguage   :Hungarian ;
  oe:subjectResponsibilityOf :Rita_Fleiner ;
  oe:subjectDegree     :bsc ;
  oe:subjectTerm       :semester-2014-15-1 .

PC_Labor_210
  rdf:type          loc:Room ;
  loc:is_part_of    :second_floor_-_bmf_building .

: second_floor_-_bmf_building
  rdf:type          loc:Floor ;
  loc:is_part_of    :BMF_Main_Building .

: BMF_Main_Building
  rdf:type          loc:Building ;
```

```
rdfs:label          "BMF Main Building"@en ;
vcard:hasAddress    :AddressOE.
```

Listing 1: Courses with time and location from the first semester of 2014/15

```
PREFIX oe:<http://lod.nik.uni-obuda.hu/OLOUD.owl/#>
PREFIX oedata:<http://lod.nik.uni-obuda.hu/data/>
SELECT ?course ?location ?time
WHERE {
  ?course rdf:type oe:Course .
  ?course oe:courseTerm oedata:semester -2014-15-1 .
  OPTIONAL { ?course oe:locatedAt ?location }
  OPTIONAL { ?course oe:courseTime ?time }
}
```

Listing 2: Lecturers and their courses

```
PREFIX oe:<http://lod.nik.uni-obuda.hu/OLOUD.owl/#>
SELECT ?lecturer ?course
WHERE {
  ?course oe:courseTeacher ?lecturer .
  ?course rdf:type oe:Course .
  ?lecturer rdf:type oe:Person .
}
```

Listing 1. and 2. show example SPARQL queries to demonstrate some of the use cases and the simplicity of querying data. The first query generates a list of courses and their location and timing matching to the criteria, while the second one list lecturers and their courses.

6. Summary

We have described the role and importance of Linked Open Data in the educational domain. With the presented ontology we intend to establish a generic descriptive framework for university linked data, which can be reused in various scenarios. The problem of designing such an ontology is to keep its descriptive power yet to re-use widely used vocabularies and thus to stay easily understandable for new adopters. Existing vocabularies or ontologies were often too simple for our purposes and furthermore we found that their RDF formalisations were sometimes restricting their re-use in a more generic setting. It was shown that the ontology is capable for describing the main modeled entities of the university world (people, courses,

locations, events) and their various relations. In this way a timetable can be formalized in RDF which is much richer than the usual ones as it may contain the links to further information about lecturers, lecture halls and course material as well. Our future plan is to fill the ontology with sample instances for some faculties, and offer it as a starting point for online services and mobile applications helping students and lecturers. This work is already started in the form of some master theses. We expect that using our university linked data model, new services will appear with a wide range of support for university life ranging from help in navigation and timetables to the discovery of research information such as new topics and their key persons at the university.

References

- [1] FLEINER, R., MICSIK, A. Linked Open Data az egyetemen (in Hungarian), *Informatika a felsőoktatásban 2014. Debrecen, Magyarország, 2014.08.27-2014.08.29*. pp. 872-879.
- [2] B. HYLAND, D. WOOD The joy of data - a cookbook for publishing linked government data on the web, *Linking Government Data, Springer New York*, 2011., p. 3-26.
- [3] B. HYLAND, G. ATEMEZING, B. VILLAZÓN-TERRAZAS Best Practices for Publishing Linked Data, *W3C Working Group Note 09 January 2014* <http://www.w3.org/TR/ld-bp/>
- [4] T. BERNERS-LEE Linked data-design issues, 2006, <http://www.w3.org/DesignIssues/LinkedData.html>
- [5] Linked Universities, <http://linkeduniversities.org/>
- [6] STEFAN DIETZE, HONG QING YU, DANIELA GIORDANO, ELENI KALDOUDI, NIKOLAS DOVROLIS AND DAVIDE TAIBI Linked Education: interlinking educational Resources and the Web of Data, *ACM Symposium On Applied Computing (SAC-2012) Special Track on Semantic Web and Applications*, <http://oro.open.ac.uk/id/eprint/31077>
- [7] MATUSZKA T., GOMBOS G., KISS A. A New Approach for Indoor Navigation Using Semantic Webtechnologies and Augmented Reality, *15th International Conference on Human-Computer Interaction, VAMR/HCI 2013, LNCS 8021 Part I*, pp. 202-210. Springer, DOI: 10.1007/978-3-642-39405-8_24
- [8] V. TESTSOS, P. KIKIRAS, S. P. HADJIEFTHUMIADES AND C. ANAGNOSTOPOULOS, OntoNav: A Semantic Indoor Navigation System, *Proc. Managing Context Information in Mobile and Pervasive Environments*, vol. 165, (2003).
- [9] PATRICK M. DUDAS, MAHSA GHAFOURIAN, AND HASSAN A. KARIMI ONALIN: Ontology and Algorithm for Indoor Routing, *In Proceedings of the 2009 Tenth International Conference on Mobile Data Management: Systems, Services and Middleware (MDM '09)*, IEEE Computer Society, Washington, DC, USA, 720-725. DOI=10.1109/MDM.2009.123