

SEEU Study Programs Curricula in Linked Open Data

FloriJe Ismaili, Kaltrina Ukalli, Xhemal Zenuni, Bujar Raufi, Jaumin Ajdari

Abstract— WWW is moving from a Web of hyper-linked documents to a Web of linked data. Linked Open Data goes beyond borderlines by linking all the freely published RDF data in freely accessible datasets connected with each other to form the so called LOD cloud. This paper is focused on the evolution of the Linked Open Data platform of the SEE University curricula, from a research experiment to a data hub for the open content of the university. Based on Linked Data principles, loud.seeu.edu.mk is used to publish and access data about academic programs and courses offered in the university. It exposes as well a SPARQL endpoint in order to provide recommendations of resources based on user queries.

Keywords— Semantic Web, Linked Data, Linked Open Data, Ontology, Study Program Curricula, Protégé.

I. Introduction

Nowadays we are witnessing the growth of information requirements. Living in the era of technology, facilitates the possibilities of getting the desired responses. The web has become the main source of information and its access is increasing the amount of data day by day. Therefore, Linked Open Data has become one of the driving forces for the emerging Semantic Web. The main aim of this paper is to describe SEEU study programs curricula Linked Data set which will be online available to be seen and understood equally by humans and machines, contributing in this manner mainly in the Educational Domain.

FloriJe Ismaili
South East European University
Macedonia

Kaltrina Ukalli
State University of Tetovo
Macedonia

Xhemal Zenuni
South East European University
Macedonia

Bujar Raufi
South East European University
Macedonia

Jaumin Ajdari
South East European University
Macedonia

The focus will be on the evolution of the Linked Open Data platform of the SEE University curricula, from a research experiment to a data hub for the open content of the university.

As starting point, the ontology is built using Protégé [1]. This ontology stands as a guideline of how the study programs are organized into respective faculties, their location, what courses each semester offers, a brief description of each course and the corresponding number of ECTS.

In the next step, loud.seeu.edu.mk publication of open data is described:

- Converting University Data to structured data in a standard, non-proprietary format - RDF triples, using URIs to denote things which links to other datasets.
- Exposing a SPARQL endpoint in order to provide recommendations of resources based on user queries.

This paper is organized as follows: the first section shows the aim of this research, the second section depicts the problem statement, while the third section explains our work and the last section stands for the conclusion and future work.

II. Problem Statement

Building an ontology for the SEEU study programs curricula as well as publishing them to LOD is the motivation of this research. To be more specific, an ontology which includes the undergraduate study program curricula of the Faculty of CST is built, soon after to be enhanced with the curricula of the other two cycles study programs. CST undergraduate studies curricula includes all the study programs, with all the list of courses divided by semesters (from semester 1 to semester 6). Also, a description for each course is provided and the number of ECTS each course is characterized.

The current ontology will enable to review in general or specific through faculties and the study programs such as:

- Which study program belongs to each faculty,
- Which study programs belongs to each cycle,
- Distinction between master programs by their total ECTS,
- Courses of each study program divided by semesters as well as branches and sub-branches of possible elective courses,
- Description of courses and number of ECTS that each course is possessing.

A. Analyzing The SEEU Study Programs Curricula

Even SEEU website is well structured and understandable the Semantic Web to the SEEU website is the missing part of the puzzle.

The University data is organized hierarchically, starting from the five faculties that the SEEU has. Each faculty has its sub-branches, among others including: the management, the academic staff and the study programs.

The hierarchy of faculties continues with the cycle of studies, where each cycle includes its study programs. Furthermore, each study program is comprised of: Description, Learning Outcomes and List of Courses.

This work concentrates more on the last subdivision from above, which contains all the semesters of a Study Program, where each one is including its courses. In addition, each course has its description and the number of ECTS, as well. The analysis of the list of courses in the undergraduate studies showed that there is another subdivision among some courses. These courses are a group of elective courses, whether they are: courses from the field of study, languages (English, Albanian or Macedonian) in different levels, Academic Writing in these languages, and Communication Skills with its sub-branches or IT Skills and its sub-branches.

III. Building the Ontology

This section represents the methodology that was used to build the ontology and ontology itself. For developing university ontology, Protégé tool is used since it is the mainly accepted and broadly used tool for ontology development. Except the class hierarchy, object properties, data properties, and individuals, Protégé possesses different plug-ins, which enrich it. OWLViz [9] and OntoGraf [10] plug-ins are used in order to visualize the university ontology.

During the ontology development, the following major activities are taken [12]:

- **Planning:** what should be done and what resources are needed,
- **Conceptualization:** structuring the domain knowledge such as identifying existing concepts in domain and the relation between them,
- **Implementation:** this step involves the ontology implementation,
- **Evaluation:** comparing ontology performance by putting under several revisions,
- **Documentation and Maintenance:** documenting the ontology in order to facilitate sharing and reuse as well as keeping the ontology up to date.

A. Building and Visualizing the Classes of the Ontology

After analyzing the SEEU Study Programs Curricula the core requirements for the ontology were determined.

The main motivation of developing a new ontology born during the first steps of implementation: In this domain there are existing ontologies and our initial approach was to build a linked Open University dataset using these existing ones. But it resulted that the existing ontologies are not extensible. Therefore some of our requirements were not addressed at all by any of these existing ontologies.

During the development of the ontology the guidelines described in [12] were taken into consideration. The scope of the ontology is to support publishing of SEEU Linked Open Dataset, in addition, the applications built on top of this data might include applications for supplying lecturers, students and interested users with course information.

The developed ontology comes with a root level class: Thing, which is the parent class of all other SEEU LOD classes.

The main classes in this ontology are: Class Person and Class University as it can be seen in the figure 1.

The Class Person is a super-class of the classes: Student and Visitor. This class is meant for instances of students and simple visitors, which specifies the connection between people and university.

The most important class in the developed ontology is the University class, which includes the five subclasses illustrated in figure 1.

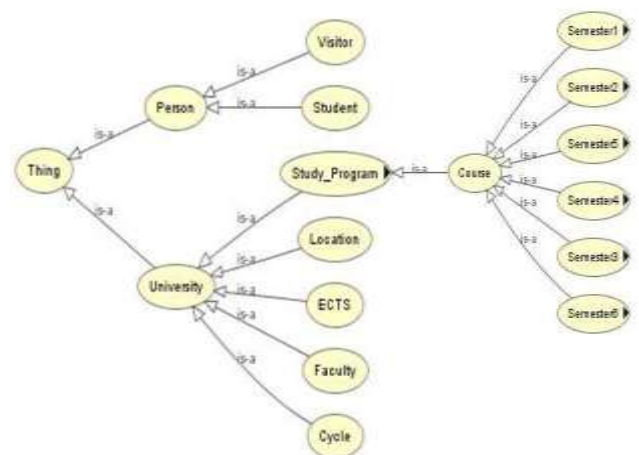


Figure1. Representation of Class Person and Class University

The class Faculty stands for the five faculties of SEEU, which are added as individuals: Contemporary Sciences and Technologies - CST, LCC¹, LAW², PAPS³ and BE⁴.

The class Location stands for the Location of the Study Programs and the role of this class is to hold the individuals: Tetovo and Skopje, which represent the cities and help to contrast each study program by its location.

¹ LCC- Faculty of Languages, Cultures and Communication;

² LAW - Faculty of Law;

³ PAPS - Faculty of Public Administration and Political Sciences;

⁴ BE - Faculty of Business and Economics.

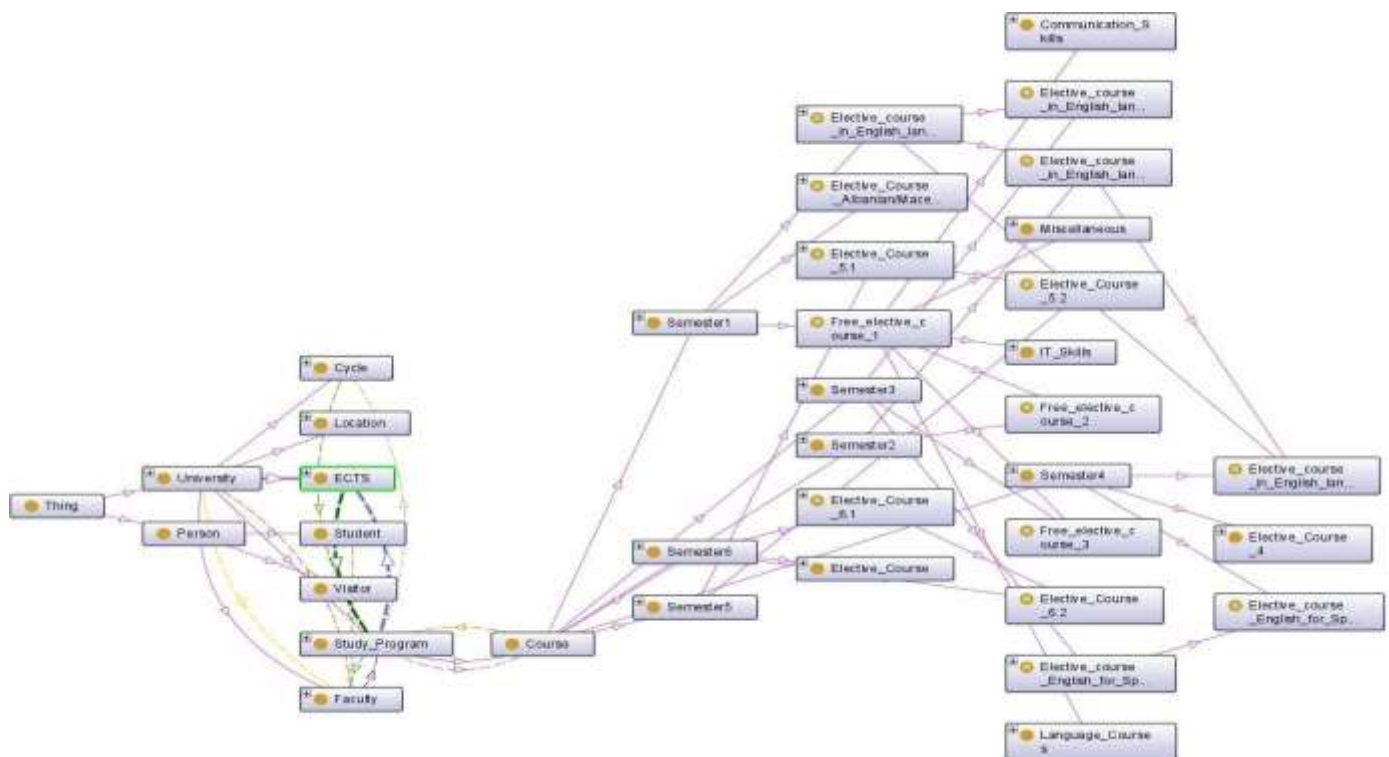


Figure 2. Representation of SEE University ontology

The class Cycle stands for the cycle of the studies by classifying the study programs into the input individuals: Undergraduate, Postgraduate or PhD studies.

The class ECTS stands for differing the postgraduate cycle of studies through the total number of ECTS. For instance, there are master study programs that have 60 ECTS or 120 ECTS.

Study_Program is a subclass of University, meanwhile the super class of the class Course, where all the study programs are listed as individuals. The course, as the most important part of each study program in real life, is of a major role synonymously in our ontology. This class consists of six other subclasses which represent the six semesters and all the courses for each study program divided by semesters. In all semesters students attend core courses and elective courses that are the same for all the SEEU study programs in the first cycle of studies. This is the reason why the elective courses are added as subclasses of the corresponding class Semester. Figure 2 represents the OntoGraph Visualization of the ontology

B. Object and Data Properties of the Ontology

Object properties, also called predicates, serve as a means to link the individuals with each other. They are characterized by domain and range, for use in reasoning.

While object properties are used in order to make a link between individuals in different classes, the role of data properties is to link the individuals with literals [13].

This ontology consists of these object properties:

- studyProgramOf,
- hasStudyProgram,
- offersMasterProgram,
- attendsStudyProgram, courseOf,
- hasCourse,
- composedOf,
- hasECTS,
- isCycle,
- isOfferedIn,
- offers,
- part of,
- thisCycleIncludes.

The data properties used in this ontology are two:

- courseECTS: adds the respective ECTS to each course
- courseDescription: gives a specific description to each course.

A very important issue to mention is the integration with other existing ontologies. In the process of creating the schema for the SEEU Linked Open curricula dataset, it was advisable to reuse as much as possible of the available ontologies or vocabularies [13]. We found that there are

relatively different vocabulary reuse strategies [14]. The two basic forms are

- reusing classes and properties from existing vocabularies directly,
- establishing links on schema-level, which means defining new classes as either subclasses or equivalent classes and properties as subproperties or equivalent properties of the classes and properties of the reused ontology.

In case of SEEU ontology, the followed integration strategy was: identifying the necessary concepts and then the list of vocabularies was chosen from the best ones that could serve to express the defined concepts

C. Querying and Evaluating the Ontology

The role of the evaluation is to verify the fulfillment of the initial goals. In Protégé there are two query systems built as plug-ins, DL Query and SPARQL Query. These serve as testing and evaluation method with the purpose of concluding while the ontology is fulfilling the needs. In this work, DL Query is used since it is sufficient in giving the exact answers to fulfill the requirements of this paper.

Figure 3 shows the case, where it is written a query that seeks as a result all the study programs that are offered in Skopje, whose result lists all the possible courses (22 instances) of the study program Computer Sciences that have 6ECTS in the first semester.

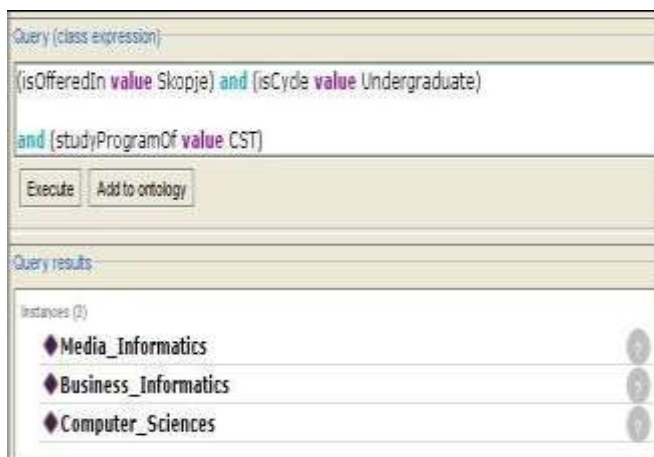


Figure 3. Query result of CST undergraduate study programs in Skopje

In order to evaluate and validate the ontology LOD triplets were created continuously. This process is done with Java programming through a Jena project in Eclipse [15]. Now, when we have our triples the following step is to see how they look like through the aforementioned RDF serialization formats in order to be more readable. If we take into account the N-Triples serialization format, we can see below the five sets of triples of the study program E-Technologies:

```
<http://www.loud.seeu.edu.mk/curricula#E-Technologies>
<http://www.loud.seeu.edu.mk/curricula#isOfferedIn>
<http://www.loud.seeu.edu.mk/curricula#Tetovo>
```

```
<http://www.loud.seeu.edu.mk/curricula#E-Technologies>
<http://www.loud.seeu.edu.mk/curricula#isCycle>
<http://www.loud.seeu.edu.mk/curricula#PhD>
```

```
<http://www.loud.seeu.edu.mk/curricula#E-Technologies>
<http://www.loud.seeu.edu.mk/curricula#isOfferedIn>
<http://www.loud.seeu.edu.mk/curricula#Tetovo>
```

```
<http://www.loud.seeu.edu.mk/curricula#E-Technologies>
<http://www.loud.seeu.edu.mk/curricula#isCycle>
<http://www.loud.seeu.edu.mk/curricula#PhD>
```

IV. SPARQL Endpoint Service

The last and final step of the project was building the SPARQL endpoint illustrated in Figure 4, in order to provide recommendations of resources based on user queries.



Figure 4. SEEU SPARQL ENDPOINT

The following is an example of SPARQL query to select the number of instances of the class semester1:

```
SELECT ?semester1 (COUNT(?s) AS ?count )
{ ?s a ?semester1 }
GROUP BY ?semester1
ORDER BY ?count
```

v. Conclusions and Future Work

In this paper we showed the process of developing the SEEU LOD ontology during the whole ontology life cycle, including: knowledge acquisition, edition, ontological mappings, integration and evaluation. The developed ontology provides classes and properties to describe public data generally available in the educational domain, to finally publish the results embedded in SPARQL endpoint. It is obvious that not all problems could be discussed in detail and there are open issues to solve in future work.

Besides, this work can serve to promote research innovations and service development in education through semantically enabled technologies.

Reference

- [1] Protege[®], Stanford University [Online], Available: <http://protege.stanford.edu/>. [Accessed 10 July 2015].
- [2] M. Shahzad, "A Semantic Web approach for dealing with University courses," University of Agder, Faculty of Engineering and Science, Grimstad, Norway, 2013.
- [3] J. Cardoso, "The Semantic Web Vision: Where are We?," IEEE Intelligent Systems, vol. 22, no. 5, pp. 84 - 88, September - October 2007.
- [4] N. S. Bawany and N. Nouman, "A Step towards Better Understanding and Development of University Ontology in Education Domain," Research Journal of Recent Sciences , vol. 2, no. 10, pp. 57-60, October - 2013.
- [5] N. Malviya, N. Mishra and S. Sahu, "Developing University Ontology using protégé OWL Tool: Process and Reasoning," International Journal of Scientific & Engineering Research, vol. 2, no. 9, September-2011.
- [6] S. A. Naser, R. R. Atallah and S. Hamo, "Building an Ontology in Educational Domain Case Study for the University of Palestine," International Journal of Research in Engineering and Science, vol. 3, no. 1, pp. 15-21, 2015.
- [7] S. K. Malik, N. Prakash and S. Rizvi, "Developing an University Ontology in Education Domain using Protégé for Semantic Web," International Journal of Science and Technology, vol. 2, no. 9, pp. 4673-4681, 2010.
- [8] M. Jhaveri, J. Pareek and J. Jha, "Assemblage of Recommendations with Constraints: A Choice Based Credit System Perspective," in Technology for Education (T4E), 2013 IEEE Fifth International Conference on, Kharagpur, 18-20 Dec. 2013.
- [9] M. Horridge, "OWLviz," Protege, 23 July 2013. [Online]. Available: <http://protegewiki.stanford.edu/wiki/OWLviz>. [Accessed 15 November 2015].
- [10] Falconer, "OntoGraf," Protege, 18 January 2013. [Online]. Available: <http://protegewiki.stanford.edu/wiki/OntoGraf>. [Accessed 16 November 2015].
- [11] K. Hadjar and N. Channane, "University Ontology: case study Ahlia University," in International Conference on Semantic Web & Web Services, Las Vegas, 2013.
- [12] Janowicz, K., Hitzler, P., Adams, B., Kolas, D., & Vardeman II, C. (2014). Five stars of Linked Data vocabulary use. Semantic Web, 5(3),173-176.
- [13] Correndo, Gianluca, et al. "Linked timelines: Time representation and management in linked data." First International Workshop on Consuming Linked Data (COLD 2010), Shanghai, China. 2010
- [14] Schaible, J., Gottron, T., and Scherp, A.. "Survey on common strategies of vocabulary reuse in linked open data modeling." The Semantic Web: Trends and Challenges. Springer International Publishing,2014.457-472.
- [15] The Eclipse Foundation, "Eclipse," Eclipse, 2015. [Online]. Available: <https://eclipse.org/>. [Accessed 27 October 2015].