

AKADEMİK ARAŞTIRMALAR İÇİN ONTOLOJİ TASARIMI

Umut KAYA (umut.kaya@kavram.edu.tr)

Kavram Meslek Yüksekokulu, İstanbul, Türkiye

Zeynep ALTAN (zeynepaltan@beykent.edu.tr)

Beykent Üniversitesi, İstanbul, Türkiye

ÖZET

Anlamsal ağ, İnternet ortamından erişilebilir bilgi ve hizmetlerin daha etkin şekilde işlenebilmesi için bazı soyutlamalar ve modeller sağlar. Anlamsal ağ ile etkin iletişim, somut nesnelere anlamındaki varlıklar kavramını oluşturan ontolojiyi verir. Bu da herhangi bir alana ait kavramsallaştırmanın biçimsel ve açık bir betimlemesidir. Bu çalışmada, akademik araştırmaların sınıflandırıldığı bir ontoloji tasarlanmıştır. Mantıksal kuralların tanımlanmasından sonra, oluşturulan ontolojinin biçimsel kavramsal analizi gerçekleştirilmiştir. Oluşturulan ontolojinin tüm sınıfları, nesne özellik ifadelerine göre birbirleri ile etkileşen örnekler içerir. Bu yapı Protege kullanılarak OWL dilinde oluşturulmuştur. Son olarak, geliştirilen sistemin bir prototipine ait SPARQL sorguları tanımlanmıştır.

Anahtar kelimeler: *Ontoloji, Semantik web, Kavramsallaştırma, Bilgi gösterimi, SPARQL.*

AN ONTOLOGY DESIGN TO REPRESENT ACADEMIC RESEARCHES

Umut KAYA (umut.kaya@kavram.edu.tr)

Kavram Vocational High School, Istanbul, Turkey

Zeynep ALTAN (zeynepaltan@beykent.edu.tr)

Beykent University, Istanbul, Turkey

ABSTRACT

Semantic Web provides models and abstractions to process web-accessible information and services to be more effectively. The effective communication with semantic web is ontology, which is the concept of the entities that means concreted objects. It provides formal and explicit specification of the conceptualization in any domain. In this paper, we designed an ontology in which the academic studies have been classified. After the description of the logic rules, we realized formal conceptual analysis of the constituted ontology. All classes of our ontology include instances interacting with each other according to the object property assertions and it has been built with OWL web ontology language by using Protege. Finally, we defined SPARQL queries of our prototype.

Keywords: *Ontology, Semantic web, Conceptualization, Knowledge representation, SPARQL.*

1. INTRODUCTION

As the semantic web works increase, the applications on ontology will grow more. Since the expandable property is the main concern of any developed ontology, building any ontology is not easy as expected. If domain issues related with all concepts, attributes and samples will correctly be processed, then an elaborated conceptualization is clear to constitute a reusable ontology. The scientific studies for an academic research are inevitable for researchers. The queries on any online library can be done writing the keywords related to the research area. For example, it is possible to display more than 18000 results from IEEE after the search of the keyword ontology between the years 1996 2005 ([http://ieeexplore.org/search/searchresult.jspText=ontology&ranges=1996 2005 Year](http://ieeexplore.org/search/searchresult.jspText=ontology&ranges=1996%2015%20Year)).

The huge number of results causes difficulties to obtain a specific publication or knowledge that researcher wants to access. Although the filters such as date, author and content type have been used with the keywords, the absence of semantic reasoning and semantic search cause waste of time and expose huge amount of data.

Ontology is determined associating with semantic web researches and it is typically used to express and represent knowledge. Using ontology, it is possible to concrete the primitively defined abstract concepts to be a part of the meaning. The requirements which express the ontology and constitute the relationships between data and knowledge entailed the web ontology language OWL. Since XML(Extensible Markup Language) and RDF (Resource Description Framework) did not represent knowledge and semantic reasoning exactly, OWL has been proposed to overcome the semantic problems. SPARQL is one of the query languages developed to search over RDF documents. It is important to know the exact names of classes and properties to advance using the SPARQL requires. Therefore conceptual framework must be inferred in detail.

In this study, we built the draft of an academic research ontology in terms of conceptual modelling. Since the ontology has usually been built cooperatively by a group of people in different locations, they are common constituted with the description of objects, attributes and relations. Ontology in terms of computer science defines the

origin of corresponding set model which includes knowledge domain or discourse[1]. Ontology has been redefined by Borst in his PhD thesis [2] as formal specification of shared conceptualization exploiting the Grubber's ontology definition [3]. We evolved our conceptual model similar to FRBR (Functional Requirements for Bibliographic Records) [4]. FRBR is also a conceptual entity-relationship model used to define bibliographic references; these are recorded and published on the Semantic Web descriptions. We also accompanied the relationships of the constituted ontology as in FRBR structure. To obtain a machine-readable ontology explicitly, it is generally agreed that conceptualization has been required for the abstraction of modelling [5].

In recent years, most of the studies have been essential to elaborate the semantic web studies. We implemented our prototype ontology by demonstrating the lattice structure of the problem and it is able to be used by online libraries. The study organized as follows: in Section 2, we briefly explain the lattice concept before constituting our ontology. Section 3 presents the details of the indicated lattice based ontology. Section 4 includes the design and implementation of the draft academic ontology. The conclusion remarks are given in Section 5.

2. LATTICE USAGE ON ONTOLOGY

It is possible to shape the ontology as the elements of lattice that is a sub discipline of the graph theory, then the ordering relationships of the lattice will represent the ontology. To build an ontology in terms of formal concept analysis, objects and attributes must firstly be determined [6]. Then, the constituted structure should be applied to the graph theory. Let an object is symbolized with O , attribute with A and the binary relation is represented with R , then the relationship R can be define as: $R:O \times A$ and $R \sqsubseteq A$. This means that object property gives us a binary relationship by connecting any attribute with an object. For the general concepts O and A , $O \sqsubseteq A$ as lattice order means that an O has any relationship A . If $B \sqsubseteq O$ and $C \sqsubseteq A$, then (B,C) is defined as lattice properties, where B is any object and C is any attribute. B is also the set of maximum objects of the attributes in the set of C . If a given logic L is the set of all

concepts in any domain, then the lattice over L is defined as partial ordering and is represented with the symbol \sqsubseteq and two dyadic operators \sqcap and \sqcup . Let A and B be any concepts, then the following definitions are the fundamentals a lattice L [7]:

i) If the concept A is true of subset of the conditions in which B is true as $A \sqsubseteq B$, then A is said to be a differentiation of B .

ii) The supremum of A and B is $A \sqcup B$, such that $A \sqsubseteq A \sqcup B$; $B \sqsubseteq A \sqcup B$. If $A \sqsubseteq C$ and $B \sqsubseteq C$, then $A \sqcup B \sqsubseteq C$.

iii) The infimum of A and B is $A \sqcap B$, such that $A \sqcap B \sqsubseteq A$; $A \sqcap B \sqsubseteq B$. If $C \sqsubseteq A$ and $C \sqsubseteq B$, then $C \sqsubseteq A \sqcap B$.

iv) The top of the lattice \top corresponded the universal concept and every pair of elements has a least upper bound as $A \sqcup B$.

v) The bottom of the lattice \perp is a differentiation of every concept and pair of elements has a greatest lower bound as $A \sqcap B$. Complete lattice requires that every subset have a top element and a bottom element in the partial ordering [8]. The top element is typically symbolized with \top and the bottom element with \perp as given in iv) and v). A complete lattice example is given in Figure 2 for the draft ontology and it also supplies all defined properties above. Detailed explanation of the developed ontology is depicted in Section 3.1. These formal definitions identifying lattice given above mean Description Logic-DL [9]. The theories on a lattice can be changed and they can be stated as any version of the logic, especially with the first order predicate. OWL DL provides the developer to obtain the best expressiveness with finite computations of reasoning systems. OWL contains differing species of the restricted construction of the language; this means that the class, individual or property names must be different to each other. OWL DL also agrees with DL searching as a specific part of the first order logic. Briefly, it is clear that DLs are family of knowledge representation language and they are also extendable for ontological design. Ontology that satisfy syntactic conditions is called OWL 2 DL ontology. In the Krotzsch and colleagues' research [10], the constituted direct semantics was compatible with the theoretical semantics model of the SROIQ description logic.

3. DESIGNING THE DRAFT ONTOLOGY

During the development of the ontology for academic researches; We primarily described the classes of our draft ontology and we added object properties to each of the production of classes. Then, we designed the lattice diagram to show relationships between classes and objects. We implemented the W3C standard while constructing our ontology and we utilized Protege platform to build our standard while constructing our ontology, and we utilized Protege 5.1 Beta Platform to build our ontology as RDF/OWL syntax. Finally, we added a few SPARQL queries of the constituted ontology.

3.1 Class Structure of draft Ontology

We firstly determined three disjoint subsets of the class thing as knowledge, person and place, which are described as the main sub classes. The class thing is the top main class of all classes in any ontology that is being accepted by W3C. Later, we constituted our concrete classes to define the universal meanings of the objects and acceptable things for each one. course, program and publication are the sub classes of knowledge class. While developing the ontology, we assumed that the universities generally include three different education types, so that the sub classes of program class in our ontology are bachelorprogram, masterprogram and phdprogram. lecturer and student classes are sub classes of the person class. Figure 1 shows the relationships between all classes and their sub classes of the draft ontology. The class place which is subclass of the class thing is evaluated from the classes city, country, department, library and university. The class knowledge which is also the subclass of the class thing evaluated from the lecturer and student classes.

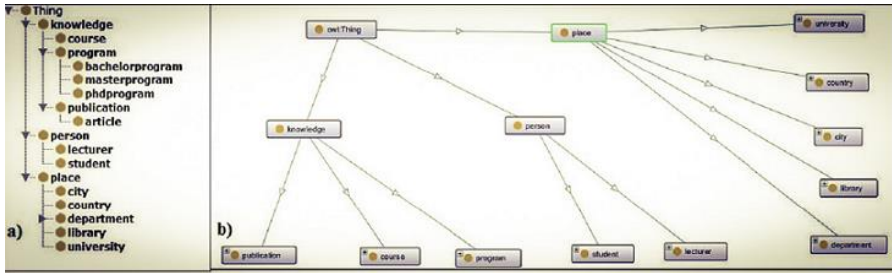


Figure1.Class hierarchies of draft ontology

We attached importance to be universal and extendable while describing the sub classes of the developed ontology. Since the students can continue to their education by registering only one training program, the class knowledge is the super class of all training programs. The extendable training programs are linked to the class program. Since any student training in a university takes different courses and any lecturer giving some courses writes articles, the classes course and publication are also the sub classes of the class knowledge

3.2 Conceptualization of the draft Ontology

As defined in Chapter 3.1, we constituted the class structure of the draft ontology from the class thing which has been determined by W3C, and knowledge, person and place classes are main classes of the class thing. To constitute the lattice configuration of the ontology, we have to define three kinds of entities according to DL description: The first is individual names that represent single individual in the domain; the second is object that represent set of individuals called as concepts, the third is object properties that represent binary relations. In Table 1, the individual names are the attributes that are the columns of the array and objects are the rows of the same array as the classes of Protege representation.

Table 1. An array representation from the draft ontology

	zeynep_altan	computer_science_student	computer_science_lecturer	impact_of_future_selection_For_corpus_based_wsd_in_Turkish
course	X	X		
university				
person	X	X	X	
article			X	X
department				
library				X

The top of the lattice in Figure 2 is class thing as the universal object accepted by W3C, and generally has been used over of all classes. Let T is top of lattice, then $T = \{course, university, person, article, department, library\}$ shows the least upper bound as the union of the objects, in other words, concept in Table 1.

$T \sqsubseteq course \sqcup university \sqcup person \sqcup article \sqcup department \sqcup library$ represents that top concept T is a special concept with every individual takes only one of the objects. The bottom lattice as set: $\perp = \{zeynep_altan, computer_science_student, computer_science_lecturer, impact_of_future_selection_for_corpus_based_wsd_in_Turkish\}$ explains the greatest lower bound as the intersection of the concepts in Table 1. The following concepts allow user to form complex concepts as follows: $course \sqcap university \sqcap person \sqcap article \sqcap department \sqcap library \sqsubseteq \perp$. This is a special statement with no individuals as instances, and the intersection of the concepts is empty.

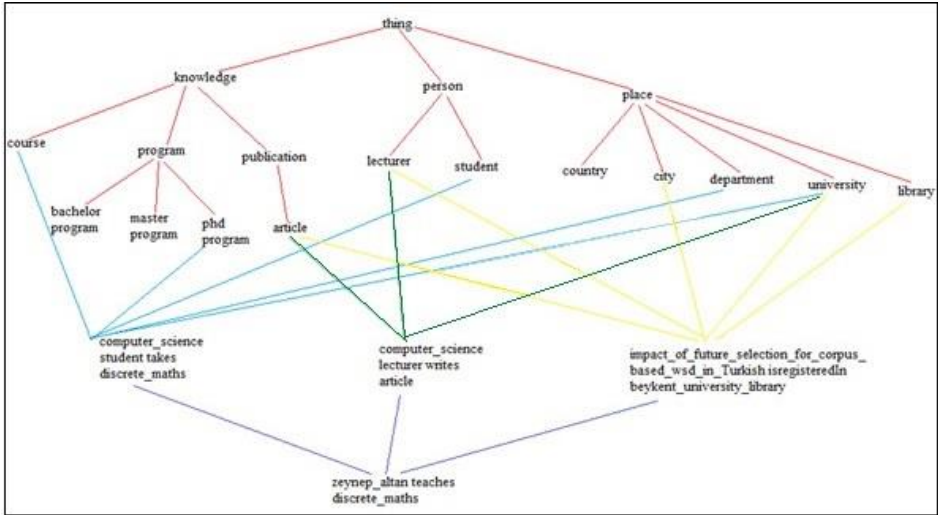


Figure 2. Conceptualization of the draft ontology example as lattice form

3.3 Illustrating the description logic axioms

In Figure 2, there are three relationships for bottom up property constitutions of the presented lattice. It shows the relationships between individuals and sub classes of the class thing as a conceptualization of the given draft ontology sample. One of the previous statements before the bottom in Figure 2 is computer_science_student takes course. In the following, we defined the ABox, TBox and RBox axioms related to DL description according to this statement. It is possible to assert the facts with ABox axioms, to determine the relationships between objects with TBox axioms. We can model the relationships between object properties with RBox axioms[11].

i)ABox samples:

student(computer_science_student),department(course), university(course), phdprogram(course).

ii)TBox samples:

student \sqsubseteq person, university \sqsubseteq place, department \sqsubseteq place, course \sqsubseteq knowledge.

iii)RBox samples:

writes \circ isRegisteredIn \sqsubseteq isMemberOf computer_science_lecturer
writes an article, article is registeredIn beykent university library,
then computer_science_lecturer isMemberOf beykent university is
another RBox example.

isPlacedIn \circ isLocatedIn \sqsubseteq isPartOf,

isTiedTo \circ consistsOf \sqsubseteq isLocatedIn,

isGivenIn \circ isRegisteredIn \sqsubseteq isPlacedIn,...

We can also get bottom inference in Figure 2 step by step. computer_science_lecturer as a lecturer is the subclass of the class person and works at university and computer_science_lecturer is attributed with the class university which is the class place. We use the object property isMemberOf to engage relationship with computer science lecturer and university. The lecturer is both a member of university and writes articles, she or he is attributed with the class knowledge. Other class course is also the subclass of knowledge. The lecture is givenIn by a lecturer at the university. Computer science lecturer with the individual zeynep_altan shows the uniqueness and intersection of all disjoint classes; in other words concepts as knowledge, person and place. Object property relationship between disjoint classes is zeynep_altan teaches discrete maths.

3.4 Specifying the Formal Properties of the Draft Ontology

The data of the draft ontology has been organized using lattice according to its taxonomy. Domain and range names of the ontology are specified by the classes place, knowledge and person. comment and seeAlso at the last column of Table 2 were used as annotated property assertions. They provide to access further information about the individuals; for example for the writer, the article or university. Shortly, these two properties direct the reader to the resource.

Table 2.Conceptual formal properties of draft ontology example

Object properties	Object properties	Data properties	Annotation properties
attends	isregisteredIn	academic_title	comment
consistsOf	istiedTo	Is_written_by	see Also
isgivenIn	takes	name	
islocatedIn	teaches	number	
isMemberOf	writes	surname	
isplacedIn			

Table 3 explains the usage of two different object properties given in Table 2.

Table3. Two Object Property examples

Object Property	Usage
isMemberOf	computer_science_lecturer isMemberOf beykent_university isMemberOf Domain lecturer Object Property: isMemberOf isMemberOf Range university
isregisteredIn	Impact_of_Feature_Selection_for_Corpus-Based_WSD_in_Turkish isregisteredIn beykent_university_library isregisteredIn Domain publication ObjectProperty:isregisteredIn isregisteredIn Range library

To define any computer science student or lecturer, we have to define name, number and surname to obtain the data property assertions (Table 4).

Table 4. Data property example

Data property	Usage
name	computer_science_lecturer name"zeynep"^^string computer_science_student name"umut"^^string DataProperty:name Functional: name name Range: string

The sentence computer_science student attends beykent_university means that beykent is university name, beykent university has computer science department and there are students which attend at beykent university. Table 5 represents the usage of individuals with different examples, which explain the property types as objects. This section gives a brief explanation of the relationships among the object, data and annotation properties and the individuals.

Table 5. Different treatments of an individual

Individual	Usage
beykent_university	beykent_university Type university beykent_universityconsistsOf beykent_university_computer_science_depa rtment Individual: beykent_university beykent_university isLocatedIn Istanbul beykent_university_libraryisPlacedIn

	beykent_university	
	computer_science_lecturerisMemberof	
	beykent_university	
	computer_science_student	attends
	beykent_university	

4. IMPLEMENTATION OF THE DRAFT ONTOLOGY

We created the ontology using Protege with 20 classes, 11 individuals, 14 object properties and 2 annotation properties. Figure 3 shows two different object property assertions between the individuals.

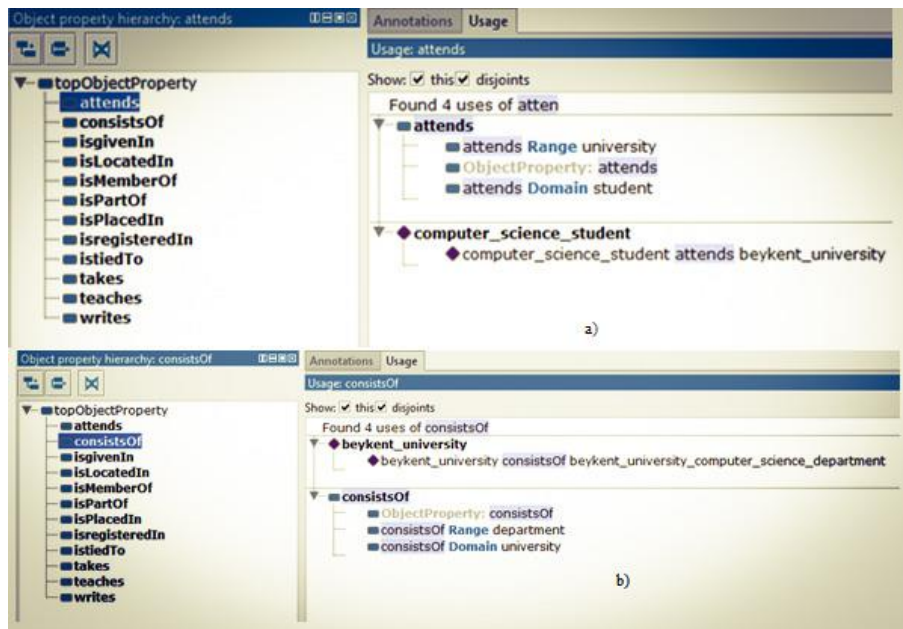


Figure 3. Different object properties of the draft ontology

When the individual has been corresponded with an object property, domain or range usage is specified to clarify the direction of the relationship. attend as an object property assertion will either include a domain from the class student or a range from the class university. In this case, there is a student which attends at Beykent

University. Figure 4 shows data property assertions of the defined ontology classes.

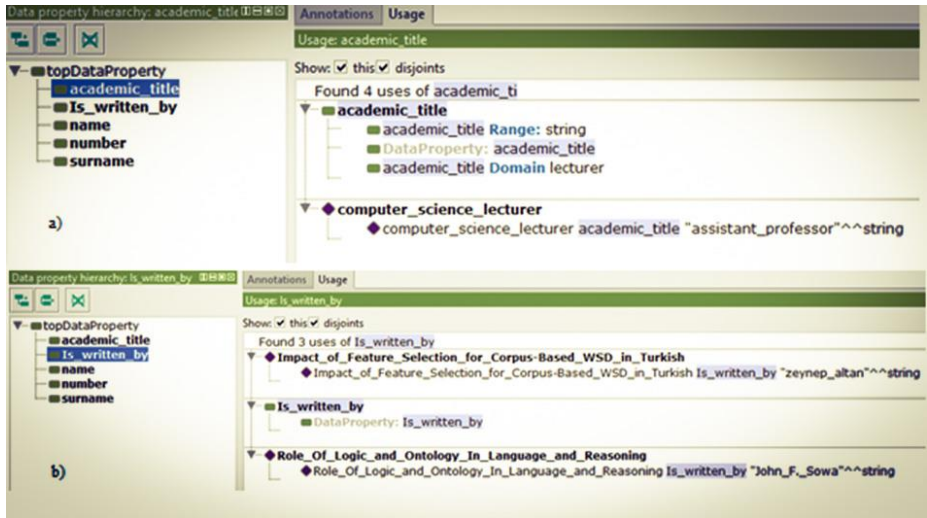


Figure 4. Data properties of the draft ontology

Since each student has a name, number and surname, these data properties can also be extended to all students attending at Beykent University. These are universal data properties that are used to define the relationships between the individuals of disjoint classes as `computer_science_lecturer` and `computer_science_student`. On the other side, Figure 5 explains all information about the class assertion Beykent University with its comment, class assertion, object property domain, object property range and sub class relation which was defined in Chapter 3 detail.



Figure 5. Different treatments of an individual

Figure 6 shows the object, data and annotation properties of the article Impact of Selection for Corpus Based WSD in Turkish as an individual.

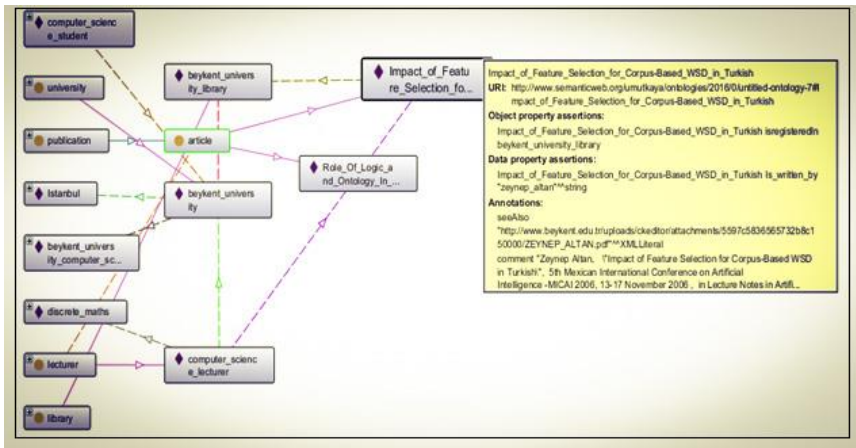


Figure 6. Exact match relations between classes and individuals with an article

Object property assertion is: Impact of Feature Selection for Corpus Based WSD in Turkish+registeredIn+beykent_universitesi_library. Data property assertion is: Impact of Feature Selection for Corpus Based WSD in Turkish Iswritten by Zeynep Altan string. The annotation property usage is: seeAlso: <http://beykent.edu.tr/uploads> and comment: Zeynep Altan. Annotation property provides the information about an article and makes it to find easily. We can also reach any other information about the writer and her other publications. The relationship between machines has been defined using OWL, and SeeAlso provides us to direct the main cite where the article has been registered, It is possible to extend this implementation for all publications at Beykent University Online Library. As we informed in the previous sections, the ontology has been implemented with OWL language using Protege. The following summarizes a small part of the OWL representation of it.

Object <http://www.semanticweb.org/umutkaya/ontologies/2016/0/untitled-ontology-7> Properties:
 attends:attends a owl:ObjectProperty
 rdfs:domain:student rdfs:range:university.

Data properties:
<http://www.semanticweb.org/umutkaya/ontologies/2016/0/untitledontology-7>

number: number number a owl: DatatypeProperty owl: FunctionalProperty rdfs:domain:student rdfs:range xsd:integer.

Classes:

<http://www.semanticweb.org/umutkaya/ontologies/2016/0/untitledontology-7> article:article a owl:Class rdfs:subClassOf:publication.

Individuals:

<http://www.semanticweb.org/umutkaya/ontologies/2016/0/untitledontology-7>

Impact_of_Feature_Selection_for_Corpus

Based_WSD_in_Turkish:Impact_of_Feature_Selection_for_Corpus- Based_WSD_in_Turkish a owl:NamedIndividual: article: isregisteredIn :beykent_university_library: Is_written_by zeynep_altan xsd:string rdfs:comment Zeynep Altan,Impact of Feature Selection for Corpus-Based WSD in Turkish, 5th Mexican International Conference on Artificial Intelligence MICA I 2006,13-17 November 2006, in Lecture Notes in Artificial Intelligence pp.868-878. rdf: XMLLiteral rdfs: seeAlso http://www.beykent.edu.tr/uploads/ckeditor/attachments/5597c5836565732b8c150000/ZEYNEP_ALTAN.pdf rdf: XMLLiteral.

Figure 7 shows there exist computer science student has a type student(class assertion), he has name, number, surname (data property assertion) and he attends beykent university(object property assertion).

Found in	Entity	Match
Display name	computer_science_student	computer_science_student
IRI	computer_science_student	http://www.semanticweb.org/umutkaya/ontologies/2016/0/untitledontology-7#computer_science_student
ClassAssertion	computer_science_student	computer_science_student Type student
DataPropertyAssertion	computer_science_student	computer_science_student name "umut"^^xsd:string
	computer_science_student	computer_science_student number 130821002
	computer_science_student	computer_science_student surname "kaya"^^xsd:string
ObjectPropertyAssertion	computer_science_student	computer_science_student attends beykent_university

Figure 7. computer science student search in draft ontology

Figure 8 shows that the university subject has rdf: type as predicate

and owl:class as object. It supports our draft ontology with query results.

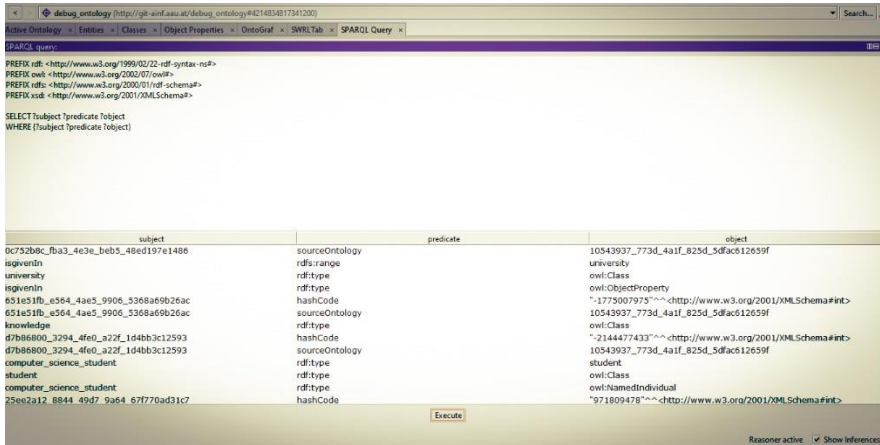


Figure 8. SPARQL query results according to subject, predicate and object

Type query as the SPARQL sample query which is shown in Figure 9 provides the different data types of our draft ontology.

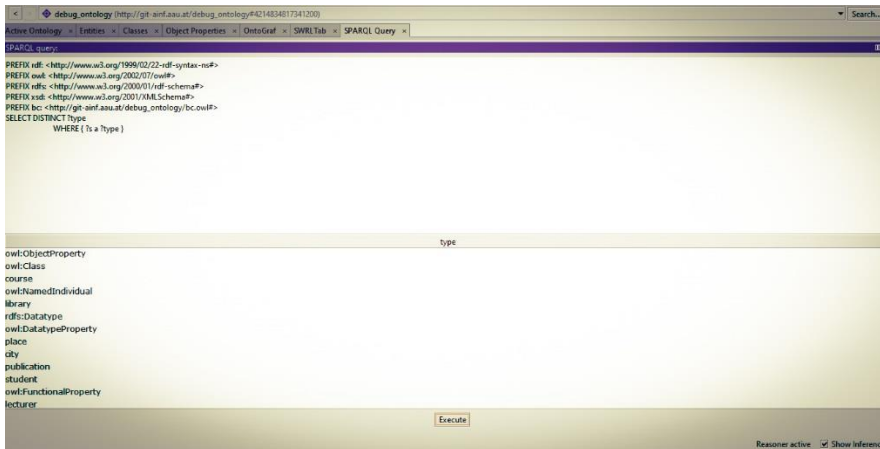


Figure 9. SPARQL query results for Distinct type criterion

In Figure 10, we use the SPARQL sample query according to the predicate and frequency that shows rdf: type, domain and range, etc.

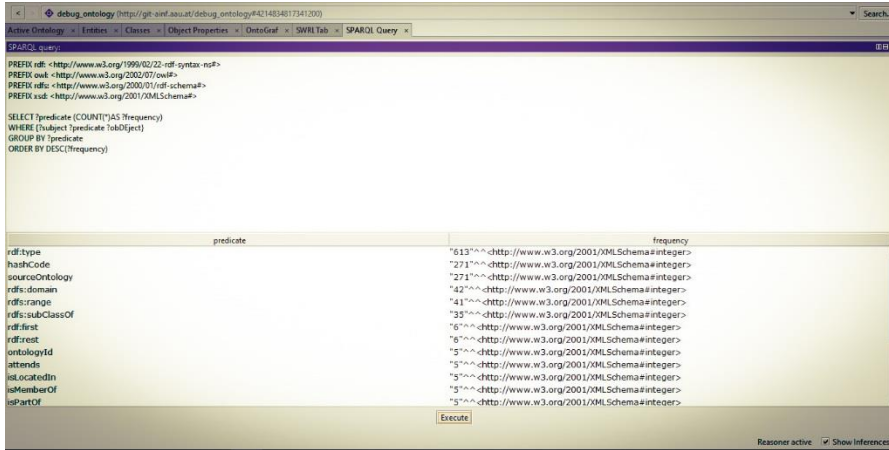


Figure 10. predicate and frequency SPARQL query results for draft ontology

5. CONCLUSION AND FUTURE WORKS

In recent years, semantic web has widely being applied in different domains and its influences on the development of information technology have gradually been increasing. Thereby, building the ontology rapidly and correctly has become an essential task for content based searching on the Internet. Google, Flickr, YouTube and e-Bay exploit these type applications. As an example, IEEE explore on-line library holds approximately 3 million documents about scientific articles, publications and conference papers. It is possible to find any document in request within the shortest time while seeking according to keywords. The main goal of the proposed ontology is to ease the article search procedure, i.e. to enable the researchers to find their articles easily. Beykent University Library has been selected as test case in which the generated data exports the OWL documents. The results obtained from the implementation of the conceptualization have produced the draft ontology with the description of 20 classes using the relationships among them. This research has completely been focused on exploring the main concepts of the underlying domain academic researches. Lastly, SPARQL queries have been provided powerfully by providing easy access to common attributes from the constituted OWL structure.

Future work needs to consider the extension of our ontology to link it with other library databases and to be more helpful to the academic publishing domain. This will suggest enriching the process of ontology development. Moreover, as another application of study, we intend to compare it with other similar ontologies. Different terminologies may be used for similar domains in the other ontology. Since we developed the draft ontology using lattice structure, it can easily be mapped with new terminologies. As the concepts belonging to same domains match semantically, it is possible to accept that the matching is materialized by two different ontology. This can be realized with one to one correspondence between two ontologies, and it is required to find similar concept in other taxonomy for each concept in the current one.

REFERENCES

- [1] Gruber, T: Ontology to appear in the Encyclopedia of Database Systems, Liu, L., Ozsu, M. T. (Eds.), Springer-Verlag, 2008. (2007). <http://tomgruber.org/writing/ontology-definiton-2007.html>
- [2] Borst, W: Construction of engineering ontologies for knowledge sharing and reuse, PhD. thesis, University of Enschede (1997).
- [3] Gruber, T : A translation approach to portable ontology specifications, Knowledge Acquisition 5(2) (1993) 199-220.
- [4] Tillet, B : What is FRBR? A conceptual model for the bibliographic universe, 2004, library of Congress cataloging distribution service.
- [5] Studer, R., Benjamins, V. R., Fensel, D.: Knowledge engineering: principles and methods, Data & Knowledge engineering 25 (1-2) (1998) 161-197.
- [6] Tang, M. T., Toussaint, Y.: Introducing metrics in the lattice to build ontology (2014).
- [7] Sowa, F : The role of logic and ontology in language and reasoning, Springer, Netherlands, Dordrecht, 2010, pp. 231-263. doi:10.1007/978-90-481-8845-1_11.
- [8] Lickly, B.: Static model analysis with lattice-based ontologies, PhD. thesis, EECS Department, University of California, Berkeley (Nov 2012).

<http://www2.eecs.berkeley.edu/Pubs/TechRpts/2012/EECS-2012-212.html>.

[9] Grosz, B., Volz, R., Horrocks, I., Deckers, S.: Description logic programs: combining logic programs with description logic, in: Proceeding of 12th International Conference on World Wide Web, 2003.

[10] Krotzsch, M., Simancik, F., Horrocks, I.: Description logics, IEEE Intelligent Systems 29 (1) (2014) 12-19.

[11] Giordano, L., Gliozzi, V.: Encoding a preferential extension of the description logic sroiq into sroiq , in: Esposito, F., Pivert, O., Hacid, M-S, Ras, Z.W, S.Ferilli, (Eds.) ISMIS Vol. 9384 of lecture notes in Computer Science, Springer, 2015, pp.248-258.