International Conference on Cyber Security and Computer Science (ICONCS'18),
Oct 18-20, 2018 Safranbolu, Turkey

An Ontology-based Approach to Terminology Management for Healtcare Systems

Y. GÜLTEPE[1]

[1] Kastamonu University, Kastamonu/Turkey, yasemingultepe@kastamonu.edu.tr

***Abstract* - The Semantic Web is a web environment that allows well defined information and services to be easily understood by machines. The main component of the Semantic Web is ontologies, which formally define a set of concepts for a domain and the relationships between concepts. One of the areas where ontologies can be used is the field of healthcare. In particular, the use of ontologies in the field of healthcare is recommended because of the formal representation of a subject area and its support for reusability. Many medical classification systems are used in the field of medical informatics. The deficiency seen in the proposed approaches for health information systems is that there is no meaningful reference and sharing system that enables the collection of classification and coding systems. Considering the general classification and coding systems, it is clear that a system is required for faster, accurate and efficient processing. In accordance with the interoperability needs of health information systems that are constituted by different ontology combinations, this study propose ontology based of Systemized Nomenclature of Medical-Clinical Terms (SNOMED CT) Concept Model. This model remarks reasonable definition of concepts in SNOMED CT. The official semantics of ontology enhance the ability to automate information management of complex terminology, facilitate the maintenance of clinical decision support materials, and significantly improve interoperability.**

***Keywords* - Healthcare Systems, Ontologies, Terminology, Snomed Ct, SPARQL.**

I. Introduction

The most important part of today's health information systems is Electronic Patient Record (EHK). EHK is the electronic retention of all kinds of data associated to health status and health care during the life of a person. Interoperability between EHK standards is one of the major challenges [1, 2]. On the other hand, there are difficulties in data exchange between different data sources used in different health information systems. In order to reduce the errors in health services due to misinterpretation and misinterpretation of health data, it is necessary to pay attention to create interoperable structures within the framework of interoperability needs and in this continuation, easy and possible integration of solutions and use for the benefit of the country.

Using medical field knowledge from ontologies such as SNOMED CT can get the better of some restrictions of keyword-based systems, thus improving the search sample of warehouse users. An machine-controlled approach based on ontological partition is an effective and practical way to support modeling, management and user pilotage in clinical archive repositories [3, 4, 5].

Biomedical ontologies are broader than taxonomies; define the relations between the concepts of an area and the constraints and share the information by using a common terminology [6]. Different ontologies and terminologies in the field of health can be accessed via web services on BioPortal. BioPortal is a collection of nearly 200 biomedical ontology (OBO, OWL, RDF, protégé frames and RRF) developed in different formats. RDF allows users to query and extract content using the SPARQL query language. SPARQL is a standard and query language for a Web of Data (Semantic Web).

BioPortal SPARQL provides a service to query BioMedical ontologies using the SPARQL standard. Ontologies are converted from their original formats (OWL, OBO and UMLs / RDF, etc.) to RDF triples and transferred to a triple store [7].

In this study, the study of questioning with SPARQL query language from the ontology which is the main component of Semantic Web is explained. In this study, the basic component of Semantic Web, ontology, SPARQL query language. After the information about semantic web and SNOMED CT was given at the entrance of the paper, BioPortal portal which is the basis of the application was mentioned. In the next section, the semantically kept information on the BioPortal SNOMED CT ontology was questioned by SPARQL. In the conclusion section, the results of the queries were evaluated.

II. Preliminaries

A. Data Analytics with SNOMED CT

Here are some of the reasons for using SNOMED CT: a) The SNOMED CT data format is in simple text format b) SNOMED CT consists of a large number of concepts and clusters of object-property-value triad c) SNOMED provides a precious set of inter-relationships between concepts. Hierarchical relations here define certain concepts as children of more general concepts.

In this study, the diabetes area was chosen as a prototype. SNOMED CT terminology values will be used in this field. Table 1 presents information on the diabetes mellitus concept in SNOMED CT.

Table 1: Diabetes mellitus concept in SNOMED CT [8].

Code System Concept Code	73211009
Code System Concept Name	Diabetes Mellitus (disorder)

For diseases/disorders, SNOMED CT uses the relationships between concepts to supply deductive, computer-readable descriptions of medical concepts.

There are some category of relationships described or modeled in SNOMED CT. The following are examples: Is-a, has finding site, causative agent, associated morphology. Table 2 shows the example of SNOMED CT relationships related to diabetes mellitus.

Table 2: Example of SNOMED CT relationships related to diabetes mellitus.

Diabetes mellitus *has finding site* Structure of endocrine system
Diabetes mellitus *is_a* Disorder of glucose metabolism

The relationships on BioPortal differ, as it provides an ontology-based application. Table 3 presents an example of the relationship of diabetes mellitus in the ontology of snomed on bioportal.

Table 3: OWL Representation SNOMED CT Diabetes mellitus.

Diabetic mellitus *subClassOf* Disorder of glucose metabolism
Diabetes mellitus *semantic_type* Disease
Diabetes mellitus *Associated finding of* Suspected diabetes mellitus

B. The Diabetes Ontology

Diabetes ontology has been established to eliminate the shortcomings of databases, to eliminate the constraints and to draw attention to the semantic integrity between the data sources and to define the semantic relations between the information in the web environment [9].

In the described ontology, the concepts of diabetes chronic diseases, concept hierarchy, different concepts and examples are explained.Thus the *diabetesOnt* ontology has been described as a chronic disease information store of diabetes.

The basic class definitions and properties of diabetes ontology are shown in Figure 1. *The diabetesOnt* ontology was developed using the Protégé ontology development editor. Ontologies are defined visually by the Protégé ontology development editor's graphical interface and thus the desired area can be modeled. In addition, it facilitates the development of ontologies and reduces the possibility of errors.

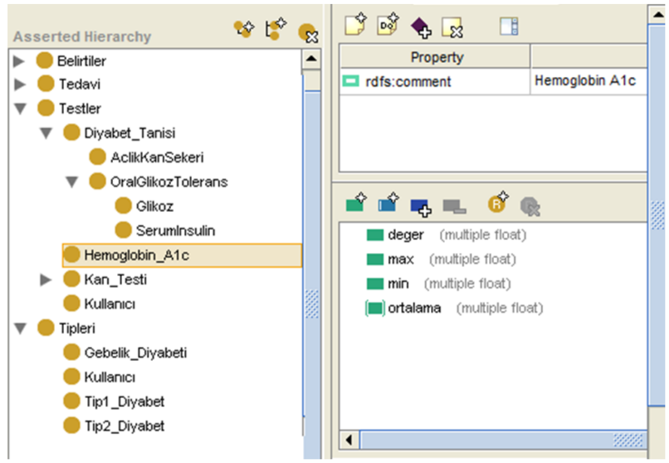

Figure 1: “diabetesOnt” ontology.

In the document search process, some semantic rules are defined. These rules are based on a rule knowledge base based on ontology and are created using the Semantic Web Rule Language (SWRL).

SWRL rules are described under the *diabetesOnt* ontology. These semantic definitions are referred to as *diabetesOnt* and represent the general ontology structure of diabetes. It was created. Table 4 presents rule definitions created in the *diabetesOnt* ontology which is the pattern ontology of diabetes.

Table 4: View of SWRL semantic search rules related to *diabetesOnt* ontology on the protégé editor.

SWRL Rulebase	*Rule 1:* foaf:Person(?x) ∧ foaf:aclikKanSekeri(?x, ?test) ∧ diyabetOnt: AclikKanSekeri (?test, ?y) ∧ swrlb:greaterThanOrEqual(?y, 126) → diyabetOnt:Diyabet(?x)
	Rule 2: foaf:Person(?x) ∧ foaf:hemoglobin_A1c(?x,? test) ∧ diyabetOnt: hemoglobin_A1c (?test, ?y) ∧ swrlb:greaterThanOrEqual(?y, 6.5) → diyabetOnt:Diyabet(?x)
	Rule 3: foaf:Person(?x) ∧ foaf:hamilelik(?x, true) ∧ foaf:aclikKanSekeri(?x, ?test) ∧ diyabetOnt:AclikKanSekeri(?test, ?y) ∧ swrlb:greaterThanOrEqual(?y, 140) → diyabetOnt:Gebelik_Diyabeti(?x)

These rules in ontology; The user accesses the search words, the semantic links to the diabetes diagnosis in the *diabetesOnt* ontology. At this point, if the rules shown in Table 4 are executed and the result of the document is deducted, the blood glucose level of an individual without diabetes becomes 120 mg / dl in fasting condition and does not exceed 140 mg / dl. The blood glucose level measured in fasting or toughness is above these values indicates the presence of diabetes.

Personal information is added to the sample of the *Person* class in the Friend of a Friend (FOAF) ontology, and the test results information is added as examples of

OralGlukozTolerans and *AclikKanSekeri* class in diabetesOnt ontology. The widely used FOAF ontology consists of classes such as person and name, surname, email address (mbox) of those classes. Although classes and properties can be defined with RDFS, complex relationships between objects cannot be modeled with RDFS. For this reason, FOAF is more qualified with Web Ontology Language (OWL).

III. BIOPORTAL SPARQL

SPARQL, whose standards are defined by W3C, a standard and query language for a Web of Data. SPARQL, which is an example of SQL syntax, is used to collect data from RDF files as well as information from the database. Therefore, it is extremely suitable for the discovery of devices and services. Queries are used by information management applications for inference operations.

Salvadores et al. [7] are community based ontology repositories for developed biomedical ontologies. BioPortal is able to offer new ontology development, effective communication and search methods. Web portals are web applications where users can find what they are looking for, customize content, and collaborate with other environments. Ontologies and metadata are published on the RDF-based serializations in the portal *sparql.bioontology.org.BioPortal*.

BioPortal SPARQL is a service to query BioMedical ontologies using the SPARQL standard. Operation logic is as follows; Ontologies are converted from their original formats (OWL, OBO and UMLS / RDF, ..) to RDF triples and stored in a triple store. This dataset include 203M triples, more than 300 ontologies and 9M mapping between terms.

The definition of SNOMED CT on BioPortal is collected under 19 sub-headings: Body structure, clinical finding, environment or geographical location, event, observable entity, organism, pharmaceutical/biologic product, physical force, physical object, procedure, qualifier value, record artifact, situation with explicit context, SNOMED CT Model Component, social context, special concept, specimen, straging and scales, substance.

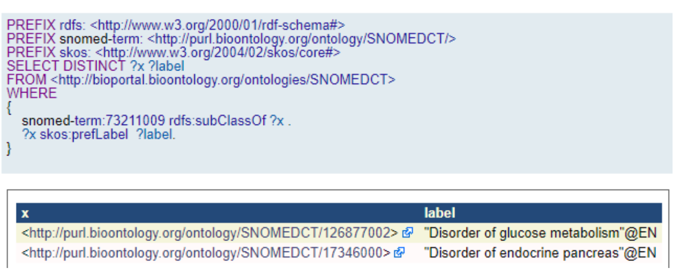

Figure 2: Retrieval of 73211009 (diabetes mellitus)'s sub classes from ontology.

After the ontology is created, one of the processes that can be done is to query the information kept in ontology with an ontology query language. In this section, SPARQL queries were run on SNOMED CT ontology in BioPortal.

The simplest example for the queries is the subclass of 73211009 diabetes mellitus in ontology. Figure 2 shows retrieval of 73211009 (diabetes mellitus)'s sub classes from ontology.

It is stated that the relationship between "73211009 (diabetes mellitus)" class and *http://purl.bioontology.org/ontology/SNOMEDCT/cause_of*. According to this, the object with the "*cause_of*" predicate from the relevant ontology is requested by the given query. The "*cause_of*" predicate describes the illness-cause relations. Figure 3 shows the operation of this query and writing the query results to the screen. As a result, it returns total 35 lines.

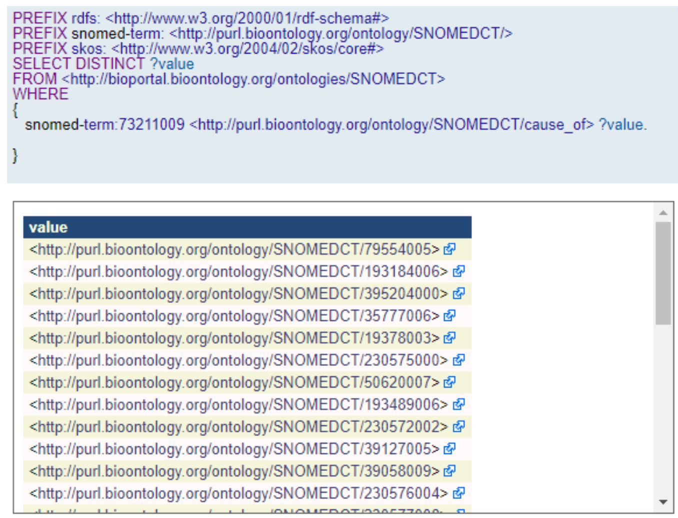

Figure 3: Retrieval of "cause_of" predicate 73211009 (diabetes mellitus) from ontology.

Table 4 lists the disorder names of some of the code in the query results in Figure 3.

Table 4: View of SWRL semantic search rules related to *diabetesOnt* ontology on the protégé editor.

Concept ID	Preferred Term
79554005	Asymmetric proximal motor neuropath co-occurrent and due to diabetes mellitus
193184006	Neuropathy co-occurrent and due to diabetes mellitus (disorder)
395204000	Hyperosmolar non-ketotic state in type 2 diabetes mellitus
35777006	Diabetic mononeuropathy multiplex
19378003	Pseudotabes co-occurrent and due to diabetes mellitus (disorder)

It is stated that the relationship between "73211009 (diabetes mellitus)" class and *http://purl.bioontology.org/ontology/SNOMEDCT/associated_finding_of*. According to this, the object with the "*associated_finding_of*" predicate from the relevant ontology is requested by the given query. The *"associated_finding_of"* predicate describes the illness-associated finding relations. Figure 4 shows the operation of this query and writing the query results to the screen.

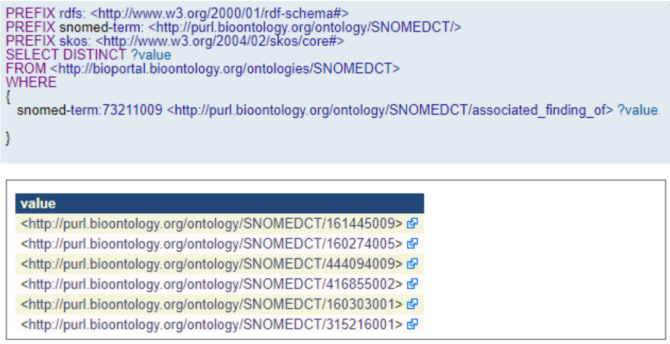

Figure 4: Retrieval of "*associated_finding_of*" predicate 73211009 (diabetes mellitus) from ontology.

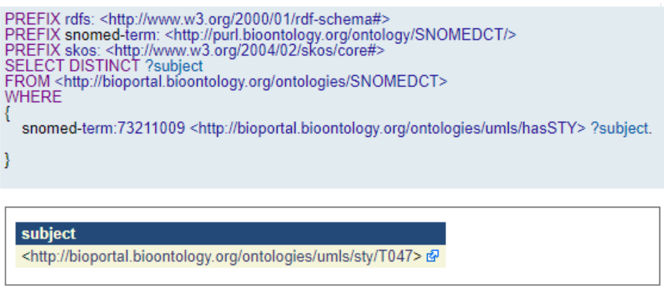

Figure 5: Retrieval of semantic type of 73211009 (diabetes mellitus) from ontology.

For each concept specific semantic type and their types and their relationships are arranged as an ontology. Figure 5 shows the semantic type value of 73211009 (diabetes mellitus). The semantic type value of 73211009 (diabetes mellitus) is T047- Disease or Symptom. The 133 semantic types are numbered in the UMLS Metathesaurus ontology from T001 to T203.

IV. Conclusion

Nowadays, there are many information systems developed independently for different purposes in the field of health. The importance of interoperability in the collection of data in these systems in certain data form in accordance with national and some international standards is increasing. Semantic Web technologies are used to solve this problem.

The advantages of ontological structure of health information systems can be listed as data and information integration, interoperability and consistency.

Thanks to the use of interpreted semantic knowledge bases that are understood by machines, developed software can make meaningful inferences. Therefore, software designed on semantic knowledge bases can give more logical and relevant results while presenting the users what they are looking for.

References

[1] J.A. Durou, *Complete interoperability in healthcare: technical, semantic and process interoperability through ontology mapping and distributed enterprise integration techniques*, Doctor of Philosophy thesis, University of Wollongong, 247p., 2009.

[2] A. Ryan, "Towards Semantic Interoperability in Healthcare: Ontology Mapping from SNOMED-CT to HL7 version 3," *In Proceedings of the Second Australasian Ontology Workshop*, pp. 69-74, 2006.

[3] N.F. Noy, "Semantic integration: a survey of ontology-based approaches," *published 2004 in SIGMOD Record*.

[4] OWL 2.0, Web ontology language overview, *W3C Recommendation 2012*, Available: http://www.w3.org/TR/owl2overview/.

[5] T. Gruber, *Ontology in Encyclopedia of Database Systems*, Springer-Verlag, pp. 1963-1965, 2009.

[6] N.F. Noy, M. Dorf, N. Griffith, C. Nyulas, and M. A. Musen, "Harnessing the Power of the Community in a Library of Biomedical Ontologies," *In Workshop on Semantic Web Applications in Scientific Discourse at the 8th International Semantic Web Conference*, 2009.

[7] M. Salvadores, M. Horridge, P. R. Alexander, R. W. Fergerson, M. A. Musen, and N.F. Noy, "Using SPARQL to Query BioPortal Ontologies and Metadata," *ISWS 2012*, pp. 180-195.

[8] Centers for Disease Control and Prevention, Application Version: 4.2.5, 2018, Available: https://phinvads.cdc.gov/vads/ViewCodeSystemConcept.action?oid=2.16.840.1.113883.6.96&code=73211009

[9] S. El-Sappagh, and F. Ali, "DDO: a diabetes mellitus diagnosis ontology," *Applied Informatics*, vol. 3, no. 5, 2016.