

Apache JENA ile OWL Uygulaması

Veysi Ertekin

Bilgisayar Mühendisliği Bölümü, İstanbul Aydın Üniversitesi

ertekin.veysi@gmail.com

Özet— Bu dökümanda WEB 3.0'ın önceki versiyonlarına binaen yenilikleri, avantajları ve de bize getireceği bazı kolaylıklarından bahsedilecektir. Bunun yanında WEB 3.0 teknolojisinin bazı temel kavramlarına değinilip örnek bir ontoloji uygulamasına değinilecektir.

Anahtar Kelimeler— Ontoloji, Semantik WEB, SPARQL, RDF, Apache JENA, OWL.

Keywords— Ontology, Semantic WEB, SPARQL, RDF, Apache JENA, OWL.

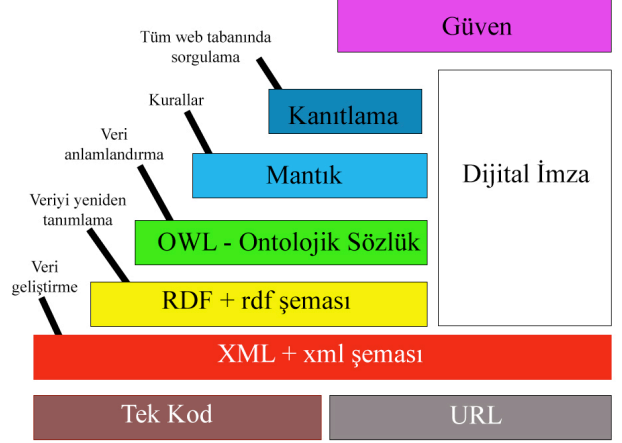
I. GİRİŞ

Günümüzde sıklıkla kullanılan web teknolojisi ile “diller arası otomatik metin çevirimi”, “bizim yerimize internet sitelerinde fiyat analizi yapıp bizim için uçak bileti alan ajanların (agent) geliştirilmesi”, “bir forum sitesi içerisindeki yazılardan insanların en çok sevdiği film kahramanının kim olabileceğinin bulunması” gibi veri analizi ve yazılan program veya makinanın içeriği bir insan gibi **tamamıyla** anlayıp daha **akıllı** davranabilmesini sağlamak adına WEB 3.0 teknolojisi *W3C* standartları altında ortaya konuldu.

II. WEB 3.0

Web 3.0 yani gelecekteki anlamsal ağ diyebileceğimiz, başta Google olmak üzere arama motorlarının şuan ki algoritmalarla başaramayacağı yepyeni bir dil yapısıdır. Kullanılmak istenen semantik dil, insan mantığının muhakeme yeteneğini hedef almaktadır. Semantik ağ dili, insan zekasının sorgulamasını prototip alan bildirimsel derleme lisansı “Web Ontology Language (OWL) yani anlamsal web dilidir.

Semantik web, html tabanlı fihristlemeye dayalı World Wide Web dilinin bilgi arama ve düzenleme eksikliğini gidermeyi amaçlar. Web 3.0 dünyası; veri-merkezli genişletilebilir işaretleme (XML), kaynak tanımlama çerçevesi (RDF) ve web ontoloji dili (OWL) dillerini kullanarak işlem yapar. OWL, XML ve RDF verileri arasında, sınıflar arası ilişkiler, en önemlilik, eşitlik, özelliklerin daha iyi sınıflandırılması, özelliklerin karakteristikleri ve numaralandırılmış sınıflar gibi yeni anlamsal bağlar kurar.



İnternetin kurucusu Tim-Berners-Lee Semantic Web'in vizyonunu şöyle ifade etmiştir: “Web için bir hayalim var. Öyle ki bilgisayarlar, web üzerindeki bütün veriyi analiz etmeye muktedir olacaklar. Ortaya çıktığı zaman anlamsal ağ; ticaretin günlük mekanizmalarını, bürokrasiyi ve günlük yaşamlarımızı birbiri ile konuşan makinalar tarafından yürütecek. İnsanlığın asırlardır konuşup durduğu “akıllı ajanlar” nihayet gerçekleşecek.”[1]

En alt katmanda yer alan XML(eXtensible Markup Language), kullanıcıtarafından tanımlanmışsöz varlığıkullanılarak yapısal web belgeleri yazılmasınaışlayan bir dildir. RDF, web kaynaklarıile ilgili olarak yalın ifadeler yazılan varlık-ilişki modeline benzer yalın bir veri modeli konumunda Anlamsal Web'in veri modelini oluşturmaktadır. Zorunlu olmamakla birlikte sıklıkla XMLile ifade edilmektedir. Bu nedenle XMLkatmanının üstünde yer almaktadır. RDF Şeması, web nesnelerini sıradüzen içerisine düzenleyen modelleme ilkeleri sağlamaktadır. Sınıflar ve özellikler, alt sınıf ve alt özellik ilişkileri, etki alanıve erişim kısıtlamalarıtemel ilkelerdir. RDF Şeması, RDFtabanlıdır ve ontoloji yazımıçin bir ilkel dil olarak görülebilir. Ancak, RDF şemasınıgenişleten ve web nesneleri arasında karmaşık ilişkilerin tanımlanmasına izin verecek daha güçlü ontoloji dillerine gereksinim vardır.

Mantıkkatmanı, ontoloji dilini güçlendirmek ve uygulamaya özel bildirim deyimi bilgisinin yazımına izin vermek için kullanılmaktadır. Kanıtkatmanı, hem tündengelimli işlemleri hem de web dillerinde kanıtların temsil edilmesini ve kanıt onaylanmasınıicermektedir. Son olarak Güven katmanı, sayısal imzalar ve güvenilir etmenler, oranlar, sertifika kuruluşlarıve tüketicilerin önerilerini temel

alan diğer bilgi türlerinin kullanımı ile ortaya çıkmaktadır. Güven katmanının, Anlamsal Web katmanlı yapısının en üstünde yer alması nedeniyle çok önemli görev üstlenmektedir. Nitekim anlamsal web'in kullanılabilirliği, kullanıcılar işlemlerinde güvencede olduklarında ve sağlanan bilginin niteliğine bağlı olarak elde edilecektir. [2]

III. ONTOLOJİ NEDİR?

Ontoloji ya da varlık felsefesi, temel sorunu varlık olan felsefi disiplin. Varlık ya da varoluş ile bunların temel kategorilerinin araştırılmasıdır. "Varlık" ve "varoluş" ayrımını; "Varlık vardır" ve "Varlık yoktur" fikirlerini tartışır.

Aristoteles'e göre ontoloji varlığın mahiyetinde varlığın bilimdir veya varlıkların incelenmesidir. Ontoloji hangi varlık kategorilerinin daha temel olduğunu belirlemekle uğraşır ve bu kategorilerdekilerden hangilerinin var olduğunu söyleyebileceğini sorar.

Değişik filozoflar temel varlık kategorileri için değişik listeler yapmışlardır. Ontolojinin temel sorunlarından biri "Temel varlık kategorileri nelerdir?" sorusudur.[3]

Ontoloji kısaca bir alan içerisindeki kavramları, bunlar arasındaki ilişkileri gösterir.[5]

IV. OWL (WEB ONTOLOGY LANGUAGE)

Ontolojileri tanımlamak ve çeşitlendirmek için kullanılan bir dildir. Bir ontoloji, ilgili özellikler ve örnekleri ile birlikte sınıfların açıklamalarını içerebilir. OWL, bilginin içeriğini sadece insanlara gösteren değil, bunun yanında işleyen uygulamalar tarafından kullanılmak üzere tasarlanmıştır. OWL, XML, RDF, and RDF Schema (RDF-S) tarafından desteklenen web içeriğinin makinalar tarafından daha iyi yorumlanabilmesini, biçimsel bir semantik ile birlikte ek sözcük kümeleri sunarak kolaylaştırır. OWL, OIL ve DAML+OIL gibi eski dilleri temel alır, ve şu anda W3C tavsiyesidir.

OWL, Semantic Web'in gelecekteki uygulamaları için temel teknolojilerden biridir. Sayıca çok ve çeşitli alanlardaki uygulamalarda önemli bir rol oynamaktadır, ve araştırmalarını araçlara, usamlama tekniklerine, biçimsel temellere ve dil eklentilerine odaklanmıştır.

OWL, anlamsal web bilgisinin işlenebilmesi için genel bir yol sağlamak üzere tasarlanmıştır. OWL, XML, RDF, ve RDF-S tarafından sağlanan anlamların açıklanmasını sağlayan imkânları kolaylaştırmak için geliştirilmiştir. Bu nedenle, makinalar tarafından çevirilip kullanılabilir web içeriğini gösterebilme yeteneği sayesinde, bu web dillerinin evrimsel devamı gibi düşünülebilir. OWL, XML tabanlı olduğu için, OWL bilgisi kolayca farklı işletim sistemlerine sahip bilgisayarlar ve farklı programlama dilleri arasında alınıp verilebilir. Dil, bilgisayarlar tarafından okunmak üzere tasarlandığından ve insanların okuyabilmesi çok fazla göz önünde bulundurulmadığından, araç meselesi olarak düşünülebilir. OWL, değer yönetimi çerçevesi, büyük ölçekli entegrasyon ve web üzerinde veri paylaşımı için standart yaratmada kullanılmaktadır.

OWL'nin, artırılmış açıklayıcılık, daha basit bir veri modeli ve serilizasyon, ve bir dizi iyi tanımlanmış alt diller içeren, genişletilmiş bir versiyonu, (bazen OWL 1.1 olarak da anılmaktadır, fakat resmileşmiş değildir) sunulmuştur.[4]

OWL üç alt dil bileşeninden oluşur:

- OWL-Lite
 - Sözdizimsel olarak basit bir alt dil olarak tanımlanabilir. Basit sınıf hiyerarşisi ve basit kısıtlamalar içerir.
- OWL-DL
 - OWL-Lite tabanlıdır. Ayrıca tanım mantığını (Description logics) içerir. Otomatik karar vermede yardımcı olabilecek temel mantıksal koşullar içerir.,
- OWL-Full
 - En etkileyici OWL alt dilidir. Dilin hesaplama ve karar verilebilirlik ölçütünü garantiler.[6]

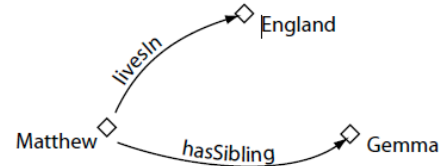


Figure 3.2: Representation Of Properties

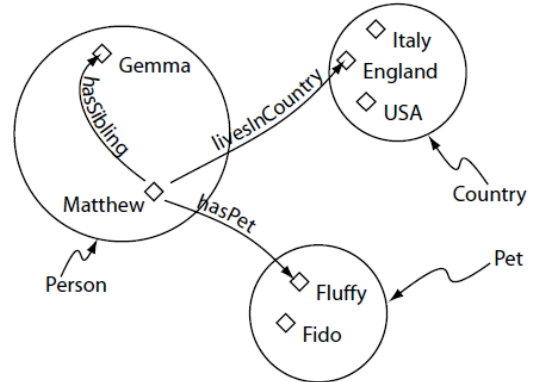


Figure 3.3: Representation Of Classes (Containing Individuals)

Örnek bir ontoloji

V. SPARQL

RDF'in "triple" denilen özne-eylem-nesne olan "predicate" mantığına dayanan üçlüleri birbirleriyle ilişkilendirildiklerinde büyük graflar ortaya çıkabilir. (RDF, XML'in ağaçlı yapısından ziyade graf yapısını daha önceki yazıda yazmıştık) İşte internet ağında bilgi tabanlarını oluşturan ve oluşturacak bu üçlüleri sorgulamak için kullanılan sorgulama diline SPARQL (telafuzu "sparkle" olarak da ifade edilmekte) denmekte. Ontolojilerin içinde barındırdığı karmaşık graf

yapıları düşünülürken sparql sorgulama dilinin önemi büyük. Aklıma gelmişken belirteyim ontolojileri rdf ile de owl ile de yazabilmekteyiz. Ancak Sparql dilini owl ile yazılan ontolojilerde tam olarak kullanamamaktayız.

Sorgulama yapabilmemiz için bir bilgi tabanımız, yani ontolojimizin üçlülere olması gerekmektedir, bunlar turtle yazımı foaf üçlülere olarak şöyle olsun:

```
@prefix foaf: <http://xmlns.com/foaf/0.1/> .
_:a foaf:name "Johan Elmander" .
_:a foaf:mbox <mailto:johan@abc.com> .
_:b foaf:name "Ahmet Çölaşan" .
_:b foaf:mbox <mailto:ahmet@abc.org> .
_:c foaf:mbox <mailto:mehmet@abc.org> .
```

SPARQL sorgusu:

```
PREFIX foaf: <http://xmlns.com/foaf/0.1/>
SELECT ?name ?mbox
WHERE {
  ?x foaf:name ?name .
  ?x foaf:mbox ?mbox }

```

ve sorgunun sonuçları :

name	mbox
"Johan Elmander"	<mailto:johan@abc.com>
"Ahmet Çölaşan"	<mailto:ahmet@abc.org>

Biraz daha açmak gerekirse Prefix kısmı foaf (friend of friend) ontolojisini kısaltarak bunu foaf a çeviriyor ve bu ontoloji içindeki name ve mbox gibi kavramları (foaf:name gibi) esnekçe kullanmamıza yarıyor.

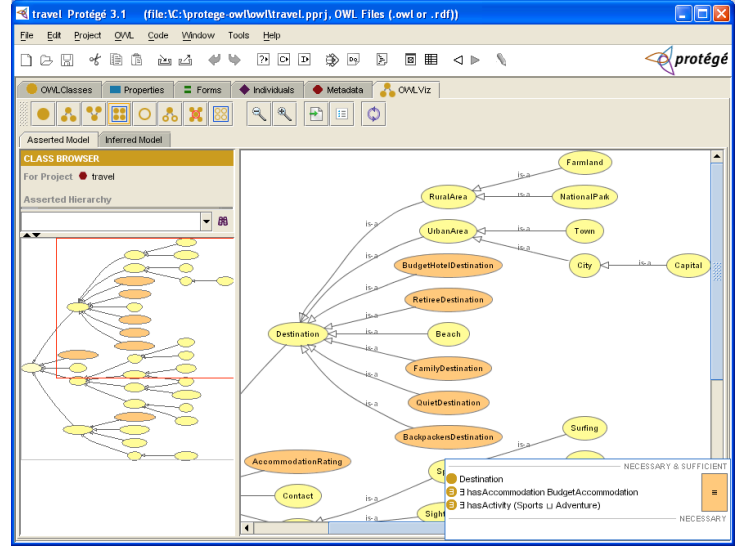
? işareti ile başlayan terimler ise tahmin edileceği gibi bizim değişkenlerimiz. Onları istediğimiz gibi belirleyebiliriz. Buradaki sorgumuzda istenen hem isim olan name kavramına hem de mail adresi olan mbox kavramına sahip nesneleri getirmek istiyoruz.[7]

VI. PROTÉGÉ

Protégé Stanford Üniversitesinde geliştirilmeye başlanan açık kaynak kodlu bir ontoloji editor ve eğitim sonucu oluşan bilginin kullanıldığı bir çatıdır (framework).

Masaüstü istemcisi veya web istemcisi ile ontolojilerin modellenmesini; ayrıca OWL, RDF(S), TTL(Turtle) ve XML

gibi bir çok formatta geliştirme yapılabilmesini de destekler.[8]

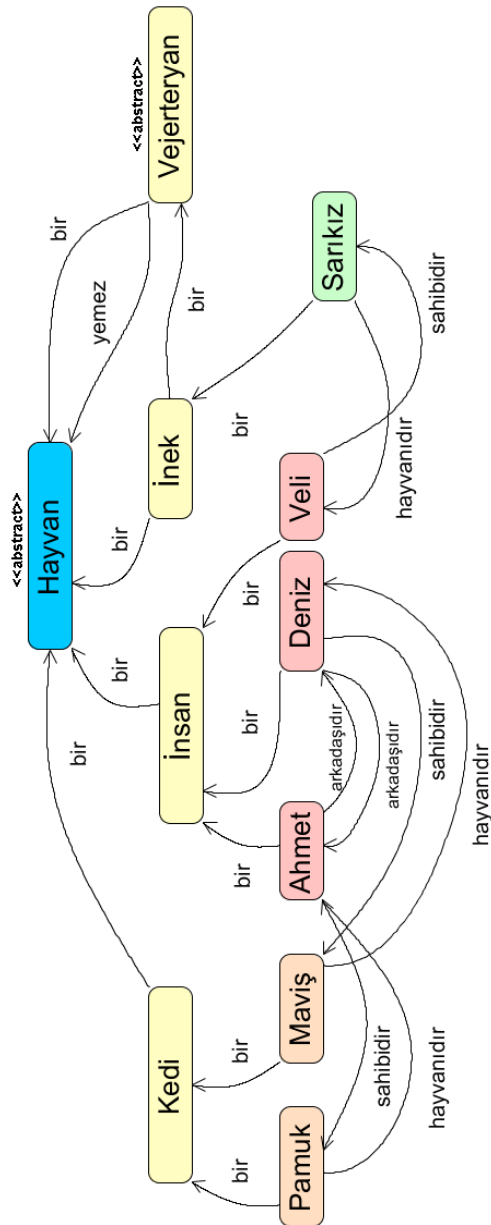


VII. APACHE JENA

Semantik Web uygulamalarının yazımı için oluşturulmuş bir çatıdır. Semantik web bağlı veri(linked-data) uygulamaları, araçları, sunucularının Java dilinde yazılmasına olanak kılar.

JENA faramework:

- RDF, XML, N-triple (triple [üçlü]: Subject + Predict + Object) ve Turtle formatlarında okuma, işleme ve yazma için bir API sunar.
- OWL, RDFS ontolojilerinin işlenmesine yardımcı olur.
- RDF ve OWL veri kaynaklarından koşullu kural tabanlı sorgulamayı destekler.
- Çok büyük miktardaki RDF üçlülerinin saklanması sağlar.



Bu ontolojide varlıklar bir birleri arasında birer triple[üçlü] oluşturacak şekilde bir birine bağlanmışlardır.

Bir triple ifadesindeki yargı iki şekilde olabilir:

1. Tek yönlü;

“Pamuk *bir* Kedi ‘dir’”

2. İki yönlü;
“Sarıkız *hayvanıdır* Veli’nin”
“Veli *sahibidir* Sarıkız’ın” gibi.

Belirtilen ontoloji tanımı incelenecek olunursa “Hayvan” ve “Vejeteryan” sınıflarının birer “abstract” yani nesnesi oluşturulamaz birer üst sınıf oldukları görülecektir.

“İnek” *bir* “Hayvan”dır. Ayrıca “Vejeteryan”dır, ‘Vejeteryanlar “Hayvan” *yemez*’. ‘O halde “İnek”ler hayvan *yemez*’ gibi bu bağlı verilerden SPARQL ve Apache JENA ile uygulamaya dönüşümü gerçekleştirdik.

Oluşturduğumuz ontolojide Apache JENA içerisinde SPARQL ile bir kaç sorgu yaparsak (***Bu örnekler iletilen proje dosyası içerisinde yer almaktadır***):

- i. “Pamuk” kimin “hayvanidir”?

```
String sparqlQueryString = "PREFIX o:  
<http://www.owl-ontologies.com/Ontology1372543966.owl#> "  
+ "SELECT ?isim WHERE { "  
+ " ?Insan o:isim ?isim . "  
+ " ?Kedi o:hayvanidiri ?Insan . "  
+ " ?Kedi o:isim \"Pamuk\" . }";
```

```
RESULT:
( ?isim = "Ahmet"^^xsd:string )
```

- ii. “Veli”nin ineği hangisi?

```
sparqlQueryString = "PREFIX o:  
<http://www.owl-ontologies.com/Ontology1372543966.owl#> "  
+ " SELECT ?isim WHERE { "  
+ " ?Inek o:isim ?isim . "  
+ " ?Insan o:isim \"Veli\" . "  
+ " ?Insan o:sahibidir ?Inek . }";
```

```
RESULT:
( ?isim = "Sarikiz"^^xsd:string )
```

IX. KAYNAKÇA

- [1] <http://www.semiyun.com/2012/11/web-3-0-nedir/>
- [2] Cihat Demirli, Ömer Faruk Küçük; *ANLAMSAZ WEB (WEB 3.0) VE ONTOLOJİLERİNE GENEL BİR BAKIŞ* (İstanbul Ticaret Üniversitesi Fen Bilimleri Dergisi Yıl: 9 Sayı: 18 Güz 2010 s. 95-105)
- [3] <http://tr.wikipedia.org/wiki/Ontoloji>
- [4] http://tr.wikipedia.org/wiki/Web_Ontoloji_Dili
- [5] <http://owl.cs.manchester.ac.uk/tutorials/protegeowltutorial/>
- [6] Matthew Horridge; *A Practical Guide To Building OWL Ontologies Using Protege 4 and CO-ODE Tools Edition 1.3*
- [7] <http://www.canberkozdemir.com/sparql-nedir/>
- [8] <http://protege.stanford.edu/>