

# Development of Semantic Web Application Architecture for Natural Language Based Querying

Ceren Ocal Tasar<sup>1</sup>[0000-0002-0652-7386], Murat Komesli<sup>2</sup>[0000-0002-8240-5540],  
Murat Osman Unalir<sup>3</sup>[0000-0003-4531-0566]

<sup>1</sup> Yasar University, Software Engineering Department  
Universite Cad. 35, 35100, Bornova, Izmir, Turkey  
ceren.ocal@yasar.edu.tr

<sup>2</sup> Yasar University, Management Information Systems Department  
Universite Cad. 35, 35100, Bornova, Izmir, Turkey  
murat.komesli@yasar.edu.tr

<sup>3</sup> Ege University, Computer Engineering Department  
Universite Cad. 9, 35100, Bornova, Izmir, Turkey  
murat.osman.unalir@ege.edu.tr

**Abstract.** With the developing technology and increased demand on practical usage of knowledge, users require to interact with systems by using natural language. Researchers discover the necessity of natural language queries should be converted to machine understandable format to meet the requirement of users. The necessity of the conversion triggered the researches on studies on question answering systems that is one of the systems mostly interact with natural language queries. For question answering systems that utilizes structural knowledge sources, understanding user intention accurately consists of the processes of natural language processing, representing the user intention and forming structured query to generate the answer. A technique and an architecture to convert natural language query to ontology query language by utilizing from both linguistic and semantic technologies are proposed in this study. Automatic conversion of an unstructured natural language query to a structured ontology query language SPARQL and process of producing answer over linked data are explained in the proposed technique and architecture. A different approach other than the related studies is followed in terms of defining a “Query Semantization” layer in the proposed architecture and developing a natural language aware ontology for the technique proposed. Contribution to the literature is expected by combining natural language processing technique and semantic web technologies with a different approach.

**Keywords:** Natural Language Processing, Ontology Development, Semantic Web Application Architecture.

# Doğal Dil Tabanlı Sorgulama İçin Anlamsal Web Uygulama Mimarisinin Geliştirilmesi

**Özet.** Gelişen teknoloji ve giderek artan bilginin pratik kullanım ihtiyacı doğrultusunda kullanıcıların sistemler ile doğal dil kullanarak etkileşime geçme ihtiyacı doğmuştur. Bu ihtiyacın karşılanması için doğal dil sorgularının makine tarafından anlaşılır bir yapıya dönüştürülmesi gerekliliği araştırmacılar tarafından gözlemlenmiştir. Bu gereklilik doğal dil sorgusu ile en çok etkileşime geçen sistemlerden biri olan soru cevaplama sistemleri alanındaki birçok araştırmayı tetiklemiştir. Yapısal bilgi kaynakları kullanan soru cevaplama sistemleri için kullanıcının bilgi ihtiyacının doğru olarak anlaşılması süreci doğal dil sorgusunun işlenmesi, kullanıcının bilgi ihtiyacının temsili ve yapısal sorgu üretilerek cevabın oluşturulması süreçlerini kapsamaktadır. Bu çalışmada doğal dil sorgusunun, hem dilbilimsel, hem de anlamsal teknolojilerden yararlanılarak ontoloji sorgu diline çevrilmesi için bir yöntem ile mimari önerilmiştir. Ontolojilerin bilgi kaynağı olarak kullanıldığı önerilen yöntem ve mimaride, yapısal olmayan doğal dil sorgusunun otomatik olarak yapısal bir sorgu dili olan SPARQL'a dönüştürülmesi ve üretilen sorgu ile bağlı veri üzerinden cevap üretilmesi süreçleri açıklanmıştır. Önerilen mimaride açıklanan "Sorgu Anlamsallaştırma" katmanı, bu katmanda tanıtılan işlemler ve doğal dil farkında bir ontoloji geliştirilerek tasarlanan yöntem ile ilgili çalışmalardan farklı bir yaklaşım izlenmiştir. Doğal dil işleme teknikleri ve anlamsal web teknolojilerinin birleştirilerek tasarlanan bu farklı yaklaşım ile literatüre katkı hedeflenmektedir.

**Anahtar Kelimeler:** Doğal Dil İşleme, Ontoloji Geliştirme, Anlamsal Web Uygulama Mimarisi.

## 1 Giriş

Teknolojinin hayatımızdaki önemi ve depolanan veri miktarı çoğaldıkça, son kullanıcılar için bilgiye daha hızlı ve kolay erişim sürekli olarak çözümü iyileştirilmeye çalışılan bir probleme dönüşmüştür. Kullanıcıların kendilerini en rahat hissettikleri sorgu tipi günlük hayatta iletişim için kullandıkları doğal dildir. Doğal dille temsil edilen kullanıcının bilgi ihtiyacının makine tarafından anlaşılır bir forma dönüştürülmesi gerekmektedir. Bu alanda çalışan araştırmacılar, doğal dilin ayrıştırılması, her bir ögenin analizi, makine anlaşılır bir temsile çevrilmesi ve otomatik olarak cevap üretilmesi süreçlerine yoğunlaşmaya başlamışlardır. Bu gelişmeler doğrultusunda yapay zekânın bir alt birimi olan doğal dil işleme (DDİ) konusu popülerlik kazanmıştır. DDİ; verilen bir doğal dil sorgusunun bileşenlerine ayrıldığı, her bileşenin ve bileşenler arası ilişkilerin dilbilimsel olarak analiz edildiği süreçlerin bütünüdür. DDİ, verilen

sorgu ile ilgili detaylı sonuçlar çıkarsa da anlamsal analiz ile ilgili yetersiz kaldığı noktalar bulunmaktadır. Dilbilimsel analiz sürecini anlama ve uzlaşma konusunda bir üst seviyeye çıkartabilecek ve onu destekleyecek anlamsal teknolojilere ihtiyaç vardır.

Anlamsal teknolojilerin arkasındaki temel bileşen, belirli bir alanda tanımlanan kavramlar arasındaki ilişkiler ve taksonomi ile kavramsallaştırmayı temsil eden ontolojilerdir. Farklı platformlar arasında bilgi paylaşımı ve değişimi yapılabilmesinde önemli rol oynarlar. Bu sebepten dolayı ontoloji kavramı, sözdizimi, OWL ya da RDF ile temsil edilmesi ve ontoloji sorgu dilleri olan RDQL, SPARQL konularındaki çalışmalarda artış görülmektedir [1]. Anlamsal teknolojilerin yaygınlaşması ile DDİ alanındaki gelişmeler yeni bir yön kazanmıştır. Araştırmacılar arasında kullanıcıyı anlamak ve sorularına cevap üretmek için DDİ tekniklerini anlamsal teknolojiler ile birleştiren melez yöntemler tartışılmaya başlanmıştır. Bu aşamada ilk odaklanılan konu anlamsal teknolojilerin DDİ yöntemlerini güçlendirmeye nasıl yardımcı olacağı olmuştur. İkinci önemli konu ise, anlamsal web geliştirmede kullanılan ontoloji öğrenmesi, ontoloji sorgulama, çok dilli ontoloji eşleştirme yöntemlerinde DDİ'nin hangi noktalarda katkı sağlayabileceği olmuştur [2].

Cümle anlama, içerisinde temel olarak iki süreç barındırır. Bunlardan ilki morfolojik olarak incelenmesi, ikincisi ise anlamsal analizdir [3]. Morfolojik analiz kelimelerin yapılarıyla ilgilenirken anlamsallığa ulaşmak için kelimeler arası ilişkiler, kelimelerin POS ("part of speech", kelime türü) etiketleri (isim, fiil, sıfat, vb.) ve cümle içindeki görevlerini (özne, yüklem, nesne, vb.) çıktı olarak üretir. Anlamsallık seviyesine katkıda bulunmak için, varlık ismi tanıma yönteminden faydalanılır. Bu yöntem doğal dil sorgusunda bulunan varlıkların önceden tanımlanmış kategorilerle (kişi, organizasyon, tarih, lokasyon, vb.) işaretlenmesini sağlar [4]. Üretilen bu çıktılar, doğal dil sorgusunun zenginleştirilmesi için ontolojilerle birleştirilerek kullanılır. Bu yaklaşım özellikle doğal dil ile sorgulama imkanı sağlayan soru cevaplama sistemleri için ontolojilerin bilgi kaynağı olarak kullanılmasını sağlar. Doğal dil sorgusu ile temsil edilen bilgi ihtiyacı ontolojinin sorgu diline dönüştürülür. Kullanıcılar ontoloji sorgu dilini, yapıyı ve ontoloji sözlüğünü öğrenme ihtiyacı olmaksızın ontolojilerdeki bilgiye ulaşabilirler. Kullanıcılar ile sistemler arası köprü kuran bu yaklaşım sayesinde ontolojilerin pratik kullanımı da sağlanmış olmaktadır [5].

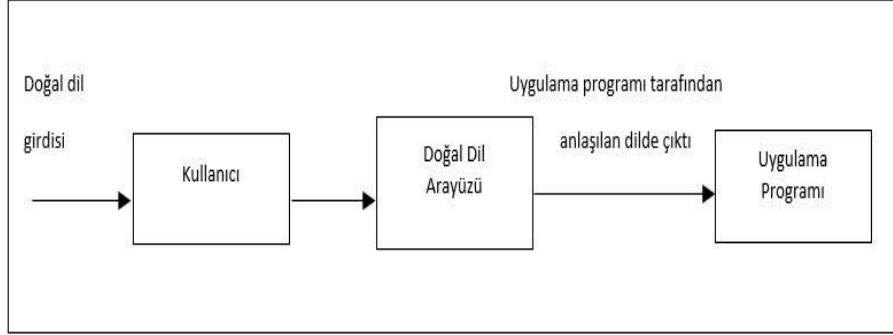
Bu çalışmada DDİ teknikleri ve anlamsal web teknolojilerinin birleştirilmesi yaklaşımı ile tasarlanan mimari sunulmuştur. Bu bölümde konuya bir giriş yapılarak Bölüm 2'de temel kavramlar açıklanmış ve benzer çalışmalar sunulmuştur. Bölüm 3'de doğal dil farkında ontoloji geliştirme açıklanmıştır. Bölüm 4'de, doğal dil sorgularının otomatik olarak anlamsal web sorgularına dönüştürülmesi ve çalışmada önerilen mimari açıklanmıştır. Son bölümde ise, çalışmanın katkısı, diğer yaklaşımlar ile farkları ve benzerlikleri, sağlayacağı katkılar ve bu katkıların genişletilerek gelecek çalışmalara hangi şekilde yol gösterebileceği tartışılmıştır.

## 2 Temel Kavramlar ve İlgili Çalışmalar

### 2.1 Doğal Dil İşleme

Yapay zeka ve hesaplamalı dilbilim konularına ait bir araştırma konusu olan doğal dil işlemenin temel konusu içeriği çözümlenme, kelime yapılarının ve kelimeler arası ilişkilerin anlaşılması için kelime, cümle ve metin işlemedir. Doğal dil işlemenin temel prensibi; doğal dil verisinin dilbilim yöntemleri kullanılarak zenginleştirilmesi olarak özetlenebilir. Doğal dilde kullanılan olası kelime sayısı oldukça fazladır ve verilen bir kelime için belirli eklerle birleştirilip anlamları birbirinden farklı kelimeler türetilmesi olasıdır. Bu sebepten dolayı doğal dilde bulunan kelime kümesinin geniş ve karmaşık olduğu görülmektedir. [6].

Doğal dil girdisi kabul ederek, makine tarafından anlaşılır çıktıya dönüştüren doğal dil işleme sistemlerinin blok diyagramı Şekil 1’de gösterilmiştir.



Şekil 1. Bir DDİ sistemi için sistem blok diyagramı [7].

Hayes ve Carbonell [8], doğal dil işlemeyi, doğal dille iletişim için hesaplama etkili mekanizmaların formüle edildiği ve araştırıldığı alan olarak tanımlamıştır. Doğal dil işlemenin temel motivasyonu insan dilinin modellenmesi ve bilgisayar sistemleri tarafından oluşturulan modellerin verimli ve etkin bir şekilde kullanılması olarak belirtilmiştir [8].

Liddy [9], doğal dil işlemeyi bir dizi görev veya uygulama için insan benzeri dil işlemenin gerçekleşmesi amacıyla metnin bir ya da daha fazla seviyede dilbilimsel analizi ve temsili için bir dizi teorik hesaplama teknikleri olarak tanımlamıştır. Bu tanıma yaparken bir dizi hesaplama teknikleri kavramını birden fazla seviyede çeşitli dilbilimsel analiz tekniklerini temsil etmek adına vurgulamıştır. Her bir dilbilimsel analiz seviyesi anlama ulaşmak için ayrı bir görev üstlenir. Bu sebepten dolayı sistemlerin daha doğru ve etkin sonuçlar üretebilmesi için uygulama ihtiyacına göre hangi seviye ya da seviyelerde analiz yapılması gerektiğinin iyi belirlenmesi gerekmektedir [9].

### 2.2 Ontoloji Geliştirme

Felsefe bilimi ontolojiyi varoluşun konusu [10] ve belirli bir alanda geçerli olan varoluşun kategorize edilmesi [11] olarak tanımlar. Bilgisayar bilimleri ve bilgi bilimleri

alanında ontolojilerde bilgi, kavramlardan oluşan bir küme ve bu kavramlar arasındaki ilişki tanımları ile temsil edilir. Bir başka deyişle, belirli bir alanı, kavramların tanımları, özellikleri ve ilişkileri ile modelleyen ortak bir sözlük paylaşımı ontolojiler tarafından yapısal bir şekilde sunulur.

Gruber'in [12] tanımında, ontolojiler paylaşılan kavramsallaştırmanın resmi ve açık tanımlamaları olarak ifade edilir. Kavramsallaştırma; belirli bir amaç doğrultusunda dünyanın soyut ve sadeleştirilmiş bir görüntüsüdür. Gruber [13], bir diğer çalışmada, kavramsallaştırmanın her bilgi tabanlı sisteminde açık veya kapalı şekilde var olduğuna değinmiştir. Tanımda kullanılan ikinci önemli ifade olan resmi ve açık tanımlama, ontolojinin makine tarafından okunur/anlaşılır olabilmesi için tanımlayıcı, açık ve belirli bir standartlarda temsil edilmesi gerektiğine atıfta bulunmaktadır.

Hendler [14], ontolojiyi sözlük, anlamsal ilişkiler, belirli bir alana özel tanımlı bazı çıkarsama ve mantık kurallarından oluşan terimler kümesi olarak tanımlar. Hendler, mantıksal kurallar ve çıkarsama mekanizması kullanılarak kavramlar arasında anlamsal tanımlamalar ile bağlantı kurulduğuna dikkat çeker.

Kalfoglou'na [15] göre, ontoloji, bazı ilgi alanlarındaki önemli kavramların ortak bir anlayışının açık bir temsilidir. Nesnellik kavramı ortak anlayış amacını destekler ve nesnellığe ulaşabilmek için açık kavramsal bir yapı kullanıldığını belirtmiştir. Ontolojinin nesnel olma özelliği, akıllı ajanlar ve uygulamalara bilgi paylaşımı ve tekrar kullanımını sayesinde anlamsal uyumluluk ve birlikte çalışabilirlik özelliğini sağlamıştır.

Noy and McGuinness [16]' in ontoloji geliştirme 101'de tanımladığı adımlar ise şu şekilde sıralanmaktadır:

- Adım 1: Ontoloji alanının ve kapsamının belirlenmesi
- Adım 2: Hâlihazır ontolojilerin yeniden kullanımının gözden geçirilmesi
- Adım 3: Ontolojideki önemli terimlerin belirtilmesi
- Adım 4: Sınıfların ve sınıf hiyerarşisinin tanımlanması
- Adım 5: Sınıfların özelliklerinin tanımlanması
- Adım 6: Özellikler ile ilgili nitelik ve kısıtların belirlenmesi
- Adım 7: Sınıf örneklerinin oluşturulması.

Bölüm 3'de bu adımlar izlenerek doğal dil tabanlı sorgulama için doğal dil farkında ontoloji geliştirilmesi açıklanmıştır.

### 2.3 Anlamsal Web Teknolojileri ile Ontoloji Kullanarak Soru Cevaplama

Bilgisayarlar tarafından doğal dil anlama ve yorumlamanın hayata geçirilmesi için ses ve dil işleme, insan dili teknolojileri, doğal dil işleme, hesaplamalı dilbilim ve konuşma tanıma/sentezleme gibi birçok farklı disiplini bir araya getirmek gerekmektedir. Temel hedef, insan-makine ve insan-insan arasındaki iletişimi iyileştirmek ve metin ya da ses verisinin kullanışlı bir şekilde işlenebilmesini sağlamaktır. Soru cevaplama sistemleri ise, doğal dil sorgularını yapısal sorgulara dönüştüren, bu sorgularına cevaplar üreten ve doğruluk oranları çeşitli metrikler ile ölçülebilen bu hedefe katkı sağlayan sistemlerdir.

Bir soru cevaplama sistemi 3 temel sürece sahiptir. Bu süreçler doğal dil girdisinin işlenmesi, kullanıcının bilgi ihtiyacının anlaşılması ve bu doğrultuda cevap üretilmesi

şeklinde sıralanır. Soru cevaplama sistemleri 2 farklı temel yaklaşımda görülmektedir. Bunlardan ilki bilgi geri getirmeye tabanlı soru cevaplama, diğeri ise, bilgi tabanına dayalı soru cevaplama [17]. Cevap üretmek için bilgi geri getirmeye tekniklerini kullanan ilk yaklaşım, metin tabanlı soru cevaplama ya da belge tabanlı soru cevaplama olarak da bilinir. Verilen bir doğal dil sorgusu için cevabı içerebilecek aday belgelerin oluşturulması bir arama mekanizması kullanılarak gerçekleştirilir.

İkinci yaklaşım olan bilgi tabanlı soru cevaplama sistemlerinde, odak noktası yapısal olmayan doğal dil sorgularının yapısal olan bir bilgi tabanı sorgu diline çevrilmesi esas alınmaktadır. Bilgi tabanlı soru cevaplama sistemleri kural tabanlı, gözetimli ya da yarı gözetimli yöntemler kullanırlar. Kural tabanlı yöntemler sorudan ilişkileri çıkarmak için tanımlı kurallar ve örüntüler kullanırlar. Gözetimli yöntemler ise, mantıksal form ile eşleşmiş soru kümelerinden oluşan işaretlenmiş eğitim veri setlerini kullanan öğrenme algoritmalarından faydalanırlar. Gözetimli öğrenme algoritmaları genel olarak doğal dil sorgu kavramları ile bilgi kaynaklarında tanımlı kavramların eşleştirilme süreci için ayrıştırma ağaçlarını kullanırlar [17].

Anlamsal web teknolojilerinin giderek gelişmesi bağlı verinin artışı sağlamıştır. Bu artış beraberinde gerçek hayat ihtiyacını temsil eden daha geniş bir spektrumda bağlı veri kullanımı sağlamıştır [1]. Ayrıca bağlı veri artışı, birbiriyle bağlantılı RDF veri setlerine sahip olan bağlı veri depolama kaynaklarına erişim ve kullanım ihtiyacını da artırmaktadır. Bu ihtiyaç doğrultusunda araştırmacılar arasında kullanıcıların yapısal bir sorgu dili bilmeden bağlı veri depolayan ontolojilerden nasıl yararlanacakları yeni bir odak noktası olmuştur [18]. Anlamsal teknolojilerin arka planındaki karmaşık rasyonellikten dolayı son kullanıcılar ve bağlı veri arasında bir boşluk bulunmaktadır. Bu boşluğu kapatacak olan yaklaşım, doğal dil arayüzleridir. Araştırmacılar kullanıcıların rahatlıkla kullanabileceği, kullanıcı dostu ve pratik arayüz gereksinimi üzerinde yoğunlaşmışlardır. Bunlara ek olarak, bu arayüzlerin bağlı veri üzerinden soru cevaplama yapan sistemler için temel bir bileşen olması gerektiğine dikkat çekmişlerdir [19].

## 2.4 İlgili Çalışmalar

Xu vd. [20], doğal dil sorgularını SPARQL sorgularını çeviren ve sorguları DBpedia'ya gönderen Xser adında bir soru cevaplama sistemi geliştirmişlerdir. İki aşamalı olan bu yöntemin ilk aşamasında cümlede bulunan varlıklar, kategoriler, ilişkiler veya değişkenler gibi ifadeler tespit edilmiştir. İfadeler arası "predicate argument" ilişkileri sorgu örüntü yapıları şeklinde temsil edilmesi ve tahminleme yapılması için kullanılmıştır.

Dima [21] tarafından sunulan çalışmada, kullanıcının bilgi ihtiyacının RDF üçlü ifadeleri ile eşleşmesi için hem sözdizimsel, hem de anlamsal bilgiyi kullanmaktadır. Bu sistemin temel amacı, anahtar kelime tabanlı arama için alternatif bir yaklaşım sunarak, bir cevap yerine bir doküman listesi döndüren bu arama mekanizmasının yarattığı problemleri çözmektir. Intui3 olarak adlandırılan bu arama yaklaşımı, çok dilli bir soru cevaplama platformudur. Sistemde varlıklar ve yüklemeler endekslenmektedir. Yüklem endeksinin kaynağı 49,714 yüklem ve varlıktan oluşan DBpedia'dır.

He vd.'nin [22] CASIA@V2 şeklinde adlandırdığı soru cevaplama sistemi, cevapların bağlı veri üzerinden üretildiği ve DBpedia'nın veri kaynağı olarak

kullanıldığı bir mimari ile tasarlanmıştır. İfadelerin ya da ifade gruplarının tespiti ve DBpedia ile eşleştirilmesi için ortak öğrenme çerçevesi sunan Markov Mantık Ağları [23] yöntemi kullanılmıştır. Doğal dil sorgusunu bileşenlerine ayırmak için sistemdeki yapılan ilk işlem aday ifadelerin tespit edilmesidir. DBpedia içerisindeki kavramların aday ifadelerle olan olası eşlemelerini bularak RDF üçlüleri yaratmak için ayrıştırılan ifadeler ile DBpedia elemanları arasında gruplama yöntemi kullanılır.

Park vd. [24] tarafından ISOFT isimli çalışmada, doğal dil cümlelerinin SPARQL örüntüleri ile eşleştirilmesi için yararlanılan anlamsal benzerlik yöntemi üzerinde durulmuştur. Sorguda anlam içeren kelimeler seçilmiş ve ontolojideki kavramların temsil edildiği tanımlı URI'lar ile eşleştirilmiştir. Anlam içeren kelimeler cümle içerisinde ipucu içeren ve cevabı üretebilmek amacıyla anlama büyük katkı sağlayan ifadeleri temsil etmektedir.

Usbeck ve Ngomo [25], HAWK adını verdikleri hibrit bir soru cevaplama platformu sunmuşlardır. Kelime segmentasyonu, POS etiketlemesi ve geçiş temelli bağıllık ayrıştırma yöntemlerini kullanarak bir yazılım zinciri geliştirilmiştir. Cümlede ayrıştırılan her bir öge daha sonra anlamsal açıklayıcı şekilde kullanılması için POS etiketlemesi işleminden geçirilmiştir. Kolektif öğrenme tabanlı bir bilgi çıkarımı çerçevesi olan FOX [26] kullanılarak cümledeki varlıklar ve aralarındaki ilişkiler tespit edilmiştir. Bağıllık analizi ve isim tamlaması bulunması ile anlamsal rol işaretleme yöntemleri birleştirilerek sözdizimsel ve anlamsal ilişkilerin yakalanması sağlanmıştır.

Ruseti vd. [27], QAnswer adında doğal dil sorgularını DBpedia üzerinden cevaplayan bir soru cevaplama çerçevesi geliştirmişlerdir. Doğal dilden SPARQL sorgu diline dönüştürülen sorgular açık kaynak kodlu bir platform olan Virtuoso [28] üzerinde geliştirilmiştir. Çalışmalarında gelen sorgu iki aşamalı bir işlemde gerçekleştirilmektedir. İlk aşamada doğal dil sorgusundaki DBpedia varlıkları tespit edilir. İkinci aşamada ise, yönlü grafik oluşturmak için bu varlıklar birbirleriyle uygun ve doğru sözdizimsel bağıllık tipi temsil edilecek şekilde bağlanır.

Beaumont vd. [29] tarafından yapılan çalışmada, bilgi tabanlı bir soru cevaplama sistemi olan SemGraphQA kolay bir erişim ile kullanıcılarının doğal dil kullanarak ontoloji tabanlı bir platformda sorgulama yapabilmelerine olanak sağlamaktadır. Çalışmalarındaki odak nokta, doğal dil sorgusunun grafik temsili olarak gösterilmesidir. Doğal dil sorgusunu olası anlamların temsil edildiği varlıkların düğümler kenarların ise ilişkileri ile gösterildiği anlamsal çizgelere çevrilir. Soruların anlamsal analizi için kullanılan bu çizge gösterimi gözetimsiz öğrenme tekniğini kullanır. Oluşan bu çizgeler içerisinde bir ilgililik sıralama algoritması kullanılarak doğal dil sorgusuna en yakın olan bulunur ve anlamsal çizge üzerinden SPARQL formüle edilir.

Baudis ve Sedivy [30], YodaQA adını verdikleri çalışmalarında modüller ve açık kaynak kodlu bir soru cevaplama sistemi sunmuşlardır. DDİ işleme teknikleri kullanılarak doğal dil sorgusu üzerinden üç adet soru cevaplama özelliği aranır. Bunlardan ilki ipuçları olarak adlandırdıkları sorgunun kapsamını belirlemede önemli bir rol oynayan ifadelerdir. İkincisi ise, cümlenin odak noktası olan, kullanıcının bilgi ihtiyacını temsil eden odak olarak adlandırdıkları sorgulanan nesnedir. Odak bulunurken bağıllık analizi yöntemine sezgisellik entegre edilmiştir. Son özellik ise, odaktan türetilen cevap tipi tanımlamasını içeren sözcük cevap tipidir. Çıkarılan ipuçları olası cevapları belirlemek için cevap üretilirken kullanılır.

İlgili çalışmalar incelendiğinde, genel olarak DDİ yöntemlerinin doğal dil sorgusu analizinde kullanıldığı ve analiz çıktılarının çeşitli yöntemler ile birleştirilerek ontolojide kavramlar ile eşleştirildiği görülmektedir. Önerilen diğer çalışmalarda ontoloji, mimari bileşen anlamında bir bilgi tabanı olarak kullanılmıştır ve kullanılan ontolojiler halihazırda geliştirilmiş ontolojilerdir. Bu çalışmada farklı bir yaklaşım üzerinde durularak ilk nokta olarak doğal dil farkında ontoloji geliştirilmeye dikkat çekilmiştir. Daha sonra ise DDİ çıktılarının ontoloji üzerindeki karşılıkları ve öğeler arası gizli kalan ilişkiler bulunarak bir örüntü yardımıyla ontoloji sorgu dili formüle edilmiştir. DDİ çıktıları üzerine ontolojinin çıkarsama kabiliyetinden de yararlanılmıştır.

### 3 Doğal Dil Farkında Ontolojilerin Geliştirilmesi

Doğal dil farkında ontolojilerin geliştirilmesi açıklanmadan önce bu konudaki en önemli kavram olan “doğal dil farkında ontoloji” ifadesini açıklamak gerekmektedir. Ontolojiler doğaları gereği belirli bir kapsama, alana hitap etmektedirler. Dolayısıyla hitap edilen bu alanda hangi doğal dil cümlelerine cevap bulunabileceği ve hangi kavramların tanımlı olduğu belirlenmektedir. Noy and McGuinness [16]’ in ontoloji geliştirme 101’de tanımladığı ilk adım, ontoloji alanının ve kapsamının belirlenmesidir. Ontolojinin doğal dil farkında olma özelliği için bu adım kritiktir. Noy and McGuinness [16] bu adımı şu şekilde tanımlar: Alan ve kapsam belirleyerek ontolojinin sınırlarının çizildiği ontoloji geliştirmedeki ilk adımdır. Bu adımı başarıyla tasarlayabilmek için ontoloji geliştiricilerin şu soruları cevaplandırmaları gerekmektedir:

- Ontoloji hangi alanı kapsayacak?
- Ontoloji ne için kullanılacak?
- Ontoloji tarafından hangi tip sorular cevaplanacaktır?

Geliştirilecek ontolojide cevaplanacak kapsama örnek teşkil etmesi açısından örnek soruların tanımlanması gerekmektedir. Doğal dilde tanımlanan bu örnek sorular ontolojinin doğal dil farkında olmasına olanak sağlamaktadır. Bu çalışmanın temel prensibi ontolojinin bu ilk adımı göz önünde bulundurularak örnek sorularda belirtilen doğal dil ifadeleri ile ontolojideki kavram tanımlamalarının eşleştirilmesidir. Kapsamı belirleme ve kavramsallaştırma süreçlerinin temeli bu eşleştirme üzerine konumlandırıldığında gelen doğal dil sorgularının cevaplarının geliştirilen ontolojide aranması kolaylaşacaktır.

Ontoloji geliştirme 101’de belirtilen halihazır ontolojilerin yeniden kullanımının gözden geçirilmesi olan 2. adım yeni bir ontoloji yaratılacağı için atlanmıştır. 3. adım ise, ontolojideki önemli terimlerin belirlenmesidir. Doğal dil farkında özelliğinin kazandırılması adımı ile de örtüşmektedir. Bir doğal dil sorgusu içerisindeki önemli terimler varlıkları, tamlamaları, çeşitli ifadeleri içerebilir ve genellikle sorgu cevabının oluşturulmasında katkıları büyüktür. Ontoloji için tanımlanan bu önemli kelimelere örnek soru tanımlamalarında rastlanacaktır. Analiz edilecek doğal dil sorgusunda ise, ya bir varlık, ya da cümle içerisinde cevabın bulunmasına katkı sağlayan önemli bir öğe (özne, yüklem ya da nesne) olarak rastlamak mümkündür. Tüm bu kavramların doğal



dil farkında olması ile eşleşmenin ve cevap üretilmesinin daha kolay yapılması sağlanacaktır.

Bu önerinin daha anlaşılır olması için birkaç ontoloji örnek sorusu ve doğal dil sorusu ile örneklendirmek gerekmektedir. Coğrafya alanında kapsamı mekânsal bir sentez olarak Türkiye’yi ele alan bir coğrafi ontolojiyi ele alalım. Tanımlanan örnek sorular:

- Türkiye’deki şehirleri gösterir misin?
- Konya’nın komşularını gösterir misin?
- Bursa ili hangi coğrafi bölgededir?
- Marmara Bölgesi’ndeki şehirlerin nüfusları kaçtır?
- Antalya’nın en yüksek dağı hangisidir?
- Türkiye’de en fazla yağış alan il hangisidir?

Bu örnek sorgulardan doğal dilde kavramlar incelenip önemli kelimeleri bir kümeye koyduğumuzu varsayalım. Elimizdeki kümede {Türkiye, şehir, Konya, komşu, Bursa, bölge, Marmara Bölgesi, nüfus, Antalya, yükseklik, dağ, yağış, il} elemanları bulunmaktadır. Geliştirilen ontolojide bu elemanlar ya bir sınıf, ya bir birey ya da bir özelliği temsil edecektir. Bu aşamada 4. adım olarak sınıfların ve sınıf hiyerarşisinin tanımlanması ele alınmaktadır. Önemli terimlerin içerisinde genel olan ifade tanımlamalarından ya da gruplamalarından olası sınıflar ya da alt sınıflar tespit edilir. Örneğin; verilen örneklerde şehir, bölge, dağ genel tanımlamaları adreslemektedirler. Dolayısıyla ontoloji içerisinde sınıf tanımlaması olarak kullanılacaklardır. Sınıf tanımlamaları ve hiyerarşi belirlendikten sonra 5. adım olan sınıfların özelliklerinin tanımlanması uygulanır. Bulunan terimler içerisinde ontolojideki varlıkların niteliklerini belirtebilecek ve aralarında ilişki kurmamıza yarayacak olan kavramlar özellikler olarak tanımlanır. 6. adım, 5. adımın ön koşuludur ve listelenen özelliklerinin nitelik ve kısıtlarının belirlenmesidir. Örneğin; komşuluk ilişkisine simetrik olması şeklinde bir nitelik tanımlanabilir. 7. Ve son adım olan sınıf örneklerinin oluşturulması için ontolojide tanımlı olacak bireyler belirlenir. Örneğin; “şehir” ontoloji içerisinde bir sınıf olarak ve “Bursa, Konya, Antalya” bu sınıfın bireyleri olarak tanımlanır. Verilen sınıf örneklerini kapsam genişletmek adına arttırmak ontoloji içerisinde mümkündür. Bu örnek için diğer şehirler eklenerek olası özellikler kontrol edilerek yeni örnekler türetilebilir.

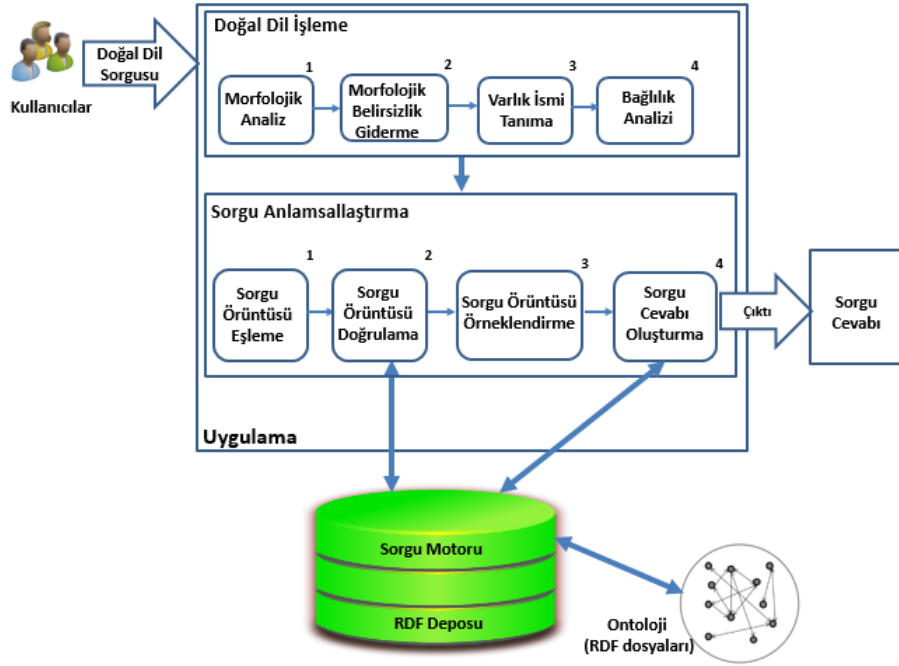
Ontoloji Geliştirme 101’de de belirtildiği üzere sınıflar, alt sınıflar ve özellikler tanımlandığı zaman artık bu kavramların özellikleri kullanarak birbirine bağlanması yani anlamlandırılması gerekmektedir. Örneğin; “komşu” şeklinde tanımlanan bir ilişki “şehir” sınıfının bir örneği olan “Konya” için tanımlanır. Bu ilişkide Konya ile komşu olan diğer şehirler eklenir. Bu kavramları içeren soruların cevapları artık ontolojide bulunmaktadır. Ontolojiden bu cevapları döndürebilmek için verilen sorguyu bir SPARQL sorgusuna dönüştüren bir doğal dil arayüzü gerekmektedir. 4. Bölümde bu dönüştürme işleminin detayları aktarılmıştır. Bunlara ek olarak, bu önemli elemanlar cümle doğal dil işleme teknikleri kullanılarak analiz edildiğinde anlama katkısı sağlayan öğeler olarak karşımıza çıkmaktadır. Morfolojik analiz ve belirsizlik giderme, varlık ismi tanıma, bağlılık analizi gibi doğal dil işleme yöntemleri ile cümlede varlıklar ve kelimeler arası ilişkiler elde edilip ontolojide bu öğelerin tanımları ve tipleri bulunarak SPARQL sorgusunun oluşturulması için gerekli altyapı bu çalışmada önerilmiştir.

Doğal dili tanıyan, formüle edebilecek yetenekte bir ontoloji doğal dilde gelen sorguların bir arayüz yardımıyla ontoloji sorgu diline çevrilmesi ile cevap üretebilen bir çerçeve oluşumunu sağlamaktadır. Gelen doğal dil sorgusunun otomatik olarak SPARQL diline çevrilmesi için önerilen yöntem ve mimari 4. Bölüm’de açıklanmıştır.

#### 4 Önerilen Yöntem ve Mimari: Doğal Dil Sorgularının Otomatik Olarak Anlamsal Web Sorgularına Dönüştürülmesi

Önerilen anlamsal web uygulaması doğal dilde sorguları kabul eden ve bilgi kaynağı olarak ontolojiyi kullanan bir platformdur. Ontolojilerin bilgi kaynağı olarak kullanılmasından dolayı doğal dil sorgusunun bir doğal dil arayüzü yardımıyla SPARQL sorgusu ile temsil edilmesi sağlanacaktır. Bu bölümde önerilen yöntem ve mimarinin tasarlanması birbirini besleyen süreçlerdir. Bu bölümün daha anlaşılır olması adına önerilen yöntem, önerilen mimari bileşenleri üzerinden açıklanmaktadır.

Doğal dil sorgusunun ilk olarak “Doğal Dil İşleme” katmanı ile bir ön işlemeden geçirilip daha sonra ise “Sorgu Anlamsallaştırma” katmanında sorguya cevap üretilmesi şeklinde önerilen sistem mimarisi Şekil 2’de gösterilmiştir.



Şekil 2. Önerilen sistem mimarisi.

Doğal dil arayüzüne gelen doğal dil girdisi ilk olarak doğal dil işleme teknikleri kullanılarak işlenir. Morfolojik analiz ve belirsizlik giderme, varlık ismi tanıma ve bağıllık analizi yöntemlerine sırasıyla başvurulur. Morfolojik analiz ve belirsizlik giderme cümledeki her kelimenin ayrıştırılıp, kelimelerin morfolojik yapılarının POS etiketleri ile çıktısı olarak temsil edildiği bir DDİ tekniğidir. Kelime kökleri, sözcük türleri (isim, fiil, sıfat, vb.) ve varsa ekleri detaylı olarak incelenir. Cümlelerin bir üst seviye analizi için çok kritiktir. Diğer DDİ süreçlerinin sonuçlarını etkiler. POS etiketleri dilbilgisi tanımlayıcıları olarak adlandırılabilirler. Ayrıştırılan her kelimenin POS etiketlerinin belirlenmesi için çeşitli farklı algoritmalar kullanılmaktadır. Cümlelerin doğru yorumlanması ve anlama ulaşılan yolda bir adım öteye taşınması için cümlelerin altında yatan sözdizimsel yapısının çözülmesi gerekir [31].

Varlık ismi tanıma yönteminde ayrıştırılan cümle içerisindeki kelimeler kişi, yer, tarih, organizasyon gibi tanımlayıcılar ya da kategoriler ile etiketlenir. Anlamsal seviyede analizin ilk aşamaları olduğu düşünülebilir. Cümle içerisinde geçen söylemler/varlıklar önceden tanımlanmış kategorilere çözümlenir. Bu kategoriler alan ihtiyacına göre özelleştirilebilir. Uygulamalar tarafından kullanılan çeşitli varlık ismi tanıma algoritmaları mevcuttur. Bu algoritmalar arasında en popüler yaklaşım öğrenme algoritmalarından etkin olarak yararlanılmasıdır [32].

Bağıllık analizi, doğal dil cümlelerini belirli yapı tiplerine yani cümleyi öğelerine ayırarak DDİ alanındaki başlıca araştırma alanlarından biridir. Kelimeler arasındaki çeşitli ilişkileri (özne, yüklem, nesne vb.) bağıllık analizi ile tanımlanır. Kelimeler arasındaki ilişkiler bağıllık yapıları ile temsil edilir. Bağıllık analizi algoritmaları düğüm ve ilişkilerden oluşan bağıllık temsili yöntemlerinden faydalanır. Başlangıç düğümü dışında diğer bütün düğümlerin birbirine bağlı olduğu bir bağıllık yapısında temsil edilir [33]. Bir cümlede fiil tipinde sözcükleri bulmak, diğer kelimeler ile aralarındaki ilişkileri çözümlenmek ve diğer terimlerinin sözcük tiplerini, görevlerini çıkarmak kritiktir. Kelimeler arasındaki ilişkiler anlama giden yolda en doğru yol göstericilerdir. Cümle içerisinde kimin kime ne yaptığı bilgisi genel anlama ulaşılrken her bir terimin kısmi olarak rolünü açıklar [34].

Şekil 2’den de anlaşıldığı üzere doğal dil sorgusu sisteme girdi olarak verilir. Verilen bu sorgu ilk olarak DDİ teknikleri kullanılarak analiz sürecinden geçirilir. Analiz sonrasında cümlede anlam içerisinde katkı sağlayan terimler belirlenir ve ontoloji tarafıyla bulunan terimlerin tanımlamaları tekrar kontrol edilir. Bölüm 3’de belirtildiği üzere doğal dil farkında bir ontoloji yardımıyla anlamsal web uygulaması oluşturmak mümkündür. Örnek bir cümle analizi konunun daha anlaşılır olmasına katkı sağlayacaktır. Örneğin; “Marmara Bölgesi’ndeki şehirlerin nüfuslarını göster” cümlesinin ilk olarak DDİ teknikleriyle işlenmiş çıktıları Tablo 1’ de gösterilmiştir. DDİ analiz ve çıktıları göstermek için ITU Türkçe Doğal Dil İşleme Web Servisi’nden [35] faydalanılmıştır.

DDİ çıktılarına baktığımızda bağıllık analizinin cümledeki kelimelerin ve ilişkilerin çıkarılması için kritik olduğu görülmektedir. Doğal dil sorgusunda kullanıcının bilgi ihtiyacını temsil eden ögenin, hedefin çoğunlukla ya nesne ya özne ya da yüklemde görüldüğü gözlenmektedir. Önerilen yöntemde bu öğeler bulunduktan sonra doğal dil farkında ontoloji ile karşılaştırmasının yapılmasının anlamlı olacağı düşünülecektir.

Örneğin bu cümlede nesne (OBJECT) olarak işaretlenen kök kelime “nüfus” iken yüklem (PREDICATE) olarak işaretlenen kelime kökü “göster” olarak görülmektedir.

**Tablo 1.** DDİ çıktıları

Örnek Sorgu	“Marmara Bölgesi’ndeki şehirlerin nüfuslarını göster”
Morfolojik Analiz	Marmara+Noun+Prop+A3sg+Pnon+Nom bölge+Noun+A3sg+P3sg+Loc^DB+Adj+Rel bölge+Noun+A3sg+P3sg+Loc^DB+Pron+Rel+A3sg+Pnon+Nom şehir+Noun+A3pl+Pnon+Gen şehir+Noun+A3pl+P2sg+Nom nüfus+Noun+A3pl+P3pl+Acc nüfus+Noun+A3pl+P2sg+Acc nüfus+Noun+A3pl+P3sg+Acc nüfus+Noun+A3sg+P3pl+Acc göster+Verb+Pos+Imp+A2sg ?+Punc
Morfolojik Belirsizlik Giderme	Marmara Marmara+Noun+Prop+A3sg+Pnon+Nom Bölgesi’ndeki bölge+Noun+A3sg+P3sg+Loc^DB+Adj+Rel şehirlerin şehir+Noun+A3pl+Pnon+Gen nüfuslarını nüfus+Noun+A3pl+P3pl+Acc göster göster+Verb+Pos+Imp+A2sg ? ?+Punc
Varlık İsmi Tanıma	Marmara Marmara+Noun+Prop+A3sg+Pnon+Nom <b>B-</b> <b>LOCATION</b> Bölgesi’ndeki bölge+Noun+A3sg+P3sg+Loc^DB+Adj+Rel <b>I-</b> <b>LOCATION</b> şehirlerin şehir+Noun+A3pl+Pnon+Gen <b>O</b> nüfuslarını nüfus+Noun+A3pl+P3pl+Acc <b>O</b> göster göster+Verb+Pos+Imp+A2sg <b>O</b> ? ?+Punc
Bağlılık Analizi	1 Marmara Marmara Noun Noun Prop A3sg Pnon Nom _ 2 <b>POSSESSOR</b> 2 Bölgesi’ndeki bölge Noun Noun A3sg P3sg Loc^DB Adj Rel _ <b>5 MODIFIER</b> 3 şehirlerinşehir Noun Noun A3pl Pnon Gen _ 4 <b>POSSESSOR</b> 4 nüfuslarını nüfus Noun Noun A3pl P3pl Acc _ 5 <b>OBJECT</b> 5 göstergösterVerb Verb Pos Imp A2sg _ 0 <b>PREDICATE</b> 6 ? ? Punc Punc _ _ 5 <b>PUNCTUATION</b>

Cümlede herhangi bir öge özne (SUBJECT) olarak işaretlenmemiştir. Her iki öge için doğal dil farkında ontoloji üzerinde tanımlamaları kontrol edilir ve “nüfus” için şehirler ve bölgelerde tanımlı bir özellik tanımlı döndürülür. Bunlara ek olarak, varlık ismi tanıma yöntemi “Marmara Bölgesi” ifadesini lokasyon (LOCATION) olarak

işaretlemiştir. Ontoloji üzerinden bu birey bulunur ve bu bireyin ait olduğu sınıf döndürülür. SPARQL formülasyonunda bu elemanların hepsi bir örüntü üzerine oturtulup SPARQL sorgusunun son haline ulaşılabilecektir. Eğer sadece ontoloji tabanlı bir yaklaşım ile bakılırsa sorgu Marmara Bölgesi'nin nüfusunu hedeflemiştir olarak algılanabilirdi. Bu noktada DDİ teknikleri yardımı koşturmaktadır. Cümlelerin nesnesi ile bağlı olan herhangi bir ögenin var olup olmadığının kontrol edilmesi gerekir. 3 numaralı “şehir” ögesi cümlelerin nesnesine aitlik (POSSESSOR) ilişkisi ile bağlıdır. Bu ilişki yardımıyla kullanıcının bilgi ihtiyacının şehir-nüfus ilişkisi şeklinde temsil edildiği elde edilebilmektedir.

Bir sonraki adımda mimaride sorgu örüntüsü doğrulama olarak gösterilen adım için, cümlelerin öğeleri ile ontolojide tanımlı kavramlar arasında ek bir eşleme kontrolü olacaktır. Cümlede ve ontolojide tanımlı ortak kavramlar arasında olası başka ilişkilerin temsil edilip edilmediğine bakılacaktır. Bu kontrol sonrası verilen örnek için her şehrin bir bölgede konumlanmış olduğu ilişkisi ontolojide tanımlı “konumVar” ilişkisi ile temsil edilmektedir. Bölge ile şehir sınıfları arasında temsil edilen bu özellik o bölgede bulunan şehirlerin elde edilmesini sağlayacaktır. Dolayısıyla bu ilişkinin şehir ve bölge sınıfları arasında SPARQL sorgusuna eklenmesi gerekmektedir. Aşağıda gösterilen örnek bir SPARQL örüntüsünde SPARQL üretmek için ihtiyacımız olan öğeler gösterilmektedir. Sorgu örüntüsü örnekleştirme işlemi sayesinde sorgu elemanları örüntüdeki doğrulanmış eşleme doğrultusunda konumlandırılır.

```
SELECT ?y
WHERE { ?x rdf:type ontoloji_ismi: varlık-ismi-sınıfı.
        ?y rdf:type ontoloji_ismi: hedef-sınıfı.
        ?y özellik: özellik-adı ?x
        FILTER(regex(str(?x), "varlık ismi", "i")) }
```

Örnek cümlemizden elde ettiğimiz SPARQL sorgusu aşağıdaki şekilde olacaktır:

```
SELECT ?x ? nufus-miktari
WHERE { ?x rdf:type geo_turkce: Bölge .
        ?y rdf:type geo_turkce: Şehir .
        ?x ins: konumVar ?y .
        ?y ins: nüfus ?nufus-miktari .
        FILTER(regex(str(?x), "Marmara Bölgesi", "i")) }
```

Uygulama içerisinde 2 temel katman bulunmaktadır. Bunlardan ilk olanı DDİ yöntemlerini sırasıyla uygulayan “Doğal Dil İşleme” katmanıdır. DDİ çoğu soru cevaplama sisteminde de bulunur. Bu çalışmanın temel katkısı “Sorgu Anlamsallaştırma” katmanı olarak temsil edilmiştir. Bir üst katmandan DDİ çıktıları mimarideki sorgu örüntüsü eşleme adımı için kullanılır. Örüntü için gereken sorgu elemanları DDİ çıktıları yorumlanarak elde edilir. Daha sonra, mimaride gösterilen sorgu örüntüsü doğrulama adımı için bulunan varlık isimleri, varlık ismi ve hedef sınıfları ve ilişkiler ontoloji üzerinden tekrar kontrol edilir. Çıkarılma mekanizması sayesinde varlıklar arası olası diğer ilişkiler de tespit edilir. Mimarideki bir sonraki adım olan sorgu örüntüsü örnekleştirme sürecinde elde edilen sorgu elemanları ile sorgu formüle edilir. Sorgu

cevabı oluşturulması adımı için sorgu motoru kullanılarak ontoloji üzerinde sorgu koşturularak cevap üretilir.

## 5 Değerlendirme ve Gelecek Çalışmalar

Çalışma kapsamında doğal dil sorgularının, kullanıcının bilgi ihtiyacını karşılamak için makine anlaşılır bir şekle dönüştürülmesi probleminde çözüm aranmıştır. Dilbilimsel analize yardımcı anlamsal teknolojilerin kullanımı için hibrit bir yaklaşım tanıtılmıştır. Yapısal olmayan doğal dil formundaki sorgulara otomatik olarak yapısal bir form kazandırılarak ontoloji sorgu diline dönüştürüldüğü bir anlamsal web mimarisi önerilmiştir. Ontoloji geliştirme sürecinde üretilen örnek sorulardan ile deneysel bir test çalışması yapılmıştır. Sisteme gönderilen 100 adet örnek doğal dil sorgusundan 78 adedine cevap dönülmüş ve bu cevaplar içerisinde 60 adedi doğru cevap olarak işaretlenmiştir. Deneysel sonuçlar kesinlik değeri (precision), hassasiyet değeri (recall) ve F- ölçüsü değerleri cinsinden %76.92, %60 ve %67.41 olarak hesaplanmıştır.

İlgili çalışmalardan farklı olarak tasarlanan ve katkı sağlanan noktalar ise, mimaride “Sorgu Anlamsallaştırma” katmanında tanıtılan işlemler ve doğal dil farkında bir ontoloji geliştirilerek tasarlanan yöntem ve mimarideki bileşen kullanımından doğan farklılıklardır. İlgili çalışmalarda bilgi kaynağı olarak ontoloji, mimari bileşeni olarak tanıtılmıştır. Bilgi kaynağı olarak kullanılan ontolojiler Freebase, DBpedia gibi halihazırda geliştirilmiş ontolojilerdir. Bu çalışmada, doğal dil farkında bir ontoloji geliştirilerek, geliştirilen ontolojinin DDİ yöntemleri ve anlamsal analiz yöntemleriyle uyumlu çalışması tasarlanan algoritmalar sayesinde mümkün kılınmıştır. İlgili çalışmalardan farklı olarak geliştirilen doğal dil farkında ontoloji sayesinde, DDİ çıktılarının yorumlanması ve sorgu öğeleri arasındaki gizli kalmış ilişkilerin çıkarılarak sorgu örüntüsüyle eşleştirilip ontoloji sorgu dilinin formüle edilmesi sağlanmıştır. Bunlara ek olarak, ontolojinin doğal dil farkında olması bu algoritmaların kapsam alanlarının genişletilebilmesi ve sonuçlarının iyileştirilmesine olanak sağlamıştır. Gelecek çalışmalar için aynı mimari kullanılarak tasarlanacak sistemde doğal dil farkında yeni ontolojiler geliştirilerek doğal dil tabanlı sorgulamanın dilden bağımsız hale dönüştürülmesi fikri önerilmektedir. Bu yaklaşım, birden fazla ontolojinin kullanıldığı bir platformda aynı mimarinin geçerli olması imkanını da sağlayacaktır. Farklı alanlarda doğal dil farkında geliştirilmiş ontolojiler ile doğal dil sorgularına cevap üreten sistemlerin tasarlanması fikri, gelecek çalışmalara yön verebilir.

Önerilen mimaride doğal dil işleme süreçleri için ITU Türkçe Doğal Dil İşleme Web Servisi’nden [35] faydalanılmıştır. Araştırma çalışmaları için ücretsiz kullanılan bu servis, ticari kullanım için lisans gereksinimini beraberinde getirecektir.

Halihazır yöntemin içerisinde tek tip olarak kullanılan sorgu örüntüsünün daha karmaşık doğal dil sorgularını cevaplayabilmesi için yeni örüntüler eklenerek geliştirilmesi ihtiyacı doğacaktır. İçeriğinde birden fazla varlık, karşılaştırmalı ifade ya da uzun tamlamalar bulunduran sorgular için yöntemin makina öğrenmesi yaklaşımı ile iyileştirilmesi ve örüntü tiplerinin genişletilmesine başvurmak gerekecektir. Gelecek çalışmaların hedefine, bu ihtiyaç da dahil edilebilir. Bu çalışmanın bir diğer önemli

katkısı ise, doğal dil farkında ontolojiler geliştirilerek çok dilli soru cevaplama platformu oluşturma imkanı tanımaktadır. Birden fazla dilde geliştirilmiş olan her bir ontolojinin doğal dil farkında olması çok dilli soru cevaplama imkanının sağlanması için kritiktir.

## Kaynaklar

1. Berners-Lee, T., Hendler, J., Lassila, O.: The semantic web. *Scientific American* 284(5), 34-43 (2001).
2. Guo, R., Ren, F.: Towards the relationship between Semantic Web and NLP. In: 2009 IEEE International Conference on Natural Language Processing and Knowledge Engineering. pp. 1-8. IEEE, China (2009).
3. Guo, Y., Li, Y., Shao, Z.: A semantic processing model for sentence understanding based on cognitive learning. In: 2011 IEEE 3rd International Conference on Communication Software and Networks. pp. 106-110, IEEE, China (2011).
4. Collobert, R., Weston, J., Bottou, L., Karlen, M., Kavukcuoglu, K., Kuksa, P.: Natural language processing (almost) from scratch. *Journal of Machine Learning Research*, 12, 2493-2537 (2011). <https://doi.org/10.1.1.231.4614>
5. Bernstein, A., Kaufmann, E., Göhring, A., Kiefer, C.: Querying ontologies: A controlled English interface for end-users. In: International Semantic Web Conference. pp. 112-126. Springer, Berlin, Heidelberg. (2005).
6. Hoque, M. M., Rahman, M. J., KumarDhar, P.: Lexical semantics: A new approach to analyze the Bangla sentence with semantic features. In: 2007 International Conference on Information and Communication Technology. pp. 87-91, IEEE, US (2007).
7. Kumar, E.: Natural language processing. IK International Publishing House. India (2011).
8. Hayes, P. J., Carbonell, J. G.: A tutorial on techniques and applications for natural language processing. Carnegie- Mellon University, (1983).
9. Liddy, E. D.: Natural language processing. *Encyclopedia of Library and Information Science*. 2nd edn. Marcel Dekker, NY (2001).
10. Gašević, D., Djuric, D., Devedžić, V.: Model driven architecture and ontology development. 1st edn. Springer, Berlin, Heidelberg (2006). <https://doi.org/10.1007/3-540-32182-9>
11. Sowa, J. F.: Ontology, metadata, and semiotics. In: International Conference on Conceptual Structures. pp. 55-81, Springer, Berlin, Heidelberg, Germany (2000).
12. Gruber, T. R.: A translation approach to portable ontology specifications. *Knowledge Acquisition*, 5(2), 199-220 (1993). <https://doi.org/10.1006/knac.1993.1008>
13. Gruber, T. R.: Toward principles for the design of ontologies used for knowledge sharing. *International Journal of Human-Computer Studies*, 43(5-6), 907-928 (1995). <https://doi.org/10.1006/ijhc.1995.1081>
14. Hendler, J.: Agents and the semantic web. *IEEE Intelligent Systems*, 16(2), 30-37 (2001). <https://doi.org/10.1109/5254.920597>
15. Kalfoglou, Y.: Exploring ontologies. In *Handbook of Software Engineering and Knowledge Engineering, Volume I: Fundamentals*, 863-887 (2001).
16. Noy, N. F., McGuinness, D. L.: Ontology development 101: A guide to creating your first ontology. Stanford Knowledge Systems Laboratory. (2001). <https://doi.org/10.1016/j.art-med.2004.01.014>
17. Jurafsky, D., Martin, J. H.: Speech and language processing. 2nd edn. Pearson, New Jersey (2014).

18. Unger, C., Freitas, A., Cimiano, P.: An introduction to question answering over linked data. In: Reasoning Web International Summer School. pp. 100-140, Springer, Cham (2014).
19. Lopez, V., Uren, V., Sabou, M., Motta, E.: Is question answering fit for the semantic web?: a survey. *Semantic Web*, 2(2), 125-155 (2011). <https://doi.org/10.3233/SW-2011-0041>
20. Xu, K., Feng, Y., Zhao, D.: Xser@ qald-4: Answering natural language questions via phrasal semantic parsing. In: Working Notes for CLEF 2014 Conference. pp. 15-18, (2014).
21. Dima, C.: Answering natural language questions with Intui3. In: Working Notes for CLEF 2014 Conference. pp. 1201-1211, (2014).
22. He, S., Zhang, Y., Liu, K., Zhao, J.: CASIA@ V2: A MLN-based question answering system over linked data. In: Working Notes for CLEF 2014 Conference. pp. 1249-1259, (2014).
23. Richardson, M., Domingos, P.: Markov logic networks. *Machine Learning*, 62(1-2), 107-136 (2006). <https://doi.org/10.1007/s10994-006-5833-1>
24. Park, S., Kwon, S., Kim, B., Lee, G. G.: ISOFT at QALD-5: Hybrid question answering system over linked data and text data. In: Working Notes for CLEF 2015 Conference. (2015).
25. Usbeck, R., Ngomo, A. C. N.: HAWK@ QALD5-Trying to answer hybrid questions with various simple ranking techniques. In: Working Notes for CLEF 2015 Conference. (2015).
26. Speck, R., Ngomo, A. C. N.: Ensemble learning for named entity recognition. In: International Semantic Web Conference. pp. 519-534, Springer, Cham (2014).
27. Ruseti, S., Mirea, A., Rebedea, T., Trausan-Matu, S.: QAnswer-Enhanced entity matching for question answering over linked data. In: Working Notes for CLEF 2015 Conference. (2015).
28. Dolan-Gavitt, B., Leek, T., Zhivich, M., Giffin, J., Lee, W.: Virtuoso: Narrowing the semantic gap in virtual machine introspection. In: IEEE Symposium on Security & Privacy 2011. pp. 297-312, IEEE, U.S. (2011).
29. Beaumont, R., Grau, B., Ligozat, A. L.: SemGraphQA@ QALD5: LIMSI participation at QALD5@ CLEF. In: Working Notes for CLEF 2015 Conference. (2015).
30. Baudiš, P., Šedivý, J.: Biomedical question answering using the YodaQA system: Prototype notes. In: Working Notes for CLEF 2015 Conference. (2015).
31. Indurkha, N., Damerau, F. J.: Handbook of natural language processing. 2nd edn CRC Press, U.S. (2010).
32. Celikkaya, G., Torunoglu, D., Eryigit, G.: Named entity recognition on real data: a preliminary investigation for Turkish. In: The 7th International Conference on Application of Information and Communication Technologies. pp. 1-5, IEEE, Azerbaijan (2013).
33. Nivre, J. Dependency parsing. *Language and Linguistics Compass*, 4(3), 138-152 (2010). <https://doi.org/10.1111/j.1749-818X.2010.00187.x>
34. Shehata, S., Karray, F., Kamel, M.: Enhancing search engine quality using concept-based text retrieval. In: International Conference on Web Intelligence. pp. 26-32, IEEE Computer Society. (2007).
35. Eryigit, G.: ITU Turkish NLP web service. In: Proceedings of the Demonstrations at the 14th Conference of the European Chapter of the Association for Computational Linguistics. pp. 1-4, Sweden (2014).