T.C.
YILDIZ TEKNİK ÜNİVERSİTESİ
FEN BİLİMLERİ ENSTİTÜSÜ

ONTOLOJİ TABANLI BİLGİ HARİTALAMA YÖNTEMİ

NURGÜL YÜZBAŞIOĞLU

YÜKSEK LİSANS TEZİ
BİLGİSAYAR MÜHENDİSLİĞİ ANABİLİM DALI
BİLGİSAYAR MÜHENDİSLİĞİ PROGRAMI

DANIŞMAN
YRD. DOÇ. DR. MEHMET SIDDIK AKTAŞ

İSTANBUL, 2016

T.C.

YILDIZ TEKNİK ÜNİVERSİTESİ

FEN BİLİMLERİ ENSTİTÜSÜ

ONTOLOJİ TABANLI BİLGİ HARİTALAMA YÖNTEMİ

Nurgül YÜZBAŞIOĞLU tarafından hazırlanan tez çalışması 11.07.2016 tarihinde aşağıdaki jüri tarafından Yıldız Teknik Üniversitesi Fen Bilimleri Enstitüsü Bilgisayar Mühendisliği Anabilim Dalı'nda **YÜKSEK LİSANS TEZİ** olarak kabul edilmiştir.

Tez Danışmanı

Yrd. Doç. Dr. Mehmet Sıddık AKTAŞ

Yıldız Teknik Üniversitesi

Jüri Üyeleri

Yrd. Doç. Dr. Mehmet Sıddık AKTAŞ

Yıldız Teknik Üniversitesi _____

Doç. Dr. Banu DİRİ

Yıldız Teknik Üniversitesi _____

Doç. Dr. Ahmet SAYAR

Kocaeli Üniversitesi _____

ÖNSÖZ

Bu çalışma boyunca yaptığı önerilerden ve katkılardan dolayı Danışmanım Yrd. Doç. Dr. Mehmet Sıddık AKTAŞ'a çok teşekkür ederim. OKM sisteminin anket çalışmasına katılım sağlayan Bilgisayar Mühendisliği Bölümü araştırma görevlisi arkadaşlarıma ve bölüm memurlarına teşekkürü borç bilirim. Ayrıca, çalışma süresince tüm zorlukları benimle göğüsleyen ve hayatımın her evresinde bana destek olan sevgili aileme ve arkadaşlarıma sonsuz teşekkürlerimi sunarım.

Temmuz, 2016

Nurgül YÜZBAŞIOĞLU

İÇİNDEKİLER

Sayfa

SİMGE LİSTESİ....vii

KISALTMA LİSTESİ....viii

ŞEKİL LİSTESİ....ix

ÇİZELGE LİSTESİ x

ÖZET....xi

ABSTRACT....xiii

BÖLÜM 1

GİRİŞ 1

1.1 Literatür Özeti 3
1.2 Tezin Amacı 5
1.3 Hipotez 6
1.4 Araştırma Problemi 7
1.5 Organizasyon Yapısı 7

BÖLÜM 2

ANLAMSAL WEB VE ONTOLOJİLER 8

2.1 Web' Genel Bakış 8
2.2 Anlamsal Web 11
2.3 Anlamsal Web Uygulamaları 12
2.4 Anlamsal Web Standartları ve Dilleri 13
2.4.1 URI ve IRI 14
2.4.2 XML (Genişletilebilir İşaret Dili) 15
2.4.3 RDF (Kaynak Tanımlama Çerçevesi- Resource Description Framework) 16
2.4.4 RDF Şema (RDFS) 17
2.4.5 OWL (Web Ontology Language-Web Ontoloji Dili) 19

2.5 Ontolojiler 21
2.5.1 Ontoloji türleri 22
2.5.2 Ontoloji Bileşenleri 23
2.5.3 Ontoloji Editörleri 24
2.5.4 Ontoloji Yönetim Sistemleri 25
2.5.5 SPARQL (Simple Protocol and RDF Query Language) 25

BÖLÜM 3

TEMEL BİLGİLER 29

3.1 Temel Bilgiler 29
3.1.1 SKOS Ontolojisi 29
3.1.2 Bilgi Haritalama 30

BÖLÜM 4

OKM SİSTEMİ REFERANS ONTOLOJİLERİ 31

4.1 Sözlük Ontolojisinin Oluşturulması 31
4.1.1 Üst kavramların çıkarılması 32
4.1.2 TDK Türkçe Sözlüğü'ndeki Kelimelerin Elde Edilmesi 37
4.2 OKM Ontoloji Sınıfları ve Sınıf Hiyerarşisinin Belirlenmesi 37
4.3 OKM Referans Ontolojisinin Özelliklerinin (Properties) Belirlenmesi 39
4.4 Ontoloji Örneklerinin Oluşturulması 40
4.5 Veri Seti 41

BÖLÜM 5

UYGULAMA DETAYLARI 43

5.1 Sistem Mimarisi 43
5.2 Doğal Dil İşleme (NLP) Modülü 44
5.3 Anlamsal Web Dönüştürücü Modülü 45
5.4 Görselleştirme ve Kullanıcı Arayüzü Modülü 46

BÖLÜM 6

YÖNTEMİN DEĞERLENDİRİLMESİ 51

6.1 Kullanıcı Değerlendirme Çalışması 51
6.2 Değerlendirme 52

BÖLÜM 7

SONUÇLARI ETKİLEYEN FAKTÖRLER 55

BÖLÜM 8

SONUÇLAR VE GELECEKTE PLANLANAN ÇALIŞMALAR 57

KAYNAKLAR 58

EK-A 62

EK-B 69

EK-C 71

ÖZGEÇMİŞ 73

SİMGE LİSTESİ

H_o	Null hipotez
H_A	Alternatif hipotez
$\bar{X}$	Grup ortalamaları
SS	Standart Sapma
n	Gruptaki örneklem sayısı

KISALTMA LİSTESİ

DL	Description Logic
D3.js	Data Driven Document
IRI	Internatialized Resource Identifier (Uluslararası Kaynak Tanımlayıcı)
FOAF	Friend of a Friend
NLP	Natural Language Processing
OKM	Ontology Aided Knowledge Mapping (Ontoloji Tabanlı Bilgi Haritalama)
OWL	Web Ontology Language (Web Ontoloji Dili)
p	Anlamlılık değeri
RDF	Resource Description Framework (Kaynak Tanımlama Çerçevesi)
RDFS	Resource Description Framework Schema (Kaynak Tanımlama Çerçeve Şeması)
SKOS	Basit Bilgi Organizasyon Sistemi (Simple Knowledge Organization System)
SPSS	Statistical Package for the Social Sciences
TDK	Türk Dil Kurumu
URI	Uniform Resource Identifier (Uniform Kaynak Tanımlayıcı)
W3C	Wide Web Consortium
XML	Extensible Markup Language (Genişletilebilir İşaret Dili)

ŞEKİL LİSTESİ

Sayfa

Şekil 2.1 Web'in Gelişimi 11
Şekil 2.2 Anlamsal Web Katman Pastası 14
Şekil 2.3 XML doküman örneği 16
Şekil 2.4 Ontoloji türleri 23
Şekil 2.5 SPARQL sorgu yapısı 26
Şekil 2.6 ASK sorgu ifadesi örneği ve sorgu sonucu 27
Şekil 2.7 CONSTRUCT sorgu ifadesi ve sorgu sonucu 27
Şekil 2.8 DESCRIBE sorgu ifadesi örneği 27
Şekil 2.9 SELECT sorgu ifadesi ve sorgu sonucu 28
Şekil 4.1 OKM Referans Ontolojisi sınıfları ve hiyerarşisi 38
Şekil 4.2 RDF üçlüleri 41
Şekil 5.1 Önerilen sistemin mimarisi 44
Şekil 5.2 Zemberek kütüphanesi kullanılarak RDF üçlüsü oluşturulması 45
Şekil 5.3 Sistem kullanıcı arayüzü 46
Şekil 5.4 "öğrenci" kavramının görselleştirilmesi 47
Şekil 5.5 "güvenlik birim" kavramının görselleştirilmesi 48
Şekil 5.6 "almak" kavramının görselleştirilmesi 48
Şekil 5.7 SPARQL sorgulama 49
Şekil 5.8 Kavram dizini 49
Şekil 5.9 Metinsel kavram arama 50
Şekil 5.10 Yardım menüsü 50

ÇİZELGE LİSTESİ

Sayfa

Çizelge 2.1 RDF sınıfları 18
Çizelge 2.2 RDF özellikleri..... 18
Çizelge 4.1 Üst Kavram belirlemede kullanılan 1.Grup 1. Kural' ın Tanımı ve Örneği . 33
Çizelge 4.2 Üst Kavram belirlemede kullanılan 2.Grup 1. Kural' ın Tanımı ve Örnekleri 34
Çizelge 4.3 Üst Kavram belirlemede kullanılan 2.Grup 2. Kural' ın Tanımı ve Örneği . 35
Çizelge 4.4 Üst Kavram belirlemede kullanılan 2.Grup 3. Kural' ın Tanımı ve Örneği . 35
Çizelge 4.5 Üst Kavram belirlemede kullanılan 2.Grup 4. Kural' ın Tanımı ve Örneği . 35
Çizelge 4.6 Üst Kavram belirlemede kullanılan 2.Grup 5. Kural' ın Tanımı..... 35
Çizelge 4.7 Üst Kavram belirlemede kullanılan 2.Grup 6. Kural' ın Tanımı ve Örneği . 36
Çizelge 4.8 Üst Kavram belirlemede kullanılan 2.Grup 7. Kural' ın Tanımı ve Örneği . 36
Çizelge 4.9 Üst Kavram belirlemede kullanılan 2.Grup 8. Kural' ın Tanımı ve Örneği . 36
Çizelge 4.10 Üst Kavram belirlemede kullanılan 2.Grup 9. Kural' ın Tanımı ve Örneği . 37
Çizelge 4.11 OKM Ontolojisinde bulunan üst seviye sınıflar 38
Çizelge 4.12 Anlamsal Bilgi Çıkarım Matrisi 39
Çizelge 4.13 OKM Ontolojisinde bulunan özellikler (Properties) 40
Çizelge 4.14 Veri seti..... 42
Çizelge 5.1 Örnek cümlenin Zemberek çözümlemeleri 45
Çizelge 6.1 Anket için kullanılan Likert ölçeği 51
Çizelge 6.2 Anket puanları ortalaması..... 52
Çizelge 6.3 Gruplara göre eğitim veri seti için istatistiksel analiz sonuçları..... 54
Çizelge 6.4 Gruplara göre spor veri seti için istatistiksel analiz sonuçları..... 54
Çizelge 6.5 Gruplara göre sağlık veri seti için istatistiksel analiz sonuçları..... 54

ÖZET

ONTOLOJİ TABANLI BİLGİ HARİTALAMA YÖNTEMİ

Nurgül YÜZBAŞIOĞLU

Bilgisayar Mühendisliği Anabilim Dalı

Yüksek Lisans Tezi

Tez Danışmanı: Yrd. Doç. Dr. Mehmet Sıddık AKTAŞ

Bilgi Haritalama, basit ve anlaşılır şekiller kullanarak bilginin nasıl gösterilebileceğiyle ilgilenir. Bilgi gösteriminde kavramlar için bloklar ve kavramlar arasındaki ilişkiler için çizgiler kullanır. İnsanlar bilgiyi nasıl oluşturabileceği, bilginin hangi ilişkilere sahip olabileceği ve bilginin kaynağının nerede bulunduğu gibi bilgileri toplu halde bilgi haritası kullanarak öğrenebilirler. Bilgi haritaları hem bilgi paylaşımın da hem de örtük bilgiyi meydana çıkarma da esneklik sağlar. Mevcut çalışmalar genellikle insan müdahaleleri gerektiren Bilgi Haritalama metodolojileri üzerine odaklanmıştır. Biz bu çalışmada, Bilgi Haritası oluşturmak için birden fazla bilgi kaynaklarından, kavramları ve ilişkilerini çıkaran ontoloji destekli bir Bilgi Haritalama Sistemi öneriyoruz.

Önerilen sistemin gereksinimlerini ve yazılım mimarisini tanıttık. Önerilen yaklaşımda 5N1K sorularını kullanmaktayız. 5N1K'nın cevapları, metin dokümanlarından bilgi çıkarmak ve Web Ontoloji Dili(OWL) aracılığıyla önemli kavramları ve ilişkilerini sunmak için bilgi toplamada temel olarak kabul edilmektedir. Aynı zamanda, önerilen sistemimizde Anlamsal Web teknolojilerinin muhakeme yeteneklerini kullanarak örtük olarak tanımlanan bilgiyi ortaya çıkarıyoruz. Önerilen bilgi haritalama sisteminin kullanılabilirliğini göstermek amacıyla, bir dizi Türkçe belgelerden Ontoloji Tabanlı Bilgi Haritaları oluşturan bir web uygulaması sunduk. Ontoloji alanına gelince, eğitim, sağlık ve spor olmak üzere 3 farklı alana ait dokümanları seçtik. Kullanıcıların sorgularını girmelerine izin verilerek sorguları sonucunda Bilgi Haritası analiz edilmiştir. Son kullanıcılarla yaptığımız değerlendirme çalışması sonucunda umut edici sonuçlar görülmüştür. Önerilen ontoloji tabanlı yaklaşım geleneksel bilgi çıkarımla

karşılaştırıldığında, elde ettiğimiz sonuçlar ontoloji tabanlı yaklaşımın daha olumlu performans sergilediğini göstermektedir.

Anahtar Kelimeler: Anlamsal Web, bilgi haritalama, ontoloji, anlamsal bilgi çıkarımı, ontoloji tabanlı bilgi haritalama

ABSTRACT

ONTOLOGY BASED KNOWLEDGE MAPPING METHOD

Nurgül Yüzbaşıoğlu

Department of Computer Engineering

MSc. Thesis

Adviser: Asst. Prof. Dr. Mehmet Sıddık AKTAŞ

Knowledge Mapping deals with how knowledge can be displayed by using simple and understandable shapes. In knowledge displaying, blocks are used for concepts and lines are used for relationships between these concepts. Using knowledge maps people can learn how to create knowledge, which relations knowledge has, where knowledge source is. Knowledge maps give flexibility not only for sharing knowledge but also revealing implicit knowledge. Existing studies has mostly focused on Knowledge Mapping methodologies that requires human interventions. In this study, we propose an ontology aided knowledge mapping system that extracts concepts and their relationships from multiple information sources to create the Knowledge Maps.

We introduce the requirements and the software architecture of the proposed system. The proposed approach uses 5W1H questions, whose answers are considered as basic in information-gathering, to extract information from text documents and represent the important concepts and their relationships in Web Ontology Language. By utilizing the reasoning capabilities of Semantic Web technologies, the proposed system is able to reveal implicitly defined knowledge as well. In order to show the usability of the proposed knowledge mapping system, we introduce a Web application implementation that creates Knowledge Maps from a set of Turkish documents. As for the ontology domain, we chose documents belonging to three different areas, including education, health and sport. In this study, the most important point is to

create an ontology as domain-independent. We present promising results from an experiment in which users were allowed to input their queries and analyze the Knowledge Maps as result of their queries. After user study, we see promising results. The proposed approach performed favorably compared to the traditional way of extracting knowledge by analyzing the textual documents.The proposed ontology-based approach performed favorably compared to the traditional way of extracting knowledge by analyzing the textual documents according to results.

Keywords: Semantic Web, knowledge mapping, ontology, semantic information extraction, ontology aided knowledge mapping

BÖLÜM 1

GİRİŞ

Günümüzde, bilgi çağının gerekliliği, az zamanda çok ve doğru bilgiye ulaşmak olarak karşımıza çıkmaktadır. Web üzerindeki dokümanlar ele alındığında, dokümanlardaki bilgiler dağınık olarak bulunmakta, birbirleriyle ilişkileri tanımlanamamaktadır. Bu eksikliği gidermek amacıyla Anlamsal Web ortaya çıkmıştır. Tim Berners-Lee ve arkadaşları Anlamsal Web'i şöyle tanımlamaktadır: Anlamsal Web, internet ortamında ortak bir veri dili oluşturularak, verilerin insanlar kadar makinalar tarafından da anlaşılabilir, yorumlanabilir, çıkarsama yapılabilir ve kullanılabilir hale getirilmesidir. Böylece makinalar bilgiyi kolayca bulabilir, paylaşabilir ve bilgiler arasında bütünleşim sağlayabilir [1].

Ontoloji, felsefe biliminde, temel kategorilerinin araştırılması olarak tanımlanmaktadır. Bilgisayar biliminde ontoloji, belirli bir tanım kümesi içindeki kavramların ve bu kavramlar arasındaki ilişkilerin kurallı bir şekilde temsili olarak tanımlanabilir. Günümüzde, Webdeki içeriklerin ontolojilerle tanımlanabilmesi, Anlamsal Web tarafından sunulan teknolojilerin kullanılmasıyla gerçeklenmektedir. Anlamsal Web'in doğuşu ile birlikte, ontolojiler kullanılarak, bilgi varlıkları ve bilgi varlıkları arasındaki ilişkilerin tanımlanması, sağlıktan iş hayatına kadar birçok alana uyarlanmaktadır.

Bir dokümanın içerdiği bilgi daha basit ve anlaşılır hale getirildiğinde, bilgi makinalar tarafından da anlamlandırabilir olabilmektedir. Böylece kullanıcıların ihtiyaç duyduğu bilgiye daha kısa sürede ulaşması sağlanabilmektedir. Dokümanlar içinde yer alan sözcükler (bilgi varlıkları), dokümanın yazarı tarafından belirli bir mantık zincirine dayalı olarak, yüklemler (ilişkiler) kullanılarak aynı cümle içinde bir araya getirilmektedir. Belirli bir konu üzerindeki yazılmış dokümanlarda, dokümanların yazarlarının, genellikle

bu konuda yaygın olarak kullanılan sözcükler ve yüklemleri tercih ettiklerini görmekteyiz. Aynı konu üzerindeki farklı dokümanlarda yer alan cümlelerdeki, bilgi varlıklarının ve ilişkilerinin, ortaya çıkartılması ve Anlamsal Web tarafından sunulan, ontoloji tanımlama dilleri gösteriminin sağlanması, belirtilen konunun daha iyi anlaşılmasına olanak sağlayacaktır. Ancak günümüzde, aynı konuda yazılmış bir doküman kümesi içindeki bilgi varlıklarının ve ilişkilerinin, ontolojik tanımlama dilleriyle gösterimini sağlayan bir metodolojinin olmadığı görülmektedir.

Bilgi haritalama, tanımlanmış kavramların ve bu kavramlar arasındaki ilişkilerin görsellik yardımıyla kullanıcıya sunulması olarak düşünülebilir. Bilgi haritaları, veri yığınlarından, bilgi üretim sürecini basitleştiren, bilgi varlıklarının daha kolay bir şekilde anlaşılmasına olanak veren araçlar olarak tanımlanabilir. Bilgi haritalamada, ontoloji kullanımı ise, ontoloji dilleriyle tanımlanmış bilgilerin, son kullanıcılara daha kolay sunumunu sağlamak amaçlıdır. Bunun yanı sıra, ontoloji tanımlama dillerinin desteklediği çıkarsama yapabilme yeteneği, Bilgi Haritalarına, örtük bilgiyi de anlatabilme yeteneği sağlayabilmektedir. Günümüzde, dokümanlarda metinsel olarak yer alan kavramlar ve aralarındaki doğrudan ve/ veya dolaylı ilişkileri, ontoloji tanımlama dilleri kullanarak, Bilgi Haritaları üzerinde gösterimini sağlayabilecek metodolojilerin eksikliği görülmektedir.

Son yıllarda birçok Web uygulaması, verinin açık olarak paylaşımı için standartlaştırılmış ontolojileri kullanarak, alt yapısını Anlamsal Web'e uygun hale getirmektedir. En bilinen standart ontolojiler FOAF [2] ve SKOS [3] ontolojileridir. Standartlaştırılmış ontolojileri kullanmak, yeni bilgiyle Web' de bulunan bilginin kolaylıkla bağlanabilmesini, bilginin zenginleşmesini ve daha fazla anlamlı çıkarsamalar yapabilmesini sağlamaktadır. Standartlaştırılmış ontolojileri kullanmanın bir diğer faydası ise insanlara uygulamayı daha rahat bir şekilde kullanma ve anlama yetisi vermesidir. Günümüzde, bir çok alanda (eğitim, siyaset, spor) standart ontolojiler kullanarak, Türkçe dilindeki sözcükler ve sözcükler arasındaki dar-kavram, geniş-kavram ilişkileri; eş-anlamlılık, zıt-anlamlılık ilişkileri gibi; ilişkileri gösteren, bir sözlük ontolojisinin eksikliği görülmektedir.

Günümüzde geliştirilmiş olan, mevcut ontolojiler göz önüne alındığında, genellikle belirli bir alan için geliştirilmiş olan, ontolojiler olduğu görülmektedir. Belirli bir alan için ontoloji tanımlanırken, ilgili alana yönelik kavramların ve kavramlar arasındaki ilişkilerin tanımlanması, daha iyi çıkarsama sonuçları vermekte ve veri kontrolünü kolaylaştırmaktadır. Ancak bu ontolojilerin çıkartılması için kullanılan yöntemler incelendiğinde, bu yöntemlerin sadece tanımlandıkları alan için çözüm sunabildikleri görülmektedir. Günümüzde, belirli bir alan/konuda yer alan doküman kümesinden, ontolojilerin çıkartılmasına olanak sağlayan "alan/konu bağımsız"metodolojilere ihtiyaç olduğu görülmektedir.

Bu çalışma kapsamında yukarıda belirtilen eksiklere çözümler sağlayacak bir sistem önerilmekte ve önerilen sistemin bir prototip uygulaması geliştirilerek, yöntemin başarısı kullanıcı anketlerine dayalı bir ölçümleme yöntemiyle ortaya konmaktadır. Bu çalışmanın sunduğu bilimsel ve teknolojik katkılarını aşağıdaki şekilde özetleyebiliriz. 1) Aynı konuda yazılmış bir doküman kümesi içindeki bilgi varlıklarının ve ilişkilerinin, ontolojik tanımlama dilleriyle gösterimini sağlayan, "alan/konu bağımsız" bir metodoloji geliştirilmektedir. 2) Eğitim, sağlık ve spor alanlarına ait doküman kümeleri için, SKOS gibi standart bir ontoloji kullanılarak, biz sözcük ontolojisi geliştirilmektedir. 3) Bilgi Haritalama yöntemi kullanılarak, belirli bir doküman kümesi içinde yer alan bilgi varlıkları ve aralarındaki ilişkilerin, görsel yolla daha iyi anlaşılabilmesini sağlayacak bir yazılım için yazılım mimarisi önerilmekte ve bu mimariyi gerçekleyen bir prototip uygulama gerçeklenmektedir. 4) Geliştirilen, Ontoloji çıkarma ve Bilgi Haritalama yöntemlerinin başarısını ölçmek için, son kullanıcılarla yapılan bir kullanıcı anketi çalışmasına dayalı bir yöntem önerilmektedir. Bu yönteme dayalı olarak, geliştirilen metodolojinin başarısı ortaya konmaktadır.

1.1 Literatür Özeti

Musa ve diğerleri, yaptıkları çalışmada, iş süreci optimizasyonu alanında ontoloji odaklı bilgi haritası kullanıldığında organizasyon yapısı anlayışının eksikliğinden kaynaklanan hataların azaldığını göstermişlerdir [4]. Bizim çalışmamız, belirli bir kaç alanda yazılmış bir doküman kümesinden, bir ontoloji çıkartılması için alan/konu bağımsız bir metodoloji sunmaktadır. Ortaya çıkan ontoloji dili ile tanımlanmış bilginin, daha kolay

anlaşılabilmesi için bilgi haritalama yöntemi kullanılmıştır. Bu açıdan, çalışmamız, iş süreçlerinin bilgi haritaları yoluyla daha kolay anlaşılmasına yönelik olan Musa ve diğerlerinin çalışmasından ayrılmaktadır.

Kawtrakul ve diğerleri, yaptıkları çalışmada, çiftçilerin problemlerine çözüm üretebilmek ve ilgili uzmanları bulabilmek için Web sitelerinden bilgi toplayarak pirinç kavramlaştırması üzerine ontoloji tabanlı bilgi haritası oluşturmuşlardır [5]. Biz bu çalışmamızla, ontolojinin aynı konu üzerindeki doküman kümesinden otomatik olarak çıkartılması üzerine alan-bağımsız bir metodoloji önermekteyiz. Bu metodolojinin çalıştığını göstermek için eğitim, sağlık ve spor alanlarına ait doküman kümesini kullanarak, ontoloji tabanlı Bilgi Haritası oluşturmaktayız. Bu açılardan yaptığımız çalışma, Kawtrakul ve diğerlerinin çalışmasından ayrılmaktadır.

Ontoloji tabanlı bilgi yönetimi sistemlerine bakıldığında, Kim ve diğerlerinin, bilginin yöntemi için finansal firma çevrelerine yönelik ontolojiler tasarladığını [6] görmekteyiz. Ontoloji geliştirme yöntemlerine bakıldığında, Ushold ve Gruninger'in aşamalardan oluşan bir metodolojik yaklaşımı [7] benimsediğini görmekteyiz. Bizim bu çalışmada önerdiğimiz yöntem, doküman kümelerinden otomatik olarak ontoloji çıkartılması ve bunun Bilgi Haritasında gösterimine dayandığı için, alan-bağımsız bir yöntem olması ve iteratif/aşamalı bir ontoloji geliştirme yöntemini izlememesi nedenleriyle, daha önceki bu çalışmalardan ayrılmaktadır.

Barkschat yaptığı çalışmada [8] enerji sektörü ve yaşam bilimleri alanı ontolojisi alan bağımlı veri setlerinden anlamsal bilginin çıkartılması üzerinde yoğunlaşmıştır. Rusu ve diğerleri [9] tarafından gerçekleştirilen çalışmada "üçlü çıkarma algoritması" kullanılarak, veri setlerinden, üçlü yapılar (özne-yüklem-nesne) çıkartılması üzerinde çalışılmıştır. Bizim çalışmamız, alan bağımsız bir yöntem önermesi yönü ile Barkschat' ın çalışmasından ayırılmaktadır. Çalışmamızda, dokümanlarda yer alan cümleler üzerinde NLP kullanılarak, cümlelerin özne, yüklem ve nesne ayrıştırılması yapılmakta, yapılan ayrıştırma sonucunda 5W1H veri modelleme yöntemi kullanılarak, ontoloji oluşturulması üzerinde yoğunlaşılmaktadır. Çalışmamız, bu açılardan bakıldığında, Rusu ve diğerlerinin yaptıkları çalışmadan da farklılıklar göstermektedir.

Medeni, kurumların iş süreçleri ile ilişkili belgelerinden bilgi haritalama oluşturulması üzerine odaklanmıştır. Medeni de bilgi haritalama stratejisini bir kullanıcı çalışması ile soruşturmuştur [10]. Biz bu çalışmamızda alana bağımlı olmayan, uçtan uca bilgi çıkarımı metodolojisi üzerine odaklanmaktayız. Fakat, çalışmamızdaki değerlendirme stratejisi için Medeni'nin çalışmasından ilham alınmıştır. Bizde kullanıcı çalışmasına dayalı bir değerlendirme stratejisi kullanmaktayız.

Yukarıda incelenen çalışmalar kendi ontolojileri ile sınırlı kalmakta, standartlaştırılmış ontolojilerden faydalanmamaktadır. Bizim önerdiğimiz ontoloji geliştirme metodolojisi, temel ontolojilerden SKOS' u destekleyerek, kavramlar arasındaki ilişkileri destekleyecek SKOS sınıf ve özellikleri kullanmaktadır. SKOS'u kelimelerin ayrıntılarını tanımlamada kullanmaktayız. Bildiğimiz kadarıyla, Türkçe bir doküman kümesi üzerinden alandan-bağımsız bir metodoloji kullanarak, otomatik ontoloji oluşturulmasını sağlayan daha önce geliştirilmiş bir yöntem yada araç yoktur. Bu açıdan bu çalışma sonucu ortaya çıkacak olan bilimsel ve teknolojik katkının önemli olduğunu ve yaygın etkisinin olacağını düşünüyoruz.

1.2 Tezin Amacı

İnternetin, hayatımızın her noktasına nüfus etmesiyle beraber aklımıza gelebilecek birçok işimizi internet üzerinden halletmekte ve bütün gerekli bilgilere internet ortamından erişmekteyiz. Bazen de aradığımız bilgiye ulaşamamakta, anlam belirsizliği ile karşı karşıya kalmakta ve zaman kaybı yaşamaktayız. Makinalar da Web içeriklerini anlamlandırmaya başladığında bu olumsuzlukları aşabileceğimiz öngörülmektedir. Fakat günümüz Web'inde yer alan bütün içerikler insanların kullanımı için oluşturulmuş vaziyettedir. Bu sebeple makinalar, Web içeriklerini anlayamamakta ve işleyememektedir. Bu sıkıntıları aşmak için Anlamsal Web ve Ontolojiler ortaya çıkmıştır. Fakat Web'de bulunan içeriklerin devasa büyüklüğü sebebiyle alana bağımlı olarak ontolojiler oluşturulmaktadır ve bu manuel ve iteratif bir yöntemle ontoloji oluşturulması zaman almaktadır. Bu sebeplerle insanların ihtiyaç duyduğu ve makinalarında anlamlandırabildiği, çıkarım yapabildiği bilgiyi, bilgi haritalamanın sağladığı imkanlarla birleştirerek sunmak önem arz etmektedir.

Yukarıda bahsedilen nedenlerden, doküman içeriğinden anlamı ayrıştırmak için ontoloji tabanlı bilgi haritalama sisteminin oluşturulmasına ihtiyaç duyulmaktadır.

1.3 Hipotez

Bu çalışmada Ontoloji tabanlı Bilgi Haritalama sistemi geliştirilmesi hedeflenmektedir. Ontoloji, Anlamsal Web teknolojilerini kullanarak, geliştiriciler tarafından bilginin iyi tanımlanmış olmasını garanti etmekte, kavram ve kavramlar arasındaki ilişkileri insanlar kadar makinalarında anlayabileceği duruma getirmektedir. Ontolojiler genellikle alana bağımlı olarak yaratılmakta ve belli bir problemi çözme amacı gütmektedir. Bu çalışma, birbirinden bağımsız birkaç alanda yazılmış bir doküman kümesinden, otomatik olarak bir ontoloji çıkartılması için alan/konu bağımsız bir metodoloji sunmaktadır. Bu metodolojide, doküman kümesinde yer alan cümleler üzerinde NLP kullanılarak, cümleler özne, yüklem ve nesne öğelerine ayrıştırılmakta, ayrıştırma neticesinde 5W1H veri modelleme yöntemi kullanılarak alandan bağımsız bir ontoloji oluşturulması üzerine odaklanılmaktadır. Bu oluşturulacak ontolojideki kavramlar ve ilişkilerin daha geniş ve anlaşılır olması maksadıyla daha önceden oluşturulan, kavramların alt ve üst bilgilerini barındıran Sözlük Ontolojisi, oluşacak alandan bağımsız ontoloji ile birleştirilmektedir. Böylelikle makinaların kavramlar hakkında daha fazla bilgisi sahibi olması sağlanmakta ve bilgi çıkarım gücü arttırılmaktadır.

Bilgi Haritalama, grafiksel gösterim temelinde, bir kavramın başka kavramlarla olan ilişkisini kullanıcıya sunarak, bilginin daha uzun süre akılda kalıcılığını arttırmakta ve öğrenmeyi üst seviyeye çıkarmaktadır. Ayrıca Bilgi Haritalama, kavramlar arasındaki detayları ortaya çıkarmaktadır. Ontolojik olarak tanımlanan bilginin daha rahat ve kolay anlaşılması için Bilgi Haritalama'dan faydalanılmaktadır.

Yapmış olduğumuz literatür taramasında, Türkçe dilinde farklı alanlara ait bir doküman kümesi kullanılarak, alana bağımlılığı kaldırılmış bir metodolojiyle, otomatik ontoloji oluşturulmasını sağlayan daha önce geliştirilmiş bir metot yada araç bulunamamıştır. Bu bakımdan çalışmamız neticesinde ortaya çıkacak olan bilimsel ve teknolojik katkının önemli olduğunu düşünüyoruz.

1.4 Araştırma Problemi

Yapılan literatür taraması bilgi haritalamada kullanılan ontolojilerin sadece alana bağımlı olarak oluşturulduğunu, alandan bağımsız bir ontolojiden faydalanılmadığını göstermektedir. Ontoloji tabanlı bilgi haritalarının geliştirilmesinde başka alanlarda kullanımı pek mümkün olmayan, yalnız bir alana ait sınıflar ve özellikler belirlenerek, bu ontolojilere RDF üçlülerinin verilmesi yaklaşımı kullanılmıştır.

Bu tezde önerilen yöntem ile belirli bir alana ait olmayan doküman kümelerinden otomatik olarak ontoloji çıkartılması ve bunun Bilgi Haritasında gösterimi üzerinde durulmaktadır. Bu araştırmanın özgün katkısı alan-bağımsız bir yöntem olması ve aşamalı bir ontoloji geliştirme yöntemini izlememesidir. Bunlara ek olarak kavramlar arasındaki anlam ilişkisini arttırmak için standart ontolojilerden de faydalanılmasıdır.

1.5 Organizasyon Yapısı

Tezin organizasyon yapısı takip eden şekildedir. Bölüm 1'de Anlamsal Web teknolojilerinden ontolojiye ihtiyaç duyulma sebepleri, bilgi haritalamanın gerekliliği ve tezin katkıları özetlenmektedir. Bölüm 2'de Anlamsal Web ve teknolojileri, Ontoloji türleri, ontoloji tanımlama ve sorgulama dilleri hakkında bilgiler yer almaktadır. Bölüm 3'te standart ontolojilerden SKOS, bilgi haritalama ve literatürde bulunan çalışmalar incelenmektedir. Önerilen sistemin kullandığı referans ontolojiler ve bileşenleri Bölüm 4'te anlatılmaktadır. Bu çalışma kapsamında önerilen yöntemin gerçekleştirilmesi için tasarlanan mimari yapı ve kullanılan teknolojiler Bölüm 5'te verilmektedir. Bölüm 6' da geliştirilen sistemin değerlendirilmesi yapılmaktadır. Bölüm 7' de geçerlilik tehditlerinden bahsedilmektedir. Bölüm 8' de elde edilen sonuçlar ve gelecek çalışmalar anlatılmaktadır.

BÖLÜM 2

ANLAMSAL WEB VE ONTOLOJİLER

2.1 Web' Genel Bakış

Bilgi teknolojilerindeki hızlı gelişmeler ile dünya küçülmüş ve iletişim kolay hale gelmiştir. Web bütün dünyaya yayılmış bağlantılı bir bilgi havuzudur [11].Buna ön ayak olan Web'de kendi içinde doğru orantılı olarak gelişim göstermiştir. Web'in gelişimi Şekil 2.1' de gösterilmektedir.

Web 1.0 internetin başlangıcı olarak bilinir. Web 1.0 ile kullanıcılar yayınlanan içerikleri sadece okuyabilmekte, görebilmekte ve dinleyebilmekteydi. İçerikler hakkında ise yorum yapamamaktaydı. İnsanlar sadece bilgi edinmek için Web'i kullanılır ve ayrılırlardı. İnsan etkileşimi bulunmamaktaydı. Bu sebeple Web 1.0 "statik Web" olarak tanımlanmaktadır. Kısacası, Web 1.0 bilgiyi internet ortamında sunan, pasif bir platform görevindeydi. Ayrıca, insanların teknik bilgi yetersizliğinden dolayı tasarım ve kullanışlılık bakımından basit ve kötü Web sayfaları bulunmaktaydı.

Web 1.0'ın sunduğu pasif hizmetin insan ihtiyaçlarını karşılamaya yetmemesi neticesinde Web 2.0 ortaya çıkmıştır. Web ortamı, Web 2.0 ile birlikte temelinde kullanıcı etkileşimini destekleyen demokratik bir çevre halini almıştır. Web 2.0 kullanıcıya içerik üretme, düzenleme, denetleme, silme ve yorum yapma gibi etkileşim imkanını sunmuştur. Bu alandaki gelişmeler kişisel Web sitelerinin üzerinde etkisini arttırmış, birçok insanı bu ortama çekmiştir. Böylelikle insanlar istedikleri bilgiyi, içeriği kendi Web sayfalarında paylaşmaya ve diğer kişiler içinde etkileşimli bir kullanım sunmaya başlamıştır. Web 2.0 ile Web ortamı sadece bilgi edinme amacıyla

kullanılmaktan çıkmıştır. İnsanlar Web 1.0'da sadece pasif ve tüketici konumundayken, Web 2.0 ile aktif ve üretici konumuna ulaşmıştır. Web 1.0 ile kullanıcılar sadece masaüstü bilgisayarlarıyla Web'de gezinebiliyorlarken, Web 2.0 ile kullanıcılar laptop, mobil telefon, tablet gibi geniş platformlardan günlük hayatın her anında yer alan etkileşimli bir kullanım elde etmişlerdir. Böylelikle Web 2.0 "Mobil Web " olarak bilinmeye başlanmıştır. Ayrıca, Web 2.0 ile birlikte tasarım ve kullanışlılık bakımından daha etkili ve iyi Web sayfaları oluşturulmaya başlanmıştır. Facebook ve Twitter Web 2.0 için verilebilecek en iyi örneklerden sadece ikisidir.

Web 3.0 ile Anlamsal Web kavramı ortaya atılmıştır. İnternetin yaygın bir hal almasıyla beraber birçok işimizi internet üzerinden gerçekleştirmekte, ihtiyacımız olan bütün bilgileri internet aracılığıyla aramaktayız. Aradığımız bilgiye ulaşmakta bazen sıkıntı çekmekte ve zaman kaybı yaşamaktayız.

Web 3.0 teknolojisi Anlamsal Web uygulamalarının hayata geçirilme amacıyla geliştirilmiştir. Tim Berners-Lee Anlamsal Web'i şuanki Web'in bir uzantısı olarak tanımlamakta, başka bir Web olmadığını vurgulamaktadır [1].

Günümüzde var olan Web içeriklerinin çoğu insanların kullanımı için oluşturulmuştur. Bu sebeple makinalar ve bilgisayarlar bu içeriklere ulaşamamakta ve bu içerikleri anlayamamaktadır. Anlamsal Web'in amacı makinaların da insan müdahalesi olmadan, etkileşimli olarak içerikleri işleyebilir ve anlayabilir olmasını sağlamaktır[12].

Web'deki gelişmeler bizim bilgisayarlar hakkındaki düşüncelerimizi de değiştirmiştir. İlk olarak matematiksel hesaplamaları yapmak için kullanılan bilgisayarlar, şimdiler de ise bilgi işleme, veri tutma ve eğlence amacıyla kullanılmaktadır [13].

Web 3.0'ı, kullanıcı tarafından iyi tanımlanarak oluşturulan içeriklerin makinalar tarafından anlamlandırılması ve yorumlanması ile bilgi ihtiyacının en iyi şekilde karşılanması olarak düşünmek yerinde olacaktır. Bu sebeple Web 3.0 Anlamsal Web olarak adlandırılmaktadır. Web 3.0, anlamlı ve kullanışlı bilgiye erişimi esas alır. Temelinde, bilgiler kategorize edilmekte ve makinalar tarafından ayıklanarak kullanıcının kullanımına hazır hale getirilmektedir.

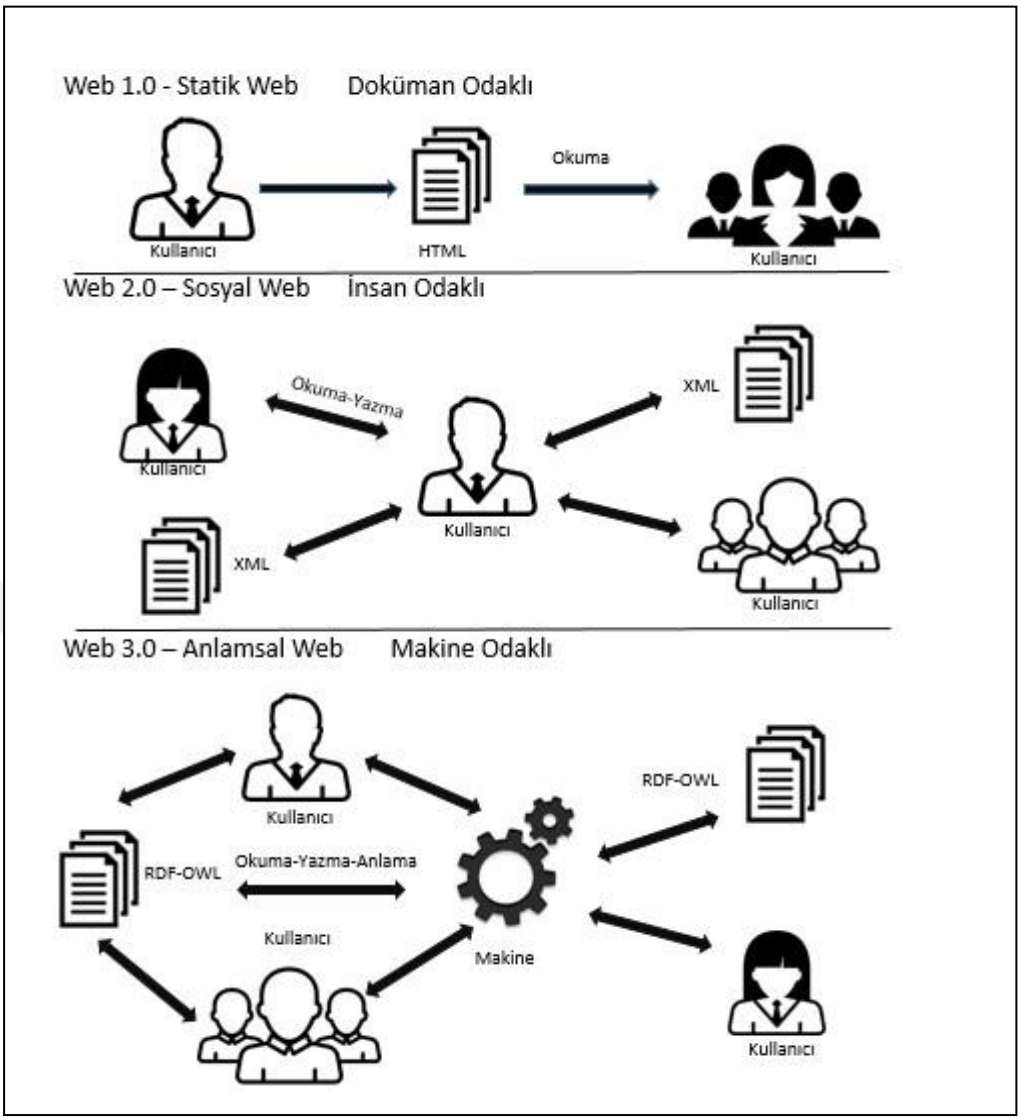
Anlamsal Web Tim Berners-Lee tarafından öne sürülmüştür. Anlamsal Web, Web'in sadece insanlar tarafından değil aynı zamanda makinalar tarafından da tam potansiyel

kullanımını hedeflemektedir. Anlamsal Web yeni bir Web olarak düşünülmemelidir. Anlamsal Web, bilgilerin iyi tanımlandığı, insanların ve makinaların beraber çalışmalarına imkan veren günümüz Web'in uzantısıdır [1].

Anlamsal Web kavramlar ve kavramlar arasındaki ilişkilerin iyi tanımlanarak, bilgilerin makinalar tarafından okunabilirliğini, anlaşılabilirliğini, yorumlanabilirliğini ve üretilebilirliğini sağlayan standartların ve teknolojilerin iyileştirilmesi ve geliştirilmesidir.

Web 3.0 ve Anlamsal Web uygulamaları günümüzde genellikle akademik ve endüstriyel alanlarda kullanılmaktadır. Anlamsal Web etkin olarak hayatımıza girmeye başladığında, bilgisayarlar Web içeriklerini yorumlama kabiliyeti kazanmış olacak ve böylelikle bize ihtiyaç duyduğumuz bilgiyi daha kısa zamanda ve daha doğru sonuçlarla sunacaktır. Web 3.0 ile birlikte bilgileri organize etmek ve kullanışlı hale getirmek mümkün görünmektedir. Anlamsal Web, doküman içeriğinden anlamı ayrıştırmamızı mümkün kılmaktadır.

Anlamsal Web'in hedefi, bütün bilgileri tek bir noktada toplamak ve makinalar tarafından bu bilgilerin yönetilmesini sağlamaktır. Bunun için veri metadatalarının oluşturulması, kavramların açıklamalarının yapılması, kategorize edilmesi ve ilişkilerin belirlenmesi gerekmektedir.



Şekil 2.1 Web'in Gelişimi

2.2 Anlamsal Web

Anlamsal Web'i anlamak için öncelikle "semantic" kelimesini açıklamamız gerekir. "semantic" kelimesi Yunanca "semantikos" kelimesinden gelmekte ve "gerçek anlam" ı ifade etmektedir. Gerçek anlam esas verinin daha etkili kullanımına olanak sağlamaktadır.

Günümüzde internet birçok Web sayfasından meydana gelmekte ve bilgisayarlar tarafından sunulan bu Web sayfaları insanlar tarafından kullanılmaktadır. Makinalar bu sundukları Web dokümanları ve içeriklerinin anlamı hakkında bilgi sahibi değillerdir, Web'teki bu içerikleri anlamlandıramamakta ve yorumlayamamaktadır. Yani bütün

Web sayfaları makinalara aynı görünür. Anlamsal Web teknolojilerinin ortaya çıkmasıyla beraber, içerikler yazılım ajanları tarafından da anlamlandırılabilir hale gelmektedir [14].

Anlamsal Web'i bir yapay zeka olarak tanımlamak yerine Anlamsal Web için yapay zeka teknolojilerini kullanır da diyebiliriz. Anlamsal Web makinaların da anlayabildiği akıllı veriler üretilmesine ön ayak olmaktadır. Buradaki akıllı veri tabiri iyi tanımlanmış veriler anlamına gelmektedir. Makinaların insan dilini bilmelerini istemekten ziyade, insanlar daha kolay ve anlaşılır akıllı veriler üretmeye çalışmalıdır [15]. Birçok Anlamsal Web uygulamalarının hala geliştirme aşamalarında olmalarına rağmen DBpedia, GeoNames gibi başarılı örnekleri de günümüzde kullanılmaktadır.

Bütün Web içerikleri Anlamsal Web teknolojileri yardımıyla dönüştürüldüğünde, o zaman uygulamalar birbirlerine bağlanabilecek ve anlam kazanacaktır. Anlamsal Web'in gücü hayatımıza kolaylıklar sunacaktır. Anlamsal Web'in gerçeklenmesi ile Web verileri makinalar tarafından doğrudan ve dolaylı olarak işlenebilir olacaktır [16].

Bu noktadaki en büyük zorluk Web'deki içeriklerin aşırı büyük ve dağınık olması, bu sebeple Anlamsal Web teknolojilerinin kullanabileceği RDF formatına dönüştürülememesidir.

2.3 Anlamsal Web Uygulamaları

Anlamsal Web teknolojileri çok çeşitli alanlarda uygulama bulmaktadır. Bu alanlardan bazılarına aşağıda değinilmiştir.

Anlamsal Web Arama Motorları: Ontolojik olarak tanımlanan Web içerikleri normal arama motorlarından daha akıllı sorgulamalar yapmaya olanak tanımaktadır. Anlamsal tabanlı arama motorlarına en iyi örnek olarak Swoogle, Kengine verilebilir.

Anlamsal Dijital Kütüphaneler: Anlamsal Web teknolojileri bilgileri düzgün bir şekilde sınıflandırma ve indeksleme imkanı sunmaktadır. Anlamsal Web teknolojilerinin sağladığı bu imkanla beraber anlamsal tabanlı dijital kütüphaneler veri erişimini kolaylaştırmaktadır. Anlamsal dijital kütüphaneye en güzel örnek JeromeDL'dir. Veriler üzerinden sorgulama ve işlem yapılabilmesi için JeromeDL'deki bütün veriler RDF formatında saklanmaktadır [17].

Ontoloji Tabanlı Kurumsal Bilgi Yönetimi: Küresel ekonomi ile birlikte üretimin verimli bir şekilde devam edebilmesi için emek, sermaye gibi geleneksel kaynakların organizasyonu ve yönetimi önem arz etmektedir. Anlamsal Web teknolojileri bu alandaki kurumsal bilgilerin etkin bir şekilde yönetimini, taranmasını ve kullanılmasını sağlamaktadır.

2.4 Anlamsal Web Standartları ve Dilleri

1994 yılından beri World Wide Web Consortium (W3C) tarafından Anlamsal Web standartları tanımlanmaya devam edilmektedir. Bu standartların bazıları XML, RDF, RDFS, OWL ve SPARQL'dir.

Şekil 2.2, Tim Berners-Lee tarafından hazırlanmış ve Anlamsal Web'in tasarımının ve vizyonunun ana katmanlarını göstermektedir. En alt katman olan XML, kullanıcıların kendi tanımladıkları etiketler yardımıyla yapısal Web belgeleri oluşturulmasına izin veren bir işaretleme dildir. XML, Web üzerinde belge yayınlamaya, aynı zamanda veri transferine uygundur. Fakat makinaların Web içeriklerini anlayabilmesi için bir etkisi veya yararı yoktur.

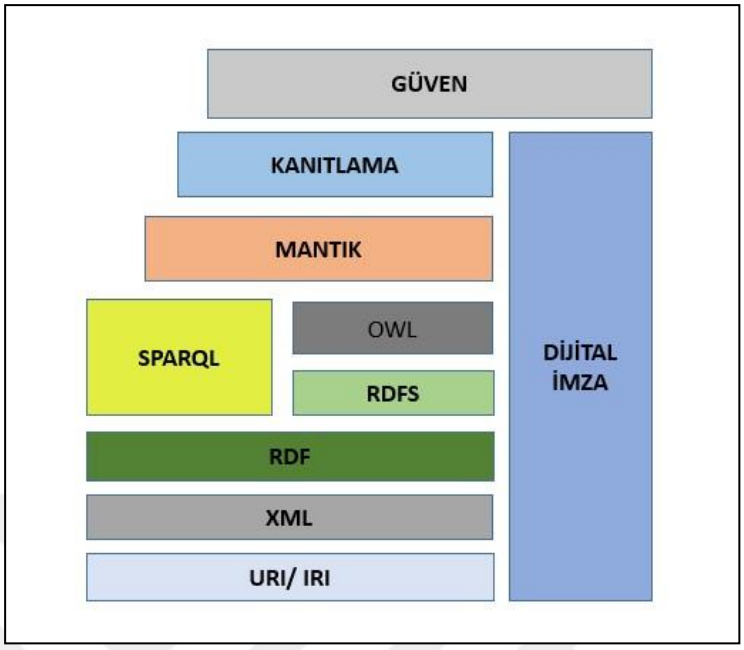
RDF, Web içeriğindeki varlıkları ve ilişkileri tanımlamaya ve bunları basit ifadeler ile göstermeye yarayan bir veri modelidir. RDF'in XML tabanlı sözdizimi mevcuttur. Bu sebeple RDF, XML katmanının üstünde bulunur.

RDF Şema (RDFS) Web kaynaklarını organize etmeye ve sınıflandırmaya yardımcı olan modelleme ilkelerini kapsar. Sınıflar ve özellikler anahtar ilişkileri, değer kümesi (domain) ve tanım kümesi (range) kısıtlamalarıdır. RDF Şema RDF modeline bağlıdır. RDF Şema ontoloji yazmaya yardımcı bir dil olarak düşünülebilir.

Web kaynakları arasında daha fazla karmaşık ilişkiler tanımlayabilmek ve RDF Şemayı genişletebilmek için daha kapsamlı bir ontoloji diline gereksinim vardır. Bu gereksinime cevap olarak Web Ontoloji Dili geliştirilmiştir. Kullanıcılara RDF Şema'dan farklı olarak daha fazla kısıtlama tanımlayabilme imkanı sunar.

Mantık Katmanı (Logic Layer) ontoloji dilini geliştirmek için kullanılır. Ayrıca uygulamaya özgü açıklayıcı bilgi yazılmasını da destekler. Kanıt Katmanı (Proof Layer) tümdengelim sürecini ve doğrulamayı içinde barındırır. Katman pastasının en üstünde

yer alan Güven Katmanı (Trust Layer) önemli bir katmandır. Sağlanan bilginin kalitesine yönelik kullanıcılara güven sunduğunda bu katman gerçeklenmiş olacaktır.



Şekil 2.2 Anlamsal Web Katman Pastası

2.4.1 URI ve IRI

Uniform Kaynak Tanımlayıcı (URI-Uniform Resource Identifier) W3C tarafından ortaya atılmış karakter kümeleridir. URI, Web kaynaklarını tanımlamak amacıyla kullanılır [18]. Her Web kaynağının kendine ait karakter veya dizelerden oluşan bir URI'ı bulunmaktadır.

Uluslararası Kaynak Tanımlayıcı (IRI-Internationalized Resource Identifier), URI'dan farklı bir amacı olmamakla beraber, URI'ın daha da genişletilmiş bir standardı olarak düşünülebilir. URI oluşturulmasında sadece ASCII karakterleri kullanılabiliyorken, bir IRI oluşturulmasında ise Çince, Korece, Kiril karakterleri gibi Evrensel Karakter Kümesi de(UNICODE) kullanabilir [19]. Bir URI veya IRI sadece bir Web kaynağını işaret edebilir. Böylelikle hangi kaynak hakkında konuşulduğu bilinir, belirsizlik karmaşası ortandan kaldırılmış olur.

URI Web kaynaklarının tanımlanması için çok önemlidir. Bu sebeple W3C tarafından Anlamsal Web için kullanılacak URI'lar için 2 standart belirlenmiştir [20]. Bunlar;

tanımlanan URI'lara hem insanların hem de makinaların ulaşabilir olması ve URI'ların tutarlı olması.

2.4.2 XML (Genişletilebilir İşaret Dili)

W3C tarafından geliştirilen XML Genişletilebilir İşaret Dili (Extensible Markup Language) anlamına gelmektedir. XML ile birlikte yarı yapılandırılmış veri terimi literatüre girmiştir [21]. XML geliştirilmeden önce verilerin veritabanlarına taşınmasında, başka yazılım veya platformlarda kullanılmasında birçok problemlerle karşılaşılmaktaydı. Verileri uyumlu bir formatta transfer etme maksadıyla XML ortaya çıkmıştır.

XML dokümanları bilgiyi sunmak için ek bilgilerle birlikte ağaç yapısını kullanmaktadır. Ağaç yapısı, etiketlerden, elemanlardan ve niteliklerden oluşur.

XML'de HTML gibi bir işaretleme dilidir. Aralarındaki fark, HTML sadece verinin görüntülenmesine odaklanırken, XML verinin farklı platformlarla paylaşılmasına ve transfer edilmesine odaklanmaktadır.

Şekil 2.3' te görüldüğü gibi "bookstore " kök düğüm etiketinin 3 adet "book" etiketli çocuk düğümü vardır. "book" düğümünün "title, author, year, price" olmak üzere 4 niteliği vardır [22].

XML dokümanlarının, yapılandırılmamış diğer türdeki dokümanlara göre bazı avantajları vardır [23].

- XML dokümanı olarak oluşturulan Web içeriklerini hem insanlar hem de makinalar kolaylıkla okuyabilir.
- Kullanıcılar kendi etiketlerini kendileri tanımlayabilirler. XML veri tipi esnek ve değiştirilebilirdir.
- XML işaretleme dili başka diller ve programlar tarafından desteklenmektedir.
- XML veri yapısı basit ve standart olması sebebiyle kolaylıkla işlenebilir.
- XML dokümanları yarı-yapılandırılmış olmasından ötürü veritabanı gibi kullanılabilir.

- XML, uygulamalar arasında veri transfer işlemini kolaylaştırmak için bir değişim formatı olarak kullanılabilir.

XML'in bu kadar avantajının yanı sıra verinin tanımlanmasında bir limitinin olmaması problemi bulunmaktadır. Bu problemin RDF kullanılarak çözüleceği öngörülmektedir.

```
<bookstore>
  <book category="cooking">
    <title lang="en">Everyday Italian</title>
    <author>Giada De Laurentiis</author>
    <year>2005</year>
    <price>30.00</price>
  </book>
  <book category="children">
    <title lang="en">Harry Potter</title>
    <author>J K. Rowling</author>
    <year>2005</year>
    <price>29.99</price>
  </book>
  <book category="web">
    <title lang="en">Learning XML</title>
    <author>Erik T. Ray</author>
    <year>2003</year>
    <price>39.95</price>
  </book>
</bookstore>
```

Şekil 2.3 XML doküman örneği

2.4.3 RDF (Kaynak Tanımlama Çerçevesi- Resource Description Framework)

Kaynak Tanımlama Çerçevesi (RDF) Anlamsal Web'le beraber ortaya çıkmıştır. RDF, W3C tarafından oluşturulan, Web üzerindeki verilerin belli bir standartlara göre tanımlanması ve diğer kaynaklarla arasındaki ilişkilerin belirlenmesi için bir veri modelidir.

RDF üç bileşenden oluşmaktadır [24]: : Özne (subject), yüklem (predicate) ve nesne (object). Bu üç bileşen literatürde ifade (statement) veya üçlü (triple) olarak adlandırılmaktadır.

RDF'in oluşturulmasındaki temel amaç insanlardan ziyade makinalara da veri sağlamaktır. Diğer bir amacı da verileri üçlüler (triples) olarak veritabanlarda kolayca saklayabilmektir. Eldeki bilgiyi parça parça açıklamak için RDF üçlüleri kullanılmaktadır [25]. RDF dosyaları birçok üçlülerden oluşmaktadır. Bu üçlüler günlük diyaloglardaki basit cümlelerle benzerlik göstermektedir.

RDF'in temel kavramları kaynaklar (resources), özellikler (properties) ve kaynaklarla özelliklerin ilişkilendirilmesiyle oluşan ifadelerdir (statements).

Kaynaklar: Kaynaklar; kitaplar, aletler, insanlar, yerler, somut ve soyut aklımıza gelebilecek, üzerinde konuşabileceğimiz her şey olabilir. Her bir kaynağın benzersiz bir URI'ı vardır. Böylelikle hangi kaynaktan bahsedildiği ve tutarlılığı sağlanmış olur.

Özellikler: Özellikler, kaynakların bir türü olup kaynaklar arasındaki ilişkileri belirtmek için kullanılır. Özelliklere örnek olarak "adres", "çalışır", "sahiptir" verilebilir. Özellikler de kaynaklarda olduğu gibi benzersiz URI ile temsil edilir.

İfadeler: Kaynakların özelliklerinin belirtilmesiyle ifadeler oluşur. İfade nesne-nitelik-değer (object-attribute-value) üçlülerinin bir arada olması halidir. Burada nesne kaynağı, nitelik özelliği gösterir. Değer ise kaynak veya literal olabilir. Literaller ise atomik dizelerdir.

RDF veri formatının avantajlarına bakacak olursak:

- Evrensel tanımlayıcılar kullanarak belirsizlikleri azaltır.
- RDF ifadeleri kolayca kullanılabilir ve saklanabilir. Önceden veri boyutunun belirtilmesi gibi bir ön koşula ihtiyaç yoktur.
- RDF modeli kurallı bir şekilde yönlendirilmiş grafiklerden (graph) oluşur. Böylece grafik kullanımının sağladığı tüm avantajları kullanıcıya sunar.
- RDF veri entegrasyonu ve veri transferi kolaylığı sunar.

2.4.4 RDF Şema (RDFS)

RDFS, RDF veri modelini genişleten bir tür sistemdir. RDFS, RDF'i genişletmek için bir dizi kaynaklar tanımlar [26]. RDF Şema, RDF kaynaklarının sınıflarını, özelliklerini ifade

etmek için kullanılan sözcükler bütünüdür. Ontoloji içinde sınıf hiyerarşisini oluşturur. Çizelge 2.1' de RDF sınıfları Çizelge 2.2' de RDF özellikleri verilmektedir.

Çizelge 2.1 RDF sınıfları

Sınıf İsmi	**Açıklamalar**
rdfs:Resource	Sınıf kaynağı, herşey.
rdfs:Literal	Atomik değerleri temsil eder, örneğin metinsel dizeler.
rdfs:XMLLiteral	XML literal sınıf.
rdfs:Class	Class kavramı.
rdf:Property	Property kavramı.
rdfs:Datatype	Veri tipi sınıfı.
rdf:Statement	RDF ifadeler sınıfı.
rdf:Bag	Sıralanmamış koleksiyonlar.
rdf:Seq	Sıralanmış koleksiyonlar.
rdf:Alt.	Alternatifler sınıfı.
rdfs:Container	Konteyner kümesinin sınıfı.
rdfs:ContainerMembershipProperty	Koyteyner üyelik özelliği, rdf:1, rdf:2, ...,
rdf:List	RDF List sınıfı.

Çizelge 2.2 RDF özellikleri

Özellik İsmi	**Açıklamalar**	**Tanım kümesi**	**Değer kümesi**
rdf:type	Öznenin bir sınıfın örneği olma durumu.	rdfs:Resource	rdfs:Class
rdfs:subClassOf	Öznenin bir sınıfın alt sınıfı olma durumu.	rdfs:Class	rdfs:Class
rdfs:subPropertyOf	Öznenin bir özelliğin alt özelliği olma durumu.	rdf:Property	rdf:Property
rdfs:domain	Özne özelliğinin tanım kümesi.	rdf:Property	rdfs:Class
rdfs:range	Özne özelliğinin değer kümesi.	rdf:Property	rdfs:Class

Çizelge 2.3 RDF özellikleri (devamı)

Özellik İsmi	Açıklamalar	Tanım kümesi	Değer kümesi
rdfs:label	Özne için insan okuyabilir isim.	rdfs:Resource	rdfs:Literal
rdfs:comment	Özne kayanğının açıklaması.	rdfs:Resource	rdfs:Literal
rdfs:member	Özne kaynağının üyesi.	rdfs:Resource	rdfs:Resource
rdf:first	Özne RDF listesinin ilk elemanı.	rdf:List	rdfs:Resource
rdf:rest	Özne RDF listesindeki ilk elemandan sonrakiler.	rdf:List	rdf:List
rdfs:seeAlso	Özne kaynak hakkında daha fazla bilgi.	rdfs:Resource	rdfs:Resource
rdfs:isDefinedBy	Özne kaynağının tanımı.	rdfs:Resource	rdfs:Resource
rdf:value	Yapılı değerler için deyimsel özellikler	rdfs:Resource	rdfs:Resource
rdf:subject	RDF ifadesinin öznesi.	rdf:Statement	rdfs:Resource
rdf:predicate	RDF ifadesinin fiili.	rdf:Statement	rdfs:Resource

2.4.5 OWL (Web Ontology Language-Web Ontoloji Dili)

Web Ontoloji Dili (OWL), kavramları daha iyi tanımlamak için en güncel teknik bileşenleri içinde barındıran ontoloji tanımlama dilidir. OWL dilinin XML, RDF ve RDFS üzerine inşa edilmesi sebebiyle kullanıcılar bu dili anlamakta güçlük çekmemektedir.

OWL gibi Anlamsal Web ontoloji tanımlama dilleri, tanımlama mantığından ("Description Logic") gelen kökleri sayesinde "subclassOf", "transitive" ve "symmetric" gibi birçok özellik tanımıyla zengin çıkarsamalar yapılmasına olanak tanımaktadırlar. Bu tanımlama dillerinde RuleML gibi farklı kural tanımlama dilleri sayesinde geliştiriciler tarafından da yeni kurallar tanımlanması mümkün kılınmıştır. Alan ontolojilerinin tanımlanması sırasında bu diller aracılığıyla alana ait bilgi kavramlarının ve bilgi kavramları arasındaki ilişkilerin tanımlanması Anlamsal Web üzerinde çalışacak uygulamaların çıkarsama yeteneğini artırmakta, bilginin daha etkin yönetilmesini, bilgi üretiminin ve denetiminin daha kolay yapılmasını sağlamaktadır.

OWL belirli topluluklara ve kullanıcılara hizmet amacıyla üç alt dile ayrılmaktadır. Aşağıda bu diller ve açıklamaları verilmektedir.

OWL Lite: Sınıflandırma hiyerarşisine ve basit kısıtlama özelliklerine ihtiyaç duyan kullanıcılar için oluşturulmuştur. Örneğin, OWL sayısallık (cardinality) kısıtlaması belirtme imkanı sunarken, kısıtlama değerleri 0 ve 1 ile sınırlandırılmıştır.

OWL DL: OWL DL (Description Logic) tanım mantığını içerir. OWL Lite tabanlıdır. Maksimum anlamlılık sağlamak amacıyla oluşturulmuştur. Makinaların otomatik olarak karar vermesine, çıkarım yapmasına yardımcı olabilecek temel mantıksal koşulları içerisinde barındırmaktadır.

OWL Full: OWL Full, maksimum anlamlılık ve hesaplama garantisi olmadan RDF sözdizimi özgürlüğünü isteyen kullanıcılar için oluşturulmuştur [27]. Örneğin, OWL Full kurallarında bir sınıf, hem bireyler topluluğu hem de kendi başına bir birey olarak düşünülür. OWL Full daha önceden tanımlanmış kelimenin anlamını arttırmaya da olanak verir. Fakat işlemlerin sonlu olmaması sebebiyle hiçbir muhakeme yazılımı OWL Full ile tanımlanan bütün özellikleri tam olarak muhakeme edememektedir.

Yukarıda bahsedilen diller sırasıyla birbirinin devamı niteliğindedir. Bu diller için aşağıdaki ilişkiler geçerlidir, fakat ilişkilerin tersi geçerli değildir [28].

- Kurallara uyan her bir OWL Lite ontolojisi, kurallara uyan bir OWL DL ontolojisi niteliğindedir.
- Kurallara uyan her bir OWL DL ontolojisi, kurallara uyan bir OWL Full ontolojisi niteliğindedir.
- Her geçerli OWL Lite sonucu, geçerli bir OWL DL sonucudur denir.
- Her geçerli OWL DL sonucu, geçerli bir OWL Full sonucudur denir.

Ontoloji geliştiricileri, kullanacağı OWL dilini seçerken “Hangi OWL dili ihtiyacımı en iyi şekilde karşılayabilir?” mantığıyla yola çıkmalıdır. Ontoloji geliştiricileri seçim yaparken daha gerçekçi kısıtlamalar tanımlamanın gerekli olup olmadığına ve muhakeme desteği durumuna göre seçim yapmalıdır.

2.5 Ontolojiler

Ontoloji terimi ilk olarak felsefe alanında "varlık bilimi" kavramına karşılık olarak kullanılmaya başlanmıştır. 90'lı yılların sonunda ise teknoloji alanında yapay zeka terimi olarak karşımıza çıkmıştır.

Gruber [29] ontolojiyi kavramlaştırmanın biçimsel ve açık belirtimi olarak tanımlamaktadır. Chen ve arkadaşlarına [30] göre, ontolojiler bağlam bağlı sistemler oluşturmak için anahtar gereksinimdir. Wang ve arkadaşları [31] için ontoloji, verinin anlamsal olarak makinaların da anlayabileceği şekilde tanımlanması ve belirtilmesidir.

Ontoloji, örtük bilgi çıkarımının yanı sıra bilgi gösteriminde de önemli bir role sahiptir. Noy ve McGuinnes [32] çalışmalarında ontoloji oluşturmanın bazı yararlarını aşağıdaki gibi belirtmektedirler:

- Bir alanda benzer ilgi ve düşünceye sahip insanlar arasında ortak bir bilgi anlayışı oluşturmayı basitleştirir.
- Örtük bilgiyi açık hale getirir.
- Alan bilgisinin yeniden kullanılabilirliğini sağlar.
- Alan bilgisinin analizini destekler.

Ontoloji, insanlara alan bilgilerini paylaşmak, değiştirmek ve tekrar kullanmak için fırsatlar sunar [33, 34]. Kısaca ontoloji, anlamsal Web' i gerçekleştirmek için kullanılan önemli bir anahtardır. Ontoloji, veri tanımlama dili olarak RDF (Resource Description Framework) yada OWL (Web Ontoloji Language) gibi farklı ontoloji tanımlama dilini kullanabilmektedir [35, 36]. Bu model veriyi Özne-Yüklem-Nesne' den oluşan RDF üçlüleri olarak saklanabilmesine olanak tanır. Özne (Subject) ve Nesne (Object), bir tanımlayıcı ile ulaşılabilecek kaynakları tanımlamak için kullanılırken, Yüklem (Property) ise Özne ve Nesne arasındaki ilişkiyi göstermek için kullanılmaktadır. Bir ontoloji sorgulama dili olan SPARQL [37] ile sorgulamalar yapılarak daha anlamlı bilgiler elde edilebilmektedir.

Herkes tarafından kolayca anlaşılabilmesi için ontolojiler iyi tasarlanmış olmalıdır. Ayrıca ontolojiler tekrar kullanılabilir, birlikte çalışabilir ve ölçeklenebilir olmalıdır.

Günümüzde ontolojiler sağlıktan iş hayatına, eğlenceden akademik çalışmalara kadar birçok alan için oluşturulmaktadır. Ontolojilerin daha etkin kullanımı için günümüz Web içeriklerinin RDF formatına dönüştürülmesi gerekmektedir.

2.5.1 Ontoloji türleri

Ontolojiler, detayları gösterme yeteneklerine ve iç yapılarının zenginliğine göre sınıflandırılmıştır. En yaygın olarak kabul gören ontoloji sınıflandırılması Guarino tarafından öngörülmüştür [38]. Şekil 2.4 bu ontoloji türlerini göstermektedir.

- **Üst düzey ontolojileri (Top Level or Upper Level Ontologies)**

Bu tip ontolojiler zaman, mekan, olay, hareket gibi genel ve soyut kavramları tanımlamak için kullanılan ontolojilerdir. Başka ontolojiler ile doğrudan veya dolaylı olarak bağlanabilirler. Bu tür ontolojilerin kapsamı çok geniştir fakat belirli bir problemi çözmek veya özel bir alan için oluşturulmazlar. Üst düzey ontolojiler için en iyi örnekler şunlardır: Wordnet [39], Sensus [40], CYC [41], SUMO [42].

- **Alan ontolojileri (Domain Ontologies)**

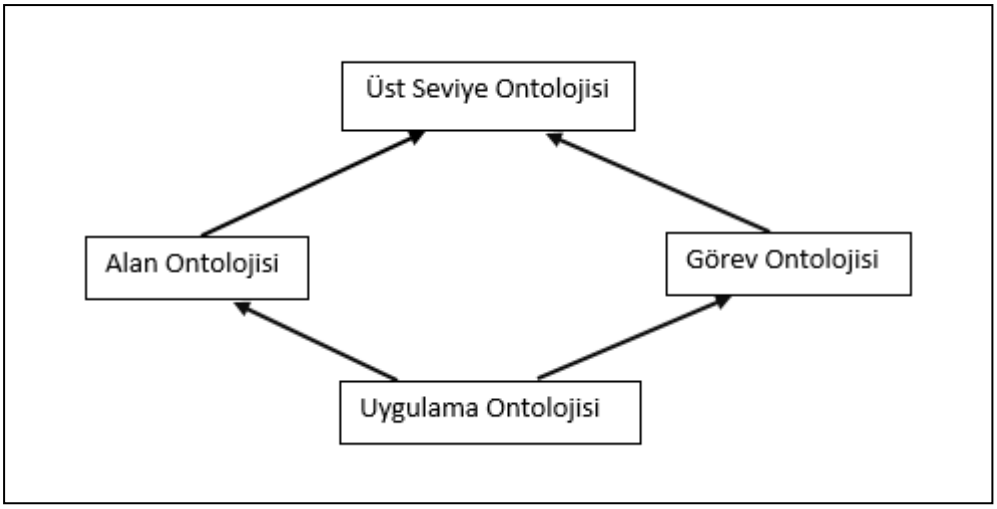
Alan ontolojileri sağlık, kültür, finans gibi belirli bir alan için tasarlanan ontolojilerdir. Alan ontolojileri üst düzey ontolojilerle ortak bir bilgi içermediği gibi kendine bu ontolojileri rahatlıkla dahil edebilir. Alan ontolojileri için en iyi örnekler şunlardır: DOLCE, OpenCyc.

- **Görev Ontolojileri (Task Ontologies)**

Görev ontolojileri, belirli görevlere ve aktivitelere yönelik üst düzey ontolojilerin tanımlanmış halidir.

- **Uygulama Ontolojileri**

Uygulama ontolojisi, belirli bir uygulama bilgisinin ihtiyacını karşılamak üzere tasarlanmış ontolojilerdir. Alan ve görev ontolojilere göre daha dar bir ontoloji olmasına karşın kendi içindeki kavramları daha detaylı olarak belirtir.



Şekil 2.4 Ontoloji türleri

2.5.2 Ontoloji Bileşenleri

Ontolojiler çeşitli bileşenlerden meydana gelmektedir. Bazı ontolojiler sadece sınıflardan, özelliklerden ve bireylerden oluşmaktadır. Bu tür ontolojiler hafif (light-weight) [43] ontolojiler olarak adlandırılır. Bazı ontolojiler ise fonksiyon terimleri, kısıtlamalar, kurallar ve aksiyomlar gibi bileşenleri de kullanmaktadır. Bu tür ontolojiler de ağır (heavy-weight) [43] ontolojiler olarak bilinir.

Aşağıda ontoloji bileşenlerinden bahsedilmektedir [44].

- **Sınıflar:** Ontolojinin en önemli bileşenlerinden biridir. Sınıflar kavramlar topluluğu, koleksiyonlar olarak da bilinir. Benzer özelliklere sahip nesnelerin bir çatı altında bulunmasını garanti eder. Sınıflar hiyerarşik yapıya sahiptirler. Bu sebeple alt ve üst sınıfları da bulunabilir. Örneğin: “Balıklar”, “Hayvan” sınıfının alt sınıfıdır, “Canlılar”, “Hayvan” sınıfının üst sınıfıdır.
- **Özellikler:** Özellikler, sınıflar arasındaki etkileşimi veya sınıf örneklerinin niteliklerini belirtmek için yaratılırlar. İki tür özellik vardır: nesne ve veri özelliği. Nesne özellikleri, bir sınıf ile başka bir sınıf arasındaki ilişkileri tanımlamak için kullanılır. Veri özellikleri, bir sınıf örneğinin alabileceği literal değerlerin atanması için kullanılır.
- **Fonksiyonlar:** İlişkilerin özel bir durumudur. Karmaşık yapıları açıklamak için ifadelerde birey (individual) terimlerinin yerine kullanılabilir.

- **Bireyler (Individuals):** Ontolojilerin en temel bileşenidir. Bireyler sınıfların örnekleri (instances) olarak da adlandırılır. Bir sınıfa ait özel üyedir.
- **Kısıtlamalar (Restrictions):** Bazı iddiaların veri girdisi olarak kabul edilmesi için önceden belirlenen sınırlamalardır.
- **Kurallar:** Kurallar if-then yapıları ile oluşturulur. Kurallar aracılığıyla daha güçlü ve özel çıkarımlar yapılabilir.
- **Aksiyomlar:** Aksiyomlar cümleleri doğru modellemek için kullanılır. Aksiyomlar ontolojideki bilginin doğruluğunu teyit ederler [45].

2.5.3 Ontoloji Editörleri

Gelişmiş araçların yardımıyla ontoloji oluşturabiliyor olmak Anlamsal Web teknolojisinin hayatımıza girmesini hızlandıracak bir etkendir. Bir ontoloji editöründen beklenen gereksinimler; ontolojileri oluşturmak için biçimsel bir dile sahip olması, ontolojileri birleştirebilmesi ve yeniden kullanabilmesi, muhakeme servislerine sahip olmasıdır.

Ontoloji editörleri insanlara ontolojilerini inşa etme kolaylığı ve esnekliği sunmaktadır. Günümüzde ticari ve ticari olmayan doksandan fazla ontoloji editörü bulunmaktadır [46]. Protégé, NeOn Toolkit, Swoop, Knoodl, Ontolingua, Neologism ve Vitro bunlardan bazılarıdır.

Protégé: Protégé, Standford Üniversitesi tarafından geliştirilmiş, basit ve karmaşık ontoloji tabanlı uygulamalar oluşturmak amacıyla kullanılan ücretsiz bir editördür. Aynı zamanda en bilinen ve yaygın olarak kullanılan ontoloji editörüdür [47]. Ontoloji oluşturmak için bir kullanıcı arayüzü bulunur. Ayrıca kullanıcıya, ontoloji dokümanını OWL, RDF ve RDF/ XML türlerinde saklama imkanı tanımaktadır. Geliştiriciler geniş bir yelpazede akıllı sistemler oluşturmak için Protégé ile ürettikleri çıktıları kural sistemlerine entegre edebilmektedir. Protégé'nin en önemli özelliği eklenti mimarisine sahip olması, buna dayalı olarak kullanışlı ve yararlı birçok eklentisinin bulunmasıdır.

Neon Toolkit: NeOn bir ontoloji mühendisliği aracıdır. Protégé'ye göre daha az eklentilere sahiptir. Genelde ağır ontolojiler oluşturmak için kullanılır.

SWOOP: Swoop basit bir ontoloji geliştirme editörüdür. Bu sebeple genelde ontoloji geliştirmeye yeni başlayanlar tarafından tercih edilir.

2.5.4 Ontoloji Yönetim Sistemleri

Ontoloji yönetim sistemleri ontoloji oluşturma, var olan RDF veri üzerinde çalışma, sorgulama, verileri saklama gibi işlemleri programlama üzerinden gerçekleştiren çerçevelerdir.

Apache Jena: Jena, Java programlama dilini kullanan, ontoloji geliştirmeye yardımcı bir araçtır [47]. Bazı özellikleri şunlardır:

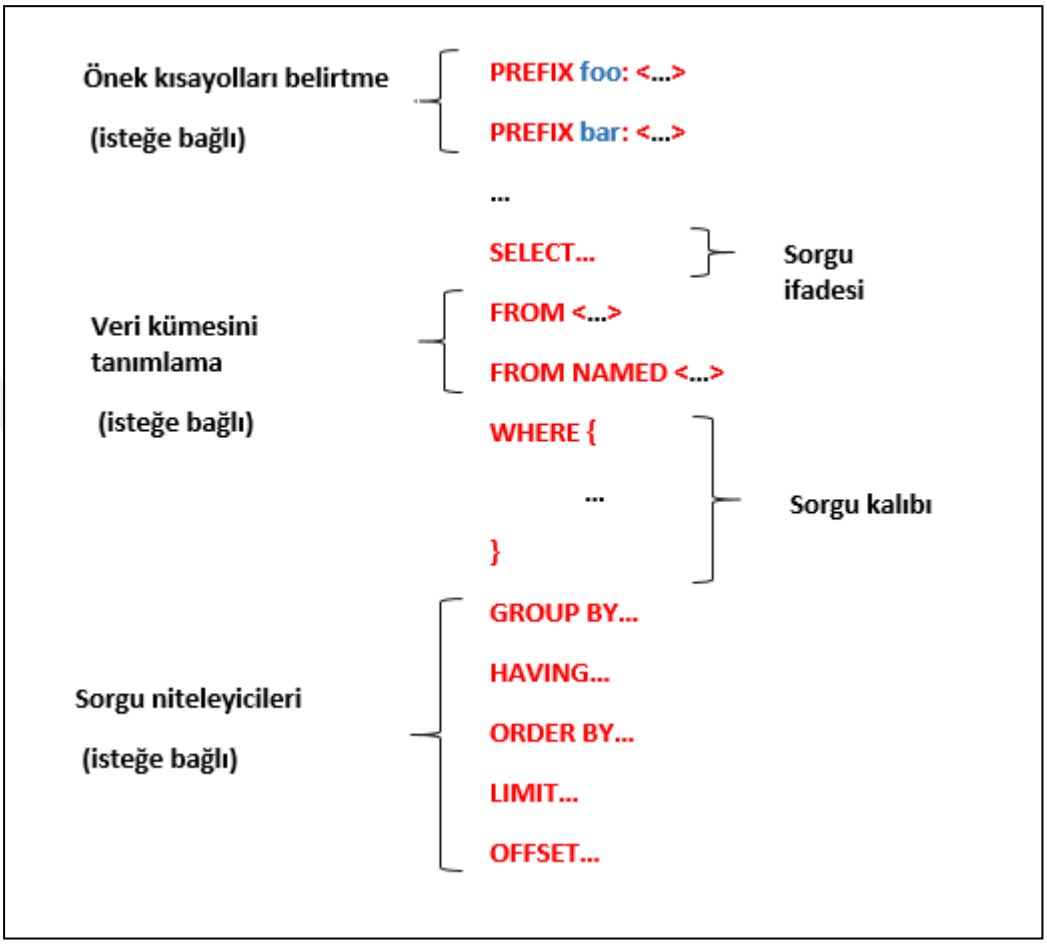
- RDF, RDF/XML, Turtle gibi veri formatlarında okuma ve yazma yapabilir.
- RDF verilerini saklamak için veritabanı bağlantı desteği sağlar.
- OWL dilini destekler.
- Ontolojileri sorgulama imkanı sunar.

Sesame: RDF verilerini depolamaya, sorgulamaya ve bilgi çıkarımına yardımcı olan açık kaynak kodlu bir çerçevedir [48]. RDF üçlü depolama (RDF Triple Store), sorgu gerçekleştirimi ve bilgi yönetimi için SPARQL Sorgu Endpoint'i ve Web uygulaması imkanı sunar.

2.5.5 SPARQL (Simple Protocol and RDF Query Language)

RDF üçlüleri (triples), Web içeriklerini belli standartlara göre düzenlemek için oluşturulmuş bir yapıdır. RDF üçlüleri kaynaklar arasındaki ilişkileri belirtmede kullanılır. RDF kaynağı (resource) ve kaynağın özellikleri (property) kullanılarak diğer RDF kaynaklarına erişilebilir. Bu sayede kullanıcılar ihtiyaç duyduğu bilgiye kısa zamanda ulaşabilmektedir. Fakat Web'deki verilerin büyüklüğü sebebiyle bilgiye erişim çok zaman ve emek gerektirmektedir. Bu sebeple RDF standartlarına göre oluşturulan içerikleri incelemek ve sorgulamak için anlamsal sorgulama dili olan SPARQL (Simple Protocol and RDF Query Language – Basit Protokol ve RDF Sorgulama Dili) geliştirilmiştir. SPARQL, RDF formatında saklanan veriler için standart bir sorgulama dili sağlar [49, 37]. Ayrıca, SPARQL birleştirmeyi (aggregation), alt sorgulamayı (subquery),

olumsuzlukları (negation), ifadelerle değer oluşturmayı ve sorgu kısıtlamayı destekleyen bir yapıya sahiptir [37]. SPARQL, ilişkisel veritabanı sistemleri sorgulama dili olan SQL ile benzerlik göstermektedir. Şekil 2.5 SPARQL sorgularının yapısını göstermektedir.



Şekil 2.5 SPARQL sorgu yapısı

SPARQL tarafından desteklenen ve yaygın olarak kullanılan 4 sorgu ifadesi vardır [37]. Bunlar; “Select”, ”Ask”, “Describe” ve “Contruct” ’tır. Bu sorgu ifadelerinin açıklamaları şöyledir:

- ASK sorgu ifadesi, sorgu kalıbının verilerle eşleşip eşleşmediğini gösteren bir boolean değer döndürmektedir.
- CONSTRUCT sorgu ifadesi, sorgu kalıbının veri setleri içinde eşleşmesiyle birlikte sonuçları tanımlanan RDF grafiği formatında döndürmektedir.
- DESCRIBE sorgu ifadesi, RDF veri kaynağını içeren bütün RDF üçlülerini döndürmektedir.

- SELECT sorgu ifadesi, sorgu kalıbına uyan verilerin isteğe bağlı olarak bir kısmını veya hepsini tablo formatında döndürmektedir.

Bu ifadelerin örnekleri Şekil 2.6, Şekil 2.7, Şekil 2.8 ve Şekil 2.9' da verilmektedir.

```
PREFIX foaf: < http://xmlns.com/foaf/01/ >.
ASK     { ?x     foaf:name      "Alice".  }
```

```
true
```

Şekil 2.6 ASK sorgu ifadesi örneği ve sorgu sonucu

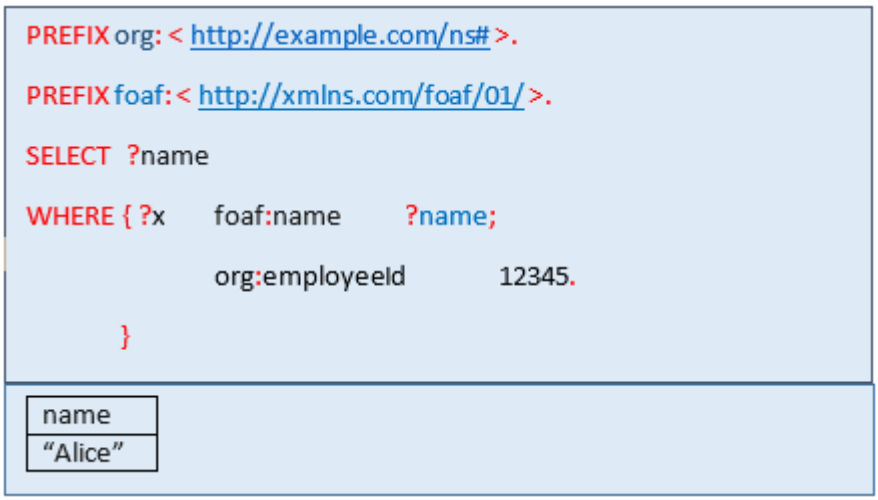
```
PREFIX org: < http://example.com/ns# >.
CONSTRUCT { ?y    org: isGrandDaughterOf     ?x.}
WHERE { ?x   org:isFatherOf    ?child.
        ?child    org:isFatherOf     ?y.
          ?y     org:gender        org:female.
        }
```

```
org:Monica org: isGrandDaughterOf org:David
```

Şekil 2.7 CONSTRUCT sorgu ifadesi ve sorgu sonucu

```
PREFIX foaf: < http://xmlns.com/foaf/01/ >.
DESCRIBE   ?audi
WHERE { ?audi          foaf:name       "AUDI Motor Company". }
```

Şekil 2.8 DESCRIBE sorgu ifadesi örneği



Şekil 2.9 SELECT sorgu ifadesi ve sorgu sonucu

Şekil 2.6 ’da ‘“Alice adında bir kişi var mı? “ sorusunun cevabı olarak “True” döndürülmektedir. Şekil 2.7’ de “CONSTRUCT” kısmında belirtilen RDF grafik formatında torun ve dede bilgisi döndürülmektedir. Şekil 2.8’ de “name” ,özelliği “AUDI Motor Company” olan kaynağın tüm bilgilerine ulaşılmak istenmektedir. Şekil 2.9’ da “employeeID” özelliği “12345” olan kişinin adına ulaşılmaktadır.

Sorgu niteleyicileri SPARQL sorgulamalarını daha da esnek hale getirmek için oluşturulmuştur. En çok kullanılan sorgu niteleyicileri şunlardır:

- GROUP BY, , bir veya daha fazla özellik üzerinden toplama (aggregation) işlemi yapılmasına olanak sağlar.
- HAVING, gruplandırılmış sonuçlar üzerinden filtreleme yapmaya yardımcı olur.
- ORDER BY, sonuç dizisinin sırasını belirler.
- LIMIT, dönecek olan sonuç sayısına üst sınır koymak için kullanılır.
- OFFSET, belirtilen sayıdan sonraki sonuçların alınması için kullanılır.

BÖLÜM 3

TEMEL BİLGİLER

Bu bölümde, OKM (Ontology Aided Knowledge Mapping) sistemini oluşturan teknolojilerle ilgili temel bilgiler yer almaktadır.

3.1 Temel Bilgiler

Temel bilgiler bölümünde standart ontolojiler ve Bilgi haritalama yöntemi anlatılmaktadır.

3.1.1 SKOS Ontolojisi

Web' de yayınlanan birçok ontoloji olmasına rağmen veri bakımından zenginleştirilen ve standart bir seviyeye getirilen sayılı ontolojiler vardır. Basit Bilgi Organizasyon Sistemi (Simple Knowledge Organization System - SKOS) W3C Recommendation tarafından standartlaştırılan bir ontolojidir. Thesauri, sınıflandırma şemaları, taksonomiler ve konu başlık listeleri gibi kontrollü sözlüklerin kavram içeriklerini ifade etmek için kullanılmaktadır [50]. SKOS, sözlük oluşturmak ve bu sözlükleri Anlamsal Web'e transfer etmek için evrensel ve kolay bir yol sunmaktadır. SKOS ontolojisi kullanılarak farklı kişiler tarafından oluşturulan sözlükler birbiriyle uyumlu bir şekilde birleştirilebilmektedir. SKOS gibi genel bir gösterim modeli kullanıldığında başka sözlüklerin birlikte kullanım ve paylaşım maliyetleri azalmakta, gereksiz zaman tüketimi engellenmektedir[50]. SKOS sözlüğünün en temel sınıfı skos:Concept' tir. Fikirler, anlamlar, nesneler bu sınıfta gösterilebilirler. Kavramlar arasındaki hiyerarşik anlamsal

ilişkileri göstermek için skos:narrower ve skos:broader propertyleri kullanılırken, ilişkisel ilişkileri göstermek için de skos:related propertysi kullanılır [50].

3.1.2 Bilgi Haritalama

Bilgi haritalama, tanımlanmış kavramların ve bu kavramlar arasındaki ilişkilerin görsellik yardımıyla kullanıcıya sunulması olarak düşünülebilir. Bilgi haritaları, bilginin anlaşılabilmesini kolaylaştıran araçlardır. Kavramları göstermek için blok, ilişkileri göstermek için çizgi gibi basit şekiller kullanılmaktadır [51, 52]. Bilgi haritalama, esas düşünce ve fikirlerin görsel tasvirini kullanıcıya sunmak bakımından önemlidir. Görsellik ise bilginin daha uzun süre akılda kalıcılığını arttırmaktadır. İnsanlar, bilgi haritalarını kullanarak bilginin nasıl oluşturulduğunu, bilginin hangi ilişkilere sahip olduğunu ve bilgi kaynağının nerede olduğunu öğrenebilir [52].

Bilgi haritaları bilgileri organize hale getirir, kavramların anlamlılığını arttırır, yanlış anlamaları en aza indirir ve öğrenmeyi üst seviyeye çıkarmaya zemin hazırlar. Bilgi haritalama, detaylar veya kavramlar arasındaki benzerlikleri ortaya çıkarmaktadır.

BÖLÜM 4

OKM SİSTEMİ REFERANS ONTOLOJİLERİ

Bu bölümde OKM sisteminin gerçeklenmesi için kullandığımız referans ontolojilerine değinilmektedir.

4.1 Sözlük Ontolojisinin Oluşturulması

Alt ve üst kavram, bir varlığın hangi üst varlık sınıfına ait olduğu bilgisini ifade etmektedir. Yani alt ve üst kavramlar, canlı ve cansız varlıkları, olayları ve düşünceleri sınıflandırmak için gruplara verilen adlardır. Bir kavramın alt ve üst kavram ilişkilerinin bilinmesi bilginin özümsenmesi, anlamlandırılması ve yapılandırılması aşamalarındaki en önemli unsurlardan biridir.

Üst kavramların oluşturulması genelleme süreci olarak düşünülebilir. Örneğin: ”Sıvı” sözcüğü “su, çay, süt ” sözcüklerinin üst kavramı olarak belirtildiğinde sütün, suyun bir sıvı olduğunu anlaşılmaktadır.

Sözlük Ontolojisi, kavramları ve kavramların alt ve üst kavram ilişkilerini tutan bir ontolojidir. Sözlük ontolojisinin oluşturulmak istenmesinin sebebi, ontoloji tabanlı bilgi haritalama sistemimizin anlaşım ve çıkarım gücünü arttırmaktır.

Ontoloji de alt ve üst kavramların kullanılması bize birçok yarar sağlayacaktır. Bunlar:

- Bir kavramın alt ve üst kavramının belirlenmesi, bilgisayarların da kavram bilgilerini daha iyi işlemesine ortam hazırlayacak ve kavramlar arasındaki bilginin anlaşılabilirliğinin arttırılmasına olanak sağlayacaktır.

- Bir kavramın alt ve üst kavramlarının belirlenmesi, kavram hakkında bize daha geniş bilgiler sunacak ve kavramın bilgisayarlar tarafından da anlaşılmasını kolaylaştıracaktır.
- Üst kavramlarının belirlenmesi neticesinde aramalarda daha iyi ve alakalı sonuçlar elde edilebilecektir.
- Kavramların alt ve üst bilgilerinin var olması makinaların bilgi çıkarım gücünü arttıracaktır.
- Alt ve üst kavramların belirlenmesi iki sözcük arasındaki ortak özelliğe erişmemizi sağlayacaktır.

Varlıkları benzer özelliklerine göre gruplara ayırmadığımız da, birbirinden ayırt edilmemiş ve birbirleriyle ilişkileri belirlenmemiş binlerce kavramla karşı karşıya kalmaktayız. Bu durum sistemli bir edinimin önünü kapatmakta ve bilgi oluşumunu engellemektedir. Buna bağlı olarak tam olarak kavranamayan bilgi transfer bakımından zorluk teşkil etmektedir.

Sonuç olarak, sözcükler arasındaki alt ve üst kavram ilişkileri kurularak yeni bilgiler inşa edilebilmektedir. Bu alt ve üst kavram ilişkilerinin, bilgisayarlarla işlenebilir sözlük kullanarak kavramlar arasındaki anlamsal ilişkilerin belirlenmesi bizim işimizi hafifletecektir. Bu sebeple Güngör O. ve Güngör T.'nin [53] çalışmalarından yola çıkarak, sistemimiz tarafından kullanılacak olan Sözlük Ontolojisi yaratımında kullanılmak üzere alt ve üst kavram ilişkilerini elde etmeye çalışmaktayız. Böylelikle ontoloji tabanlı bilgi haritalama sistemimizin anlamsal veri çıkarım gücünün arttırılması hedeflenmektedir.

4.1.1 Üst kavramların çıkarılması

[53]'daki çalışma da Türk Dil Kurumu (TDK) tarafından yayımlanan genel Türkçe sözlüğün elektronik sürümü [54] kullanılarak, sözlükteki sözcüklerin anlamlarının otomatik olarak inceleyen anlamsal bir hiyerarşik yapı oluşturan bir yöntem kullanılmıştır. Bu çalışma, Türkçe bir sözlükteki sözcüklerin anlamlarının kullanılması ile alt ve üst kavram hiyerarşisini oluşturan ilk çalışma olarak literatürde yerini almıştır. Çalışmadaki üst kavramları bulan algoritma 83.000 kavram üzerinden 78.000 tanesi için

en az bir üst kavram bulmuştur ve bunun sonucunda üst kavram çıkarma oranı %94'tür.

İngilizce Wordnet'in [39] titiz çalışmalar sonucunda oluşturulması, neredeyse hatasız bir veri tabanı olması ve en üstteki kavramların dildeki bütün kavramları kapsayacak şekilde belirlenmesi sebebiyle Wordnet'in en üst düğümlerinin Türkçe çeviri karşılıkları, hiyerarşinin en üst düğümleri olarak belirlenmiştir. Wordnet'te en üst 25 kavram bulunurken, İngilizce ve Türkçe dilleri arasındaki farklılıklardan dolayı Türkçe'deki hiyerarşik yapıda en üstte 21 düğüm bulunmaktadır. Bu düğümler şunlardır: hareket, hayvan, kişi-insan, özellik, bitki, gövde-vücut, sahip, kavram-anlama-bilgi, süreç-işlem, miktar, olay-iş, duygu-his, şekil, yiyecek-yemek, durum, grup-topluluk-küme, madde, yer, zaman-vakit, amaç.

Üst kavramları belirleyen kurallar üç gruba ayrılmaktadır; ismin yüzey biçimine göre, ismin sözlükteki kategorisine göre ve ismin sözlükteki tanımına göre. Referans alınan çalışmada toplamda 11 kural belirlenmiş, bizim çalışmamıza 10 kural adapte edilmiştir. Kuralların anlaşılabilmesi maksadıyla her kural için açıklamalar ve örnek kelimeler aşağıda Çizelge 4.1' den Çizelge 4.10' a kısımda verilmektedir.

Bilgi tekrarını ve döngü oluşumunu engellemek amacıyla hiyerarşik bir yapı olan ağaç yapısı kullanılmaktadır. Böylece arama algoritmaları performansı bakımından ağaç yapısının üstünlüklerinden faydalanmış olunmaktadır.

Çizelge 4.4 Üst Kavram belirlemede kullanılan 1.Grup 1. Kural' ın Tanımı ve Örneği

Grup 1 Kural 1: İncelenen kelime 'bilimi' ile bitiyorsa bu sözcüğün üst kavramı 'bilim'dir.
Örnek:
Kelime: Kuş bilimi
Kategori: İsim
Tanım: Kuşları inceleyen bilim, ornitoloji.
Üst kavram: Bilim

Çizelge 4.5 Üst Kavram belirlemede kullanılan 2.Grup 1. Kural' ın Tanımı ve Örnekleri

Grup 2 Kural 1: İncelenen kelimenin tanımı aşağıdaki üç grup kelime öbeğinden herhangi birisiyle bitiyorsa, grubuna göre belirtilen işlemler uygulanır. Kelime öbeğinden önceki kelime işlenip üst kavram elde edilir.

Grup 1:

'olanlardan her biri'

'olanlardan biri'

'olanlardan bazısı'

Örnek:

Kelime: Ülküdaş

Kategori: İsim

Tanım: Aynı ülkeye bağlı olanlardan her biri.

Üst kavram: Bağ

Grup 2:

'her biri'

'biri'

'bazısı'

Örnek:

Kelime: Ünite

Kategori: İsim

Tanım: Dersin bölümlerinden her biri.

Üst kavram: Bölüm

Grup 3:

'S ve/veya S herbiri'

'S ve/veya S biri'

'S ve/veya S bazısı' (S tekil kelimeleri ifade etmektedir.)

S sözcükleri işlenip üst kavram elde edilir.

Örnek:

Kelime: Sorti

Kategori: İsim

Tanım: Elektrik tesisatında lamba ve fiş konacak kolların her biri.

Üst kavram: Lamba, fiş

Çizelge 4.6 Üst Kavram belirlemede kullanılan 2.Grup 2. Kural' ın Tanımı ve Örneği

Grup 2 Kural 2: İncelenen kelimenin tanımı 'kimse' ile bitiyorsa bu sözcüğün üst kavramı 'kişi'dir. Örnek: Kelime: Fırıncı Kategori: İsim Tanım: Fırın işleten kimse. Üst kavram: Kişi

Çizelge 4.7 Üst Kavram belirlemede kullanılan 2.Grup 3. Kural' ın Tanımı ve Örneği

Grup 2 Kural 3: İncelenen kelimenin tanımı 'işi' ile bitiyorsa bu sözcüğün üst kavramı 'iş'tir. Örnek: Kelime: Dirsekleme Kategori: İsim Tanım: Dirsekleme işi. Üst kavram: İş

Çizelge 4.8 Üst Kavram belirlemede kullanılan 2.Grup 4. Kural' ın Tanımı ve Örneği

Grup 2 Kural 4: İncelenen kelimenin tanımı 'tümü ile bitiyorsa 'tümü' kelimesinden önceki kelime işlenip üst kavram elde edilir. Örnek: Kelime: Medya Kategori: İsim Tanım: İletişim iletişim araçlarının tümü. Üst kavram: Araç

Çizelge 4.9 Üst Kavram belirlemede kullanılan 2.Grup 5. Kural' ın Tanımı

Grup 2 Kural 5: İncelenen kelimenin tanımı 'değil' sözcüğü ile bitiyorsa kelimenin işlenmesi bitirilir. Böyle durumlarda, 'değil' sözcüğünden önceki kısımlarda üst kavram bulunmadığı görülmüştür.

Çizelge 4.10 Üst Kavram belirlemede kullanılan 2.Grup 6. Kural' ın Tanımı ve Örneği

Grup 2 Kural 6: İncelenen kelimenin tanımı 'hepsi' ile bitiyorsa, 'grup' kelimesi üst kavram olarak belirlenir. Örnek: Kelime: Ev halkı Kategori: İsim Tanım: Bir evde yaşayanların hepsi Üst kavram: Grup

Çizelge 4.11 Üst Kavram belirlemede kullanılan 2.Grup 7. Kural' ın Tanımı ve Örneği

Grup 2 Kural 7: İncelenen kelimenin tanımı 'SÖ ve/veya SÖ' ile bitiyorsa (SÖ sözcük öbeğini ifade etmektedir), sözcük öbeklerinin son kelimeleri işlenerek üst kavram elde edilir. Örnek: Kelime: Vakıf toprağı Kategori: İsim Tanım: Vakfın mülkiyeti altında olan toprak veya arazi. Üst kavram: Toprak, arazi

Çizelge 4.12 Üst Kavram belirlemede kullanılan 2.Grup 8. Kural' ın Tanımı ve Örneği

Grup 2 Kural 8: İncelenen kelimenin tanımı 'vb.' ile bitiyorsa 'vb.' kelimesinden önceki kelime işlenerek üst kavram elde edilir. Örnek: Kelime: Açık taşıt Kategori: İsim Tanım: Üstü örtülmemiş araba, otomobil vb. Üst kavram: Araba, otomobil

Çizelge 4.13 Üst Kavram belirlemede kullanılan 2.Grup 9. Kural' ın Tanımı ve Örneği

Grup 2 Kural 9: İncelenen kelime 'botanik' veya 'zooloji' kategorilerine aitse, tanımdaki ilk kelimeden ayrılma eki ('-dan' eki) atılarak üst kavram elde edilir.
Örnek:
Kelime: Mandalina
Kategori: İsim, botanik
Tanım: Turunçgillerden, ılıman iklimlerde yetişen ve portakala çok benzeyen bir ağaç
Üst kavram: Turunçgil

4.1.2 TDK Türkçe Sözlüğü'ndeki Kelimelerin Elde Edilmesi

Sözcüklerin sözlükteki anlamlarından yola çıkılarak alt ve üst kavramlarını bulan kural tabanlı algoritmanın uygulanması için elektronik sözlükte bulunan kelimeler, kelimelerin kategorileri ve kelimeye ait tanımlar elde edilmelidir. Elektronik sözlük olarak Türk Dil Kurumu'nun hazırlamış olduğu TDK Güncel Sözlüğü seçilmiştir.

Kelimelerin kategorilerinin ve anlamlarının TDK'nın adresinden çekilebilmesi için kelimelerin elimizde bulunması gerekmektedir. Bunun için TDK'nın Web sayfasındaki "Sözlükler" kısmındaki Yazım Kılavuzu kısmı kullanılmıştır. Girdi kısmına yazdığımız harfle, heceyle veya kelimeyle başlayan bütün kelime ve kelime grupları Yazım Kılavuzu tarafından listelenmektedir. Örneğin, "c" harfi girildiğinde "c" ile başlayan, veri tabanında bulunan bütün kelime ve kelime grupları listelenir. Bu şekilde bir program yardımıyla TDK Güncel Sözlüğü'ndeki bütün kelimeler alınmıştır. Sonrasında, hazırlanan başka bir programla bu kelimeler sırasıyla TDK Güncel Sözlüğü'nde aranarak, isim kategorileri ve anlamları kaydedilmiştir. Bu şekilde yapılmasının sebebi, aranacak kelime TDK Güncel Sözlüğü'nde yer almıyorsa listenin boş dönmesidir. Listenin boş dönmesini engelleyebilmek maksadıyla TDK'da var olan bütün kelimeler bilinmelidir.

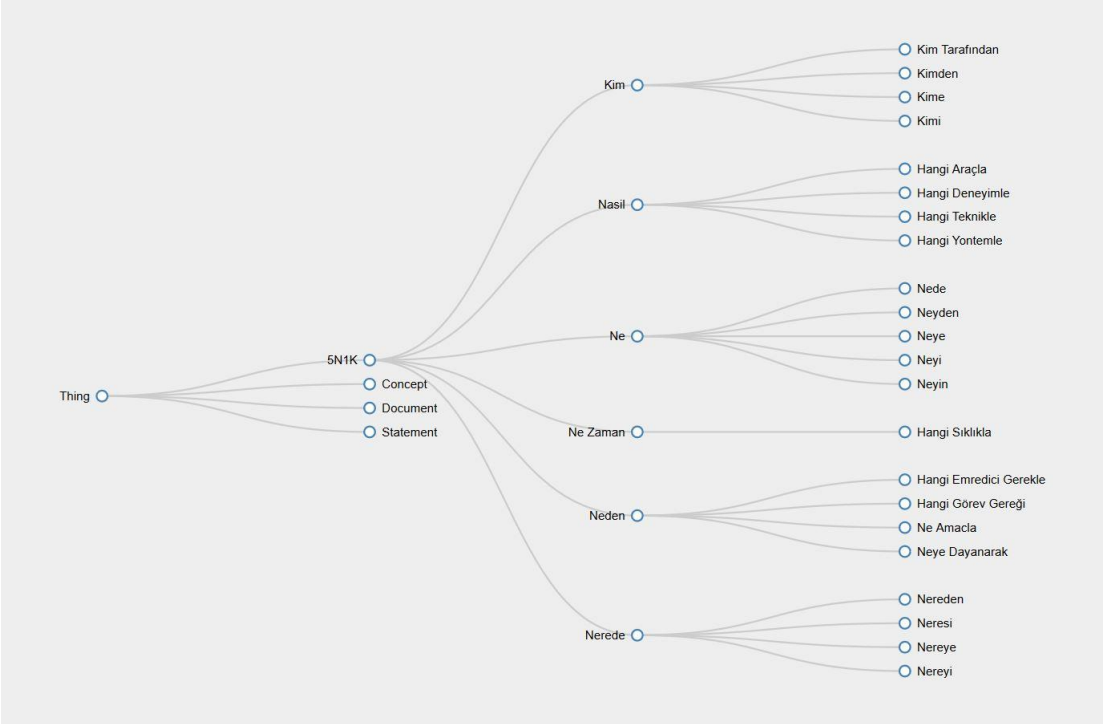
4.2 OKM Ontoloji Sınıfları ve Sınıf Hiyerarşisinin Belirlenmesi

Ontoloji oluşturmanın bir kaç adımı vardır. Bunlardan biri sınıfları ve sınıf hiyerarşisini tanımlamaktır. Sınıflar kaynakları gruplamaya yardımcı olmaktadır. Kavramların taksonomisini oluşturmak için farklı metodlar literatürde önerilmektedir [49].

Kullandığımız referans ontolojisinin belirlenmesinde, bu çalışmalardan, “Top-Down” yaklaşımı kullanılmıştır. Bu yaklaşımda, ontoloji, önce üst sınıflar tanımlanarak alt sınıflara doğru inşa edilir. Bu yol en genel sınıftan en özel sınıfa ilerlenerek ontolojinin oluşturulmasını önermektedir. Önerilen ontoloji dört temel sınıfa odaklanmaktadır; ‘‘Statement’’, “Document”, “Concept” ve ‘‘5W1H’’. Bu sınıfların ne amaçla oluşturulduğu Çizelge 4.11’ de açıklanmaktadır.

Çizelge 4.14 OKM Ontolojisinde bulunan üst seviye sınıflar

Sınıflar	**Açıklama**
Statement	Ele alınan kavramın doküman içinde hangi konumda bulunduğunu tanımlar.
Document	Kullanılan dokümanları tanımlar.
Concept	Doküman içindeki bilgi varlıklarını (sözcükleri/kavramları) tanımlar.
5W1H	Kavramların cevap olabileceği soru çeşitlerini tanımlayan üst sınıfı tanımlar.



Şekil 4.1 OKM Referans Ontolojisi sınıfları ve hiyerarşisi

Şekil 4.1, aynı zamanda, “Top-Down” yaklaşımı ile oluşturmuş olduğumuz, ontolojinin hiyerarşik yapısını da resmetmektedir. Ontolojinin örneklerinin (instances) oluşturulması için, dokümanlardan çıkartılacak RDF üçlülerinin saklanması ve takip edilmesi amacıyla, sütunlarında 5W1H soruları bulunan ve satırlarında bilgi varlıkları yer alan bir matris oluşturulmuştur. Oluşturulan matrisin örnek şematik gösterimi Çizelge 4.12’ de verilmektedir. Bu çizelgede, bir doküman içinde yer alan iki cümlenin (“İstanbul Türkiye’ de bulunmaktadır” ve “İstanbul Şehir olmuştur”), NLP kütüphaneleri kullanılarak analiz edilmesi ve matrisin ilgili hücrelerinin doldurulduğu şeklin gösterimi bulunmaktadır.

Çizelge 4.12 Anlamsal Bilgi Çıkarım Matrisi

Concept	**Predicate**	**Document**	**Statement**	**5W1H**					
				Kim	**Ne**	**Nerede**	**Ne zaman**	**Neden**	**Nasıl**
İstanbul	bulunmak	Document-1	Statement-1			Türkiye			
İstanbul	olmak	Document-1	Statement-2		Şehir				

Ontolojide yer alan, 5W1H ana sınıfının, alt sınıfları ‘‘Kim’’, ‘‘Ne’’, ‘‘Nerede’’, ‘‘Ne zaman’’, ‘‘Neden ’’ ve ‘‘Nasıl’ soruları olacak şekilde belirlenmiştir. Şekil 4.1’ de görüldüğü gibi bu alt sınıflarda kendi içlerinde alt sınıflar barındırmaktadır. Örneğin, “Nerede” sınıfı için, “Nereden”, “Nereyi”, “Nereye” gibi alt sınıflar belirlenmiştir. Bu alt-sınıflar kelime veya kelime ikililerinin hangi sorulara cevap niteliği taşıdığını gösterebilmek amacıyla tanımlanmışlardır.

4.3 OKM Referans Ontolojisinin Özelliklerinin (Properties) Belirlenmesi

Ontolojide kavramlar arasındaki ilişkileri tanımlamak için özellikler (property) kullanılmaktadır. OWL ontoloji dilinde, özellikler, object property ve datatype property olmak üzere ikiye ayrılır. Sınıflar arasındaki ilişkiyi göstermek için object property, bir sınıf ile o sınıfın alabileceği literal değerler arasındaki ilişkiyi göstermek için ise datatype property kullanılır. Tanım kümesi (domain) ve değer kümesi (range), özelliklerin kısıtlamalarını ifade etmektedir. OKM Referans ontolojisinde, 5 adet object property, 1 adet datatype property belirlenmiştir. Çizelge 4.13, bu property ve kullanım

amaçlarını göstermektedir. Önerilen referans ontolojisinin ileride başka ontolojilerle uyumluluk sağlaması amacıyla standart SKOS ontolojisinin en çok kullanılan propertyleri kullanılmıştır.

Çizelge 4.13 OKM Ontolojisinde bulunan özellikler (Properties)

Property tipi	**Property**	**Domain**	**Range**	**Açıklama**
DataType	hasText	Statement	xsd:String	Kavramın/Bilgi Varlığının varsa sahip olduğu metni ifade eder.
Object	in	5W1H	Statement	Kavramın/Bilgi Varlığının hangi dokümanda nerde olduğunu ifade eder.
Object	isContainedBy	Statement	Document	Kavramın/Bilgi Varlığının hangi dokümanda bulunduğunu ifade eder.
Object	skos:narrower	skos:Concept	skos:Concept	Bir kavramın alt kavramını ifade eder.
Object	skos:broader	skos:Concept	skos:Concept	Bir kavramın üst kavramını ifade eder.
Object	skos:altLabel	skos:Concept	skos:Concept	Bir kavramın eşanlam kavramını ifade eder.

4.4 Ontoloji Örneklerinin Oluşturulması

Sınıfların tanımlanması ve özelliklerle (property) ilişkilendirilmesi aşamalarından sonra örneklerin oluşturulması aşaması gerçekleştirilmiştir. Burada, Türkçe NLP (Natural Language Processing) kütüphanesi kullanılmıştır. NLP Kütüphanesi olarak Zemberek [55, 56] kullanılmıştır. Bir doküman içindeki tüm cümleler tespit edilmekte, aynı cümle içinde yer alan kavramlar (özne ve nesneler olarak) ve yüklemler birbiriyle ilişkilendirilmiş olarak ayrıştırılmaktadır. Nesnelerin aldıkları eklere dayalı olarak, 5W1H sorularından hangisine karşılık gelebileceği belirlenmekte ve Çizelge 4.12' de verilen Bilgi Çıkarım Matrisin ilgili hücresi güncellenmektedir. Matris' in doküman kümesinde yer alan tüm bilgi varlıkları için doldurulmasının ardından, matris' te yer alan her bir satır, RDF üçlülerine dönüştürülmektedir. Bu amaçla bir ontoloji işleme kütüphanesi kullanılmıştır. Ontoloji işleyici kütüphane olarak Apache Jena Ontoloji [56] kütüphanesi seçilmiştir. Bu yöntem kullanılarak oluşturulmuş OKM Referans Ontolojisi örnekleri

Şekil 4.2’ de verilmektedir. “Ders programı Bölüm Panosunda duyurulur” cümlesine ait RDF üçlüleri bu şekilde gösterilmektedir.

s	p	o
<okm:ders_program>	<skos:altLabel>	<okm:müfredat_program>
<okm:ders_program>	<skos:narrower>	<okm:program>
<okm:ders_program>	<okm:in>	<okm:s1>
<okm:ders_program>	<rdf:type>	<rdf:subject>
<okm:bölüm_pano>	<okm:in>	<okm:s1>
<okm:bölüm_pano>	<rdf:type>	<rdf:object>
<okm:bölüm_pano>	<rdf:type>	<okm:Nerede>
<okm:duyurulmak>	<okm:in>	<okm:s1>
<okm:duyurulmak>	<rdf:type>	<rdf:predicate>

<okm:ders_program>	<okm:duyurulmak>	<okm:bölüm_pano>

Şekil 4.2 RDF üçlüleri

4.5 Veri Seti

Bu çalışmada, sistemin uygulanabilirliğini göstermek için eğitim [58], sağlık ve spor [59] alanlarına ait doküman kümesi ele alınmıştır. Doküman veri kümesi üzerinde, Zemberek kütüphanesinin, uzun ve yapısal olarak karışık cümleleri çözümleyememesi nedeni ile cümlelerin sadeleştirilmesi işlemi gerçekleştirilmiştir. Zemberek, Türkçe dili için, tamamen Java ile geliştirilen, açık kaynak çözümleyici uygulaması olup ve imla kontrolü, hatalı kelimeler için öneride bulunma, heceleme, deascifier ve hatalı kod temizleme fonksiyonlarını içermektedir [56].

Veri setinde, eğitim, sağlık ve spor alanlarına ait küçük, orta ve büyük ölçekte sırasıyla 50, 100 ve 150 cümle yer almaktadır. Daha genel bir bilgi haritalama gerçekleştirmek adına, kavramlar çekim eklerinden arındırılarak kökleriyle kullanıldığında veri setinde toplamda 394 adet isim tipinde kelime ve kelime grubu, 116 adet yüklem bulunmaktadır. Çizelge 4.14’ de veri seti hakkındaki genel bilgiler görülebilmektedir. Ayrıca, üst kavramları bulan algoritma 83.000 kavram üzerinden 78.000 tanesi için en az bir üst kavram bulmuştur. Fakat, Zemberek’in çözümleyemediği kavramlar

içerisinden çıkarıldığında 61.400 tane alt ve üst kavram Sözlük Ontolojisinde kullanılmaktadır.

Zemberek bir sözcük için birden fazla çözümleme sunmaktadır. Sunduğu ilk öneri genellikle daha doğru kabul edildiği için, ilk çözümlemeler göz önüne alınmıştır. Zemberek, bazı sözcüklerin kök tipleri isim olmadığı halde isim olarak döndürmektedir (örneğin “ve” bağlaç olduğu halde isim olarak etiketlenmiştir). Bu kontrolü sağlamak amacıyla kelimenin, TDK [54] Türkçe Güncel Sözlükten elde edilen isim tipli kelimeler listesinde bulunmasına da bakılmaktadır.

Çizelge 4.14 Veri seti

Alan	**Doküman Sayısı**	**Cümle Sayısı**	**Kavram Sayısı**
Eğitim	5	50	131
Sağlık	7	100	190
Spor	8	150	225

BÖLÜM 5

UYGULAMA DETAYLARI

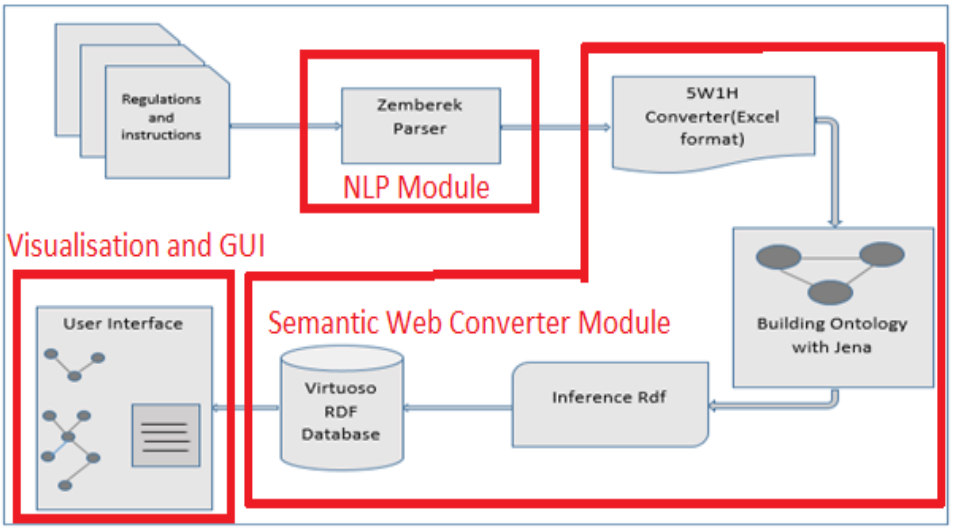
Bu bölümde, geliştirilen Bilgi Haritalama Sistemi'nin yazılım mimarisi ve yazılım gerçeklenmesinde kullanılan teknik alt yapı açıklanmaktadır.

5.1 Sistem Mimarisi

Şekil 5.1' de sistemin mimari yapısını gösterilmektedir. Sistem, 1) Doğal Dil İşleme (NLP) Modülü, 2) Anlamsal Web Dönüştürücüsü Modülü, ve 3) Görselleştirme ve Kullanıcı Arayüzü olmak üzere 3 ana bölümden oluşmaktadır. Bu bölümler aşağıda ayrı alt-bölümler halinde detaylı olarak anlatılmaktadır. Gerçeklenirken farklı teknolojiler kullanılmaktadır.

Şekil 5.1' de resmedilen sistemde, ilk olarak NLP Modülü, bir doküman kümesini girdi olarak almaktadır. NLP Modülü, doküman kümesi içinde yer alan sözcükleri ve sözcükler arasındaki ilişkileri (yüklemler) cümle cümle, yukarıda Çizelge 4.12' de şematik gösterimini verdiğimiz Bilgi Çıkarım Matrisine aktarmaktadır. NLP Modülü, bu işlemi yaparken, bu çalışma kapsamında oluşturduğumuz Sözcük Ontolojisi (Türkçe sözcükler arasındaki temel eş-anlamlılık, dar-kavram, geniş-kavram ilişkilerini içeren Sözcük Ontolojisi), Sözcük Kontrollü sözlüğü, Yüklem Kontrollü sözlüğü gibi yardımcı kaynaklar kullanmaktadır. Bu sayede, Bilgi Çıkarım Matrisine, sadece dokümandaki cümle yapılarından gelen özne-yüklem-nesne ilişkisi içeren RDF üçlüleri değil, aynı zamanda, dar-kavram geniş-kavram ve eş-anlamlılık ilişkileri de aktarılabilmektedir. Burada, Bilgi Çıkarım Matrisi bir excel dosyası olarak dosya sisteminde saklanmaktadır. Bir sonraki aşamada Anlamsal Web dönüştürücüsü Modülü, Bilgi Çıkarım Matrisinde

yer alan tüm verileri, RDF-üçlülerine dönüştürmekte ve OKM Referans Ontolojisini yeni oluşan ontoloji örnekleri (instances) ile birlikte genişletmektedir. Bu modül bir sonraki aşamada, oluşan, genişlemiş ontoloji üzerinde çıkarsama yaparak, gizli RDF üçlülerinin de oluşmasını sağlamaktadır. Bu işlemi, Şekil 5.1' de yer alan Inference-RDF alt-modülü gerçekleştirmektedir. Oluşan tüm RDF-üçlüleri Virtuoso [60] RDF saklayıcısında depolanmaktadır. Görselleştirme ve Kullanıcı Arayüzü Modülü ise, RDF saklayıcısında yer alan RDF üçlülerinin görselleştirilmesini ve bir kullanıcı arayüzü üzerinden incelenebilmesini sağlamaktadır.



Şekil 5.1 Önerilen sistemin mimarisi

5.2 Doğal Dil İşleme (NLP) Modülü

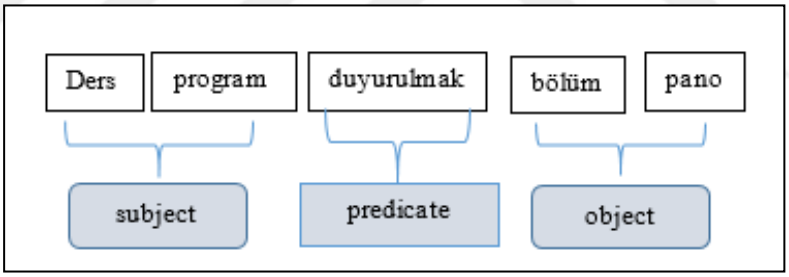
NLP Modülü cümlelerin özne-yüklem-nesne şeklinde ayrılması ile sorumlu modüldür. Türkçe için cümleleri otomatik öğelerine ayıran bir kütüphane geliştirilmiştir. Geliştirdiğimiz kütüphane Zemberek NLP kütüphanesini kullanmaktadır. Geliştirdiğimiz modül, cümleleri öğelerine ayırmak ve cümlelerdeki her bir nesnenin ontolojide hangi 5W1H sorusunun cevabı olduğunu belirlenmesi sağlamaktadır. Zemberek kütüphanesi kelime bazlı eklerine ayrıştırma yapabilmektedir. Geliştirdiğimiz kütüphanede, kelime ve kelime grupları Bilgi Çıkarım Matrisinde ontoloji sınıflarına atanırken, bu eklerden yararlanılmıştır. Örneğin, "Staj formları bölüme onaylatılır." cümlesindeki "bölüm" kelimesi "ISIM_YONELME_E " ekini aldığından "Kime-Nereye-Neye" sorularının cevabı olarak değerlendirilmektedir. Dolayısıyla bu sınıfların örnekleri olacak şekilde atanmaktadır. Bir nesnenin veya öznenin başka cümlede bir nesne veya özneyle

ilişkisini gösterebilmek için kelimelerin kökleri kullanılmaktadır. Çizelge 5.1, aşağıda verdiğimiz örnek cümlenin analizinin sonucunu göstermektedir. Örnek cümlenin öğelerine ayrılmış hali ise Şekil 5.2' de verilmiştir.

Örnek cümle: Ders programları bölüm panolarında duyurulur. (Course programmes are announced at department boards.)

Çizelge 5.1 Örnek cümlenin Zemberek çözümlemeleri

Kelime	**Kök**	**Tip**	**Ekler**
ders	ders	ISIM	ISIM
programları	program	ISIM	ISIM_KOK + ISIM_COGUL_LER + ISIM_BELIRTME_I
bölüm	bölüm	ISIM	ISIM
panolarında	pano	ISIM	ISIM_KOK + ISIM_COGUL_LER + ISIM_TAMLAMA_IN + ISIM_KALMA_DE
duyurulur	duyur	FIIL	FIIL_KOK + FIIL_EDILGEN_IL + FIIL_GENISZAMAN_IR



Şekil 5.2 Zemberek kütüphanesi kullanılarak RDF üçlüsü oluşturulması

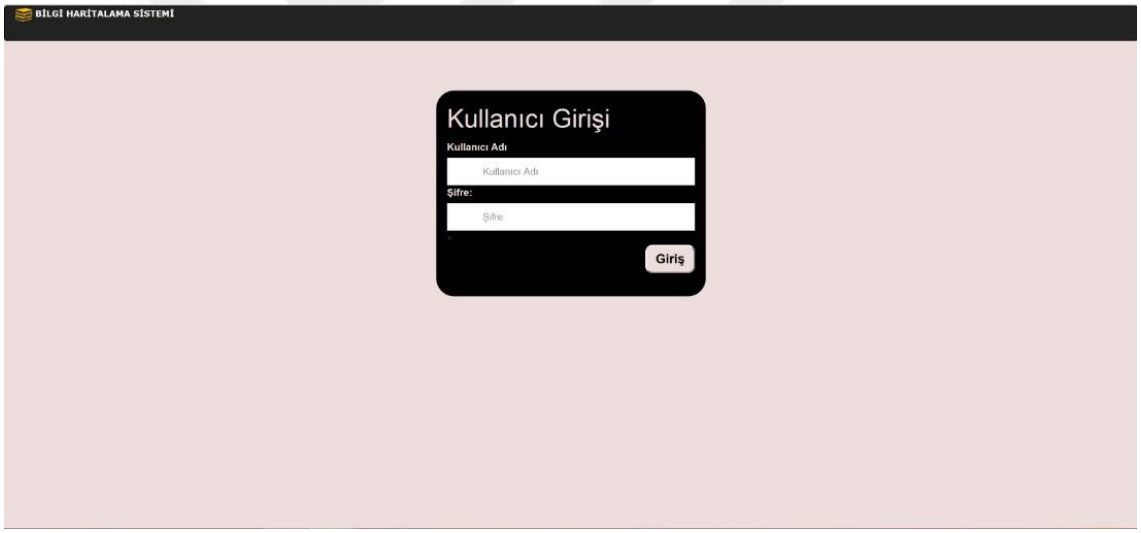
5.3 Anlamsal Web Dönüştürücü Modülü

Bu modül, Bilgi Çıkarım Matrisinde yer alan ayrıştırılmış bilgilerin, RDF üçlülerine dönüştürülmesinden sorumludur. Doküman kümesinden elde edilen RDF üçlüleri, OKM Referans Ontolojisine eklenerek, genişletilmiş bir ontoloji oluşturulmaktadır. Oluşturulan ontoloji, üzerinde bir Çıkarsama Motoru çalıştırılarak, örtük/gizli RDF üçlülerinin elde edilmesi sağlanmakadır. Tüm Ontoloji RDF üçlüleri bir RDF saklayıcısı üzerinde depolanmaktadır. RDF Saklayıcısı alt-modülü kullanılarak, SPARQL sorguları ile ontoloji verileri üzerinde sorgulamalar yapılabilmektedir.

5.4 Görselleştirme ve Kullanıcı Arayüzü Modülü

Bilginin, kullanıcıya görsel bir şekilde sunulması, bilginin ve sahip olduğu ilişkilerin anlaşılmasını kolaylaştırır. Bu amaç kapsamında, bilgiyi haritalamak için D3.js (Data Driven Document) adlı javascript kütüphanesi [61] kullanılmıştır. D3.js, veriye dayalı belgeleri işleyen bir JavaScript kütüphanesidir. Bilginin etkileşimi ve animasyon gibi dinamik davranışları desteklemenin yanında büyük veri setlerini kullanmaya da izin vermektedir. Bilgi haritalamanın gerçekleştirilmesi için elde edilen sorgu sonuçları JSON formatına dönüştürülerek D3.js'nin kullanabileceği veri dokümanı haline getirilmektedir.

Şekil 5.3'te görüldüğü gibi Ontoloji Tabanlı Bilgi Haritalama sistemi (OKM Sistemi) için bir kullanıcı arayüzü tasarlanmıştır. Arayüzün basit ve kullanıcı dostu olmasına önem verilmiştir.



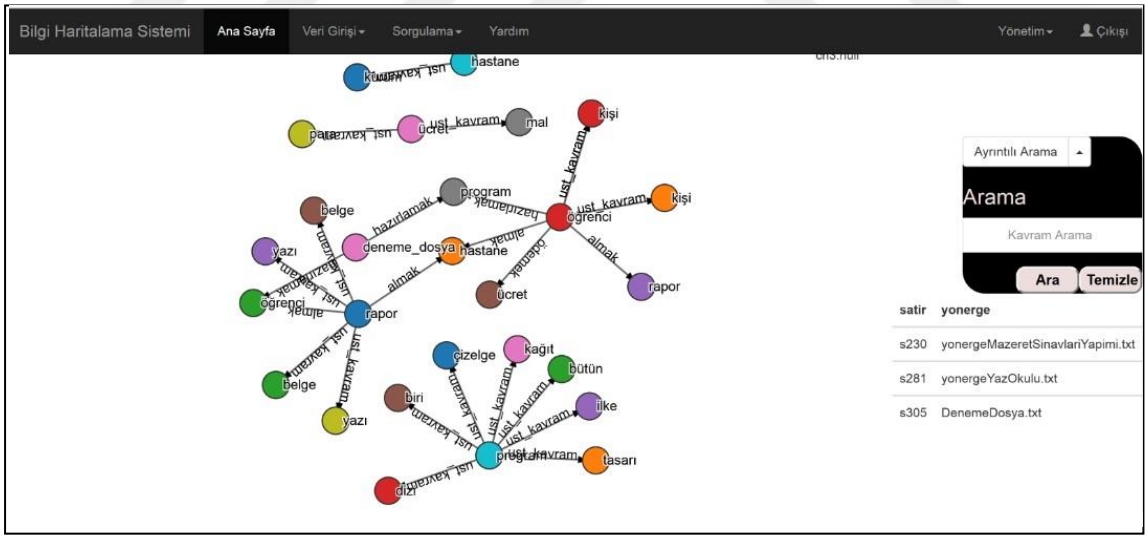
Şekil 5.3 Sistem kullanıcı arayüzü

Kullanıcı sorgusunu girdiğinde, ontoloji verileri üzerinde SPARQL sorguları işlenmektedir. Sistem, iki şekilde sorgulama yapılabilmesine olanak sağlayacak şekilde geliştirilmiştir; SPARQL sorgulama ve kelime tabanlı sorgulama. Kelime tabanlı sorgulamada, girilen kelime ile adlandırılmış kaynakların bulunması sağlanmaktadır.

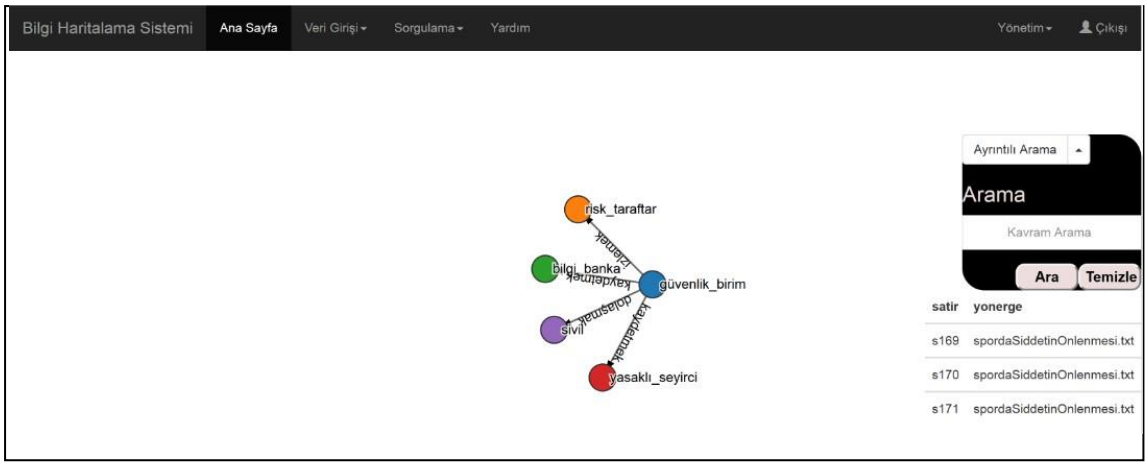
Belirli bir kelimeye yönelik sorgulama yapmak isteyen kullanıcı grafiksel arama Sorgulama menüsünden Grafik Kavram Arama linkini kullanır. Sorgulama sonucunda kullanıcı ilgili kavramın, hangi dokümanda yer aldığını görebilmekte, kavramdaki örtük

bilgileri de görmek için ise çıkarsama modeli üzerinden sorgulama yapabilmektedir. Sorgu sonuçları grafiksel olarak görselleştirilmekte ve bulunduğu dosyalar listelenmektedir.

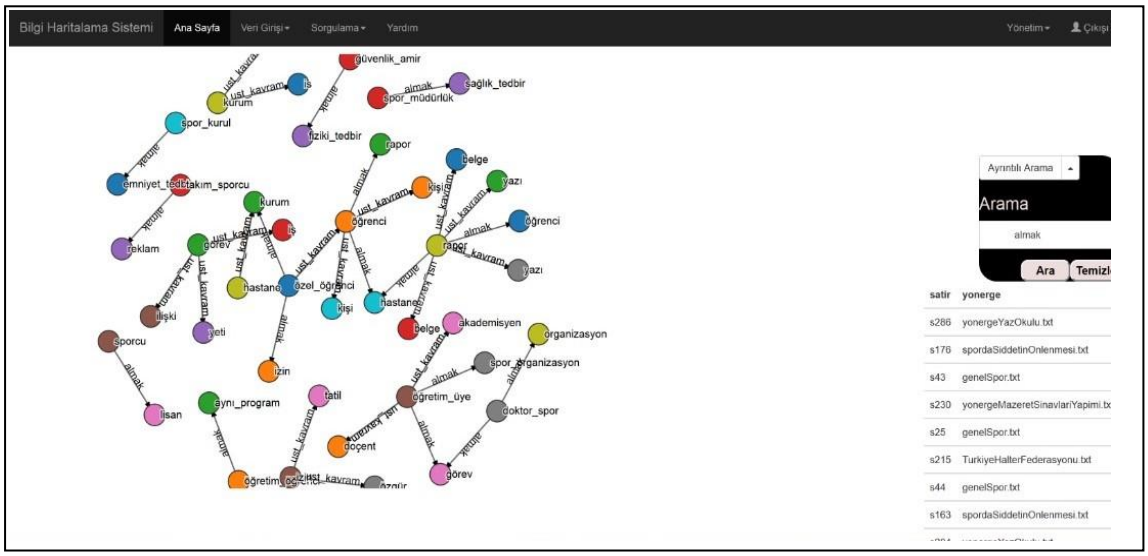
Cümlenin akışını belirtmek amacıyla haritalamada oklar özneden nesneye doğru olacak şekilde yönlendirilmiştir. Şekil 5.4' te "öğrenci " kavramının haritalanmasını görebiliriz. Burada "öğrenci" adlı düğüm özneyi ifade etmektedir. "Hazırlamak, ödemek, doldurmak ve sormak" adlı oklar (ilişkiler) yüklemi ifade etmektedir. "Staj defter, ders ücret, taahhütname form, mazeret sınav, staj form, ve sağlık beyan" adlı düğümler nesneleri ifade etmektedir. Bu grafiksel kavram arama için ayrıntılı arama ile "öğrenci" kelimesinin hangi dokümanlar ve excel dosyasının hangi satırında bulunduğu bilgileri sağ altta listelenmiştir. Ayrıca "öğrenci" kavramının ve bu kavramın ilişkide bulunduğu kavramların varsa alt-üst kavramları da grafiksel görselleştirmede gösterilmiştir. Şekil 5.5' te "güvenlik birim" kavramının haritalanmasını görebiliriz. Burada RDF üçlülerinden ne öznenin ne de nesnelerin alt-üst kavramlara sahip olmadığını görebilmekteyiz. Şekil 5.6'da "almak" kavramının görselleştirilmesi verilmektedir.



Şekil 5.4 "öğrenci" kavramının görselleştirilmesi



Şekil 5.5 "güvenlik birim" kavramının görselleştirilmesi



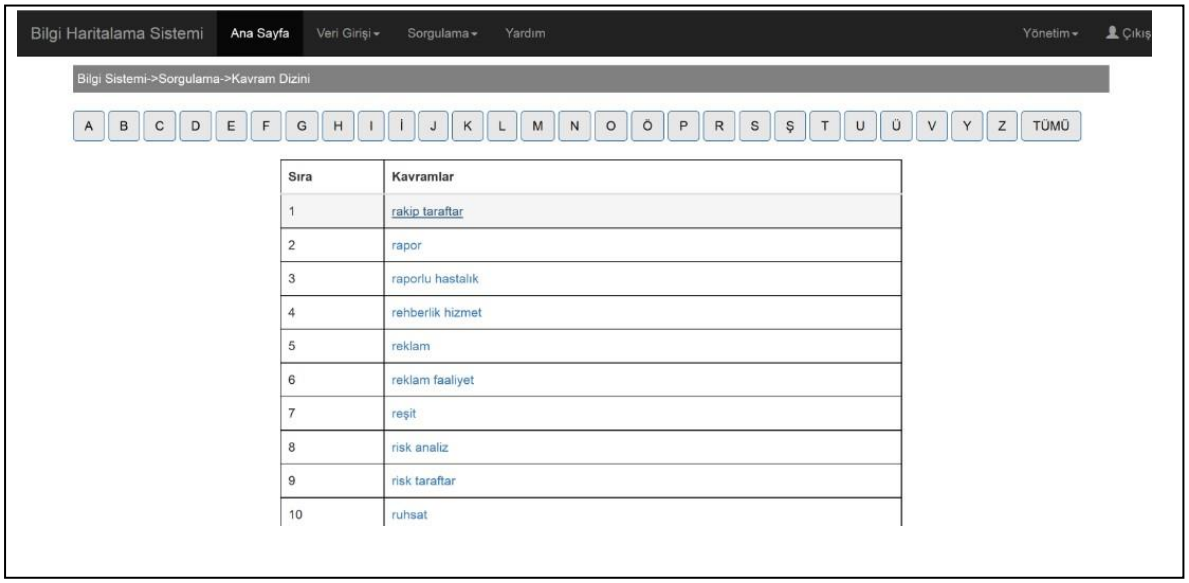
Şekil 5.6 "almak" kavramının görselleştirilmesi

SPARQL sorgu dilini bilen bir kullanıcı ilgili alana sorgusunu girerek aradığı bilgiye ulaşabilir. SPARQL bilgisi az olan bir kullanıcı “SPARQL Sorgulama” kısmındaki sağ tarafta yer alan örnek SPARQL sorgularını seçerek, isterse üzerinde değişikliğe giderek sorgusunu gerçekleştirebilir. Sorgu sonuçları tabloda listelenir. Şekil 5.7’ de SPARQL sorgulama ile ontolojide yer alan filler gösterilmektedir.



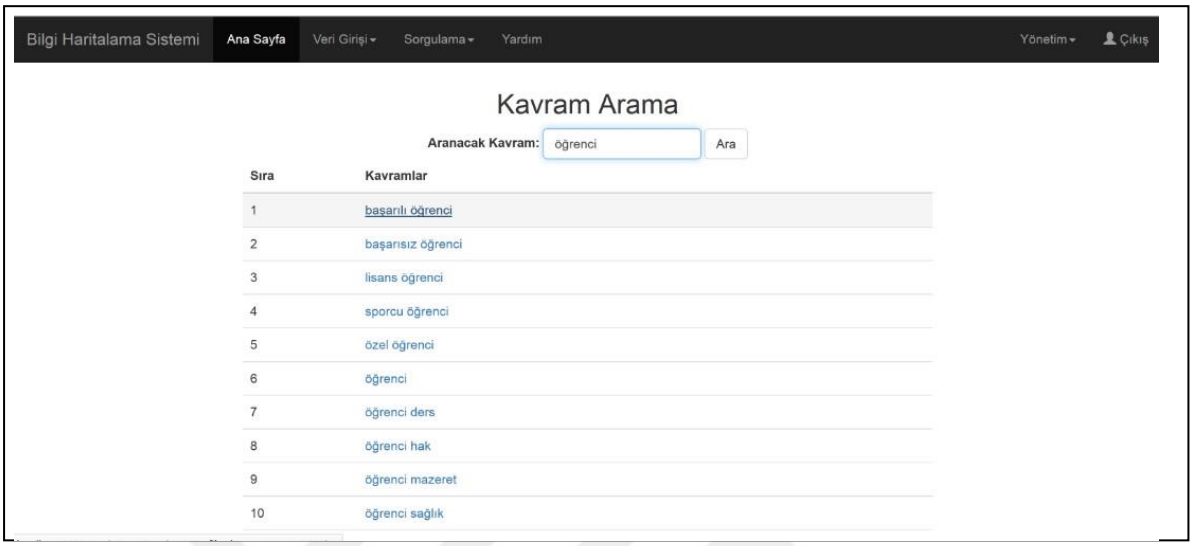
Şekil 5.7 SPARQL sorgulama

Sorgulama menüsündeki Kavram Dizini kısmından ontolojide yer alan kavramlar alfabetik olarak listelenebilmektedir (Şekil 5.8). İstenilen kavramın üzerine tıklandığında otomatik olarak Grafiksel Kavram Arama sayfasında görselleştirme gerçekleşmektedir.



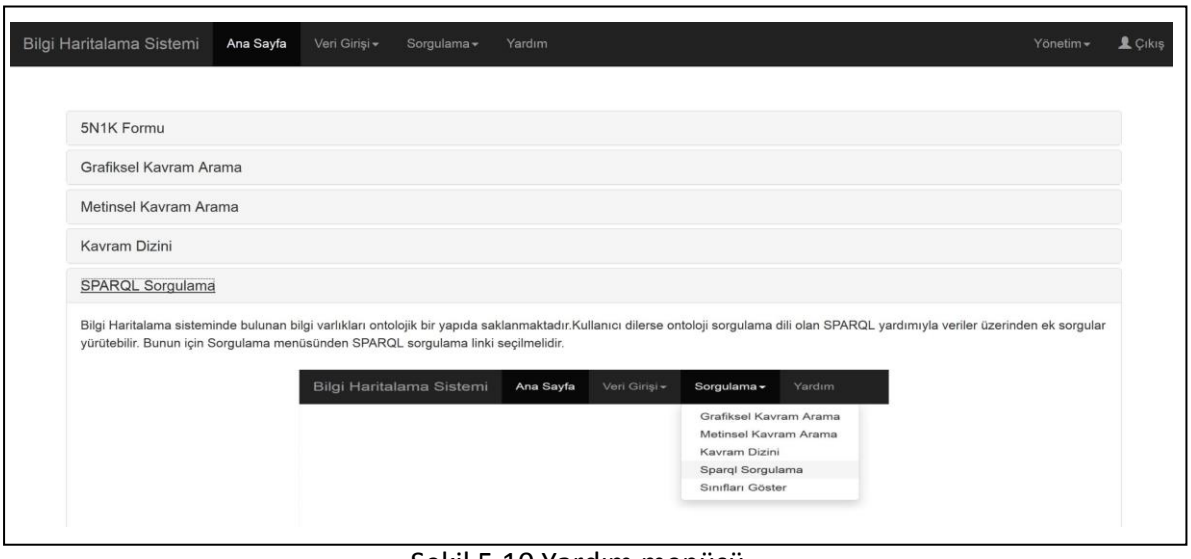
Şekil 5.8 Kavram dizini

Sorgulama menüsündeki Metinsel Kavram Arama kısmında kavram arama işlemi gerçekleştirilebilmektedir (Şekil 5.9). Aranan kelimeyi içinde barındırak kavramlar tabloda listelenmekte ve her hangi bir kavramın üzerine tıklandığında otomatik olarak Grafiksel Kavram Arama sayfasında görselleştirme gerçekleşmektedir.



Şekil 5.9 Metinsel kavram arama

Sistemin işleyişini anlamak isteyen kullanıcılar için Yardım bölümü hazırlanmıştır. Kullanıcılar burada hangi kısmın nasıl kullanıldığı, nasıl sonuçlar döndürdüğü bilgisine ulaşabilmektedir.



Şekil 5.10 Yardım menüsü

BÖLÜM 6

YÖNTEMİN DEĞERLENDİRİLMESİ

Geliştirilen ontolojinin amaca ne ölçüde hizmet ettiğinin denetlenmesi gerekmektedir. Bunu gerçekleştirmek için ontoloji ve bilgi haritalama bağlamında yeterlilik soruları hazırlanarak bir anket çalışması yapılmıştır. Anket çalışması ile bilginin anlaşılmasında sistemin kolaylık sağlamadaki yeterliliği, farklı sorgulama metotlarının bilgi edinmede fark yaratıp yaratmadığı, ontoloji tabanlı bilgi haritalamanın doğru ilişkiler getirip getirmediği ve sistemin tutarsızlık içerip içermediği öğrenilmek istenmektedir.

6.1 Kullanıcı Değerlendirme Çalışması

Sistemin bilgi ihtiyacına cevap verme yeterliliğinin ölçülmesi için bir kullanıcı değerlendirme çalışması gerçekleştirilmiştir. 18 sorudan oluşan anketimizde Çizelge 6.1' de görülen beşli Likert ölçeği kullanılmaktadır.

Çizelge 6.1 Anket için kullanılan Likert ölçeği

Cevaplar	Puanlar
Kesinlikle katılmıyorum	1
Katılmıyorum	2
Kararsızım	3
Katılıyorum	4
Kesinlikle katılıyorum	5

Son kullanıcılardan sistemimizi denemeleri ve sistem deneyimleri sonrası ankete katılımları için gönüllü olmaları istenmiştir. Anket çalışmasında, Yıldız Teknik Üniversitesi, Bilgisayar Mühendisliği Bölümü personeli arasından, 10 katılımcı ile anket gerçeklenmiştir ve toplamda 50 sorgu yapılmıştır. Her katılımcı, bölüm başkan, öğrenci, staj, güvenlik birim, hemşirelik, Sağlık bakanlık vb. gibi ifadeleri, hazırlanan sistem üzerinde kişi başı 5 adet sorgu olacak şekilde sorgulatmışlardır. Kullanıcılardan en az birer tane SPARQL ve kelime tabanlı sorgu olmak üzere toplamda 5 sorgu sormaları istenmiştir ve her sorgu için bir anket doldurmaları talep edilmiştir.

Katılımcıların 3'ü yönetmelik ve yönergeleri iş hayatında kullanan kişiler ve geriye kalan 7 kişide ontoloji ve bilgi haritalama hakkında bilgi sahibi kişilerdir. Öncelikle araştırmanın amacı katılımcılara açıklanmış ve sonra aşinalık kazanmaları için yönetmelikler, yönergeler ve kavram listesi katılımcılara verilmiştir. Katılımcılara sistemi dilediği 5 sorgu için kullanmaları ve farklı parametrelere sahip her bir sorgu için ayrı ayrı anket doldurmaları istenmiştir. Maksadımız, sistemi değerlendirirken katılımcıları sistem hakkında daha fazla düşünmeye sevk etmektir.

6.2 Değerlendirme

Anket, sistem yeterliliğini araştıran sorular (1, 3, 4, 5, 18), sistem sorgulama yeteneğini araştıran sorular (2, 8, 10), bilgi haritalamada ilişkilerin gösterimini araştıran sorular (6, 7, 9, 11, 12, 13) ve sistem tutarsızlığını araştıran sorular (14, 15, 16, 17) olmak üzere 4 gruba ayrılmaktadır. Bu yöndeki değerlendirme sonuçları, 50 sorgunun ortalaması alınarak hesaplanmıştır. Anket sonuçları Çizelge 6.2' de görülmektedir. Anket puanlarına göre, geliştirilen sistemin kullanıcıların bilgi ihtiyaçlarını karşıladığı söylenebilmektedir.

Çizelge 6.2 Anket puanları ortalaması

Sistem Yeterliliği	Sorgulama	İlişki Gösterimi	Tutarsızlık
3.81	4.18	4.02	1.85

Liao ve Cheung çalışmalarında, kullanıcıların internet tabanlı çevrimiçi bankacılık üzerindeki tutumlarını ölçmek için deneysel bir çalışma gerçekleştirmişlerdir[62]. Yedili

Likert ölçeği kullandıkları anketlerinde bay ve bayan kullanıcılar arasında internet tabanlı çevrimiçi bankacılıktan beklentilerinde anlamlı bir farklılık olup olmadığını göstermek amacıyla t testi uygulamışlardır. T testi sonucu oluşan p anlamlılık değerine göre, cinsiyetler arasında internet tabanlı çevrimiçi bankacılık beklentilerinde önemli bir farklılık olmadığını ortaya çıkarmışlardır. Bizde bu çalışmadan yola çıkarak, kullanıcıları sahip oldukları geçmişlere göre gruplara ayırdığımızda; ontoloji-bilgi haritalama bilgilerine sahip katılımcılar Grup 1'de, dokümanları iş hayatında kullanan katılımcılar ise Grup 2'de yer almaktadır. Böylelikle anketlere verilen cevaplar gruplara göre anlamlı bir farklılık gösterip göstermediği, SPSS 22 uygulaması aracılığıyla, bağımsız örneklem üzerinde t testi uygulanarak analiz edilmiştir. t testi, hipotez testlerinde en yaygın olarak kullanılan testlerden biridir. t testi, iki grup arasındaki farkın rastlantısal mı yoksa istatistiksel mi olduğuna karar vermek için kullanılır. Bağımsız örneklem t testi aşağıdaki (6.1) gibi hesaplanmaktadır . $\overline{X_1}$ ve $\overline{X_2}$, Grup 1 ve Grup 2'nin ortalamalarını, SS_1 ve SS_2 Grup 1 ve Grup 2'nin standart sapmalarını, n_1 ve n_2 Grup 1 ve Grup 2'nin örnek sayısını göstermektedir.

$$t = \frac{\overline{X_1} - \overline{X_2}}{\sqrt{\frac{(n_1 - 1){SS_1}^2 + (n_2 - 1){SS_2}^2}{(n_1 + n_2 - 2)}} - \sqrt{\frac{1}{n_1} + \frac{1}{n_2}}} \qquad (6.1)$$

Bu istatistiksel analiz için H_0 (null hipotez), H_A (alternatif hipotez) ve p (anlamlılık değeri) aşağıdaki gibi belirlenmiştir.

H_0= Grup 1 ve Grup 2 kişilerinin anket cevapları arasında anlamlı bir fark yoktur.

H_A= Grup 1 ve Grup 2 kişilerinin anket cevapları arasında anlamlı bir fark vardır.

p=0.05.

Çizelge 6.3 Gruplara göre eğitim veri seti için istatistiksel analiz sonuçları

Grup 1		Grup 2		p değeri
Ortalama	**Standart Sapma**	**Ortalama**	**Standart Sapma**	
3.57	0.61	3.53	0.14	0.911

Çizelge 6.4 Gruplara göre spor veri seti için istatistiksel analiz sonuçları

Grup 1		Grup 2		p değeri
Ortalama	**Standart Sapma**	**Ortalama**	**Standart Sapma**	
3.58	0.35	3.62	0.33	0.886

Çizelge 6.5 Gruplara göre sağlık veri seti için istatistiksel analiz sonuçları

Grup 1		Grup 2		p değeri
Ortalama	**Standart Sapma**	**Ortalama**	**Standart Sapma**	
3.67	0.29	3.77	0.44	0.711

Çizelge 6.3, Çizelge 6.4 ve Çizelge 6.5‘ te görüldüğü gibi üç alan için ayrı ayrı hesaplanan $p > 0.05$ olması gruplar arasında anket sonuçlarında anlamlı farklılıklar görülmediğini ifade etmektedir. Gruplar arasında anketlere verilen cevaplar benzer olduğu sonucuna ulaşılmaktadır. Sonuç olarak, Null hipotezi doğru olarak kabul edilmektedir.

BÖLÜM 7

SONUÇLARI ETKİLEYEN FAKTÖRLER

Bu bölümünde, çalışmanın performans değerlendirmesine yönelik geçerlilik tehditlerinden bahsedilmektedir. Bunlar; kullanılan NLP kütüphanesinden kaynaklanan tehdit, kullanılan veri setinden kaynaklanan tehdit ve kullanıcı anketlerinden kaynaklanan tehditlerdir.

Çalışmada NLP kütüphanesi olarak Zemberek'ten yararlanılmaktadır. Bilindiği üzere Türkçe'de bir kelimenin birçok morfolojik çözümlemeleri vardır. Kullandığımız NLP kütüphanesi Zemberek bu çözümlemelerin tümünü kullanıcıya vermektedir. Bu çalışma kapsamında bu çözümlemelerden birincisi kullanıldığından bazı kelimeler için doğru çözümlemeler kullanılamamıştır. Bu sebeple öğelerine ayrıştırma işleminde %100 doğruluk oranı vardır denilememektedir.

Bu çalışma ile birbirinden farklı alanlara ait küçük, orta ve büyük ölçekteki veri setleri üzerinde, basit kurallarla öğelerine ayrıştırma işlemi yapılmaya çalışılmıştır. Bu kurallar her veri setinde aynı doğrulukta çalışacaktır denilememektedir. Farklı türde bir veri seti kullanıldığında farklılık gösterebilir. Sistemin geliştirilme amacı her türlü cümleyi öğelerine ayrıştırmaktan ziyade, alandan bağımsız ontolojimiz için veri sağlamaktır.

Çalışmanın değerlendirilmesi için kullanıcı anketlerine başvurulmuştur. Kullanıcı anket analiz sonuçları değerlendirme bölümünde görülmektedir. Kullanıcıların geçmişleri, profilleri ve sayıları değişim gösterdikçe, başka deneklerle anket çalışması yapıldıkça sonuçlarda buna bağlı olarak farklılık gösterebileceği göz önünde tutulmalıdır. Ayrıca,

sistemin değerlendirilmesinde kullanıcıların sorgularına ve sayısına göre de anket cevapları çeşitlilik gösterebilmektedir.

BÖLÜM 8

SONUÇLAR VE GELECEKTE PLANLANAN ÇALIŞMALAR

Bu çalışmada makinelerin de anlayabileceği bir bilgi ortamı tanımlayabilmek ve kullanıcıların bilgiye erişimini kolaylaştırmak için, görselliğin ön planda tutulduğu bir bilgi çıkarım sistemi gerçeklenmiştir. Böylece kullanıcılar aradıkları kavramların hangi cümlede nasıl kullanıldığını, hangi dokümanda yer aldığını görebilmektedir. Ayrıca sistem, cümlelerde beraber kullanılabilen özne-yüklem, nesne-yüklem ikililerini de ortaya çıkarmaya yardımcı olmaktadır.

Gelecek çalışmada, Türkçe tümceleri otomatik öğelerine ayıran bir yazılımın gelişimiyle, bir büyük veri işleme platformu aracılığıyla büyük doküman setleri üzerinde ontoloji çıkarımı hedeflenmektedir.

KAYNAKLAR

[1] Berners-Lee, T., Hendler, J. ve Lassila, O., (2001). “The Semantic Web”, Scientific American, 284 (5):34-43.

[2] The friend of a friend (FOAF) Project, http://www.foaf-project.org/, 15 Nisan 2016.

[3] SKOS, http://www.w3.org/2004/02/skos/, 15 Nisan 2016.

[4] Musa, M. A., Othman, M. S., ve AL-Rahimi, W. M., (2013). “Ontology Driven Knowledge Map for Enhancing Business Process Reengineering”, Computer Science & Engineering: An International Journal (CSEIJ), 3(6):11-19.

[5] Kawtrakul, A., Tooumnauy, T., Khunthong, V., Thunkijjanukij, A., Lohapiyaphan, K., Pothipaki, Y., Narkwiboonwong, W. ve Pusittigul, A., (2008). “Ontology based Knowledge Map Construction for a Smart Knowledge Service”, IAALD AFITA WCCA 2008, 24 - 27 August, Tokyo, Japan.

[6] Kim, H. H., Rieh, S. Y., Ahn, T. K. ve Chang, W. K., (2004). "Implementing an ontology-based knowledge management system in the Korean financial firm environment", Proc. 67th Annu. Meeting Amer. Soc. Inf. Sci. Technol., 300-309.

[7] Uschold, M. ve Gruninger, M., (1996). “Ontologies: Principles, Methods and Application”, Knowledge Engineering Review, 11(2): 93-136.

[8] Barkschat, K., (2014). “Semantic Information Extraction on Domain Specific Data Sheets”, The Semantic Web: Trends and Challenges, Springer: 864-873.

[9] Rusu, D., Dali, L., Fortuna, B., Grobelnik, M. ve Mladenic, D., (2007). “Triple extraction from sentences”, Proceeding of the 10th International Multiconference Information Society –IS 2007 A, 218-222.

[10] Medeni, İ. T. (2014). Maintenance for Sustainable Ontologies, Doktora Tezi, ODTÜ Bilişim Sistemleri Bölümü.

[11] Bansal, A., Kona, S., Blake, M. B., ve Gupta, G., (2008). “An agent-based approach for composition of semantic web services”, In 2008 IEEE 17th Workshop on Enabling Technologies: Infrastructure for Collaborative Enterprises, IEEE, 12–17.

[12] Mohebbi, K., Ibrahim, S. ve Idris, N. B., (2012). “Contemporary semantic web service frameworks: An overview and comparisons”, International Journal on Web Service Computing, 3 (3):65–76

[13] Antoniou, G. ve Harmelen, F. V., (2008). A Semantic Web Primer, MIT Press.

[14] Frauenfelder, M., (2004). “Sir Tim Beners-Lee. Technology Review”, 107(8):40.

[15] Maedche, A. ve Staab, S., (2000). “Semi-Automatic engineering of ontologies from text”, In Proceedings of the 12th Internal Conference on Software and Knowledge Engineering, Chicago, USA, July, 5-7, 2000, 1756:231–239.

[16] Liyang, Y., (2007). Introduction to the Semantic Web and Semantic Web Services, Chapmanton & Hall/CRC: Boca Raton, FL USA.

[17] Kruk, S. R., Woroniecki, T., Gzella, A. Ve Dabrowski, M., (2007). “JeromeDL – A Semantic Digital Library”. In: GolbeckJ, MikaP, editors. Semantic Web challenge. Busan, South Korea: CEUR, 295:139–150.

[18] Shahzad, M., (2013). A semantic web approach for dealing with university courses, MS thesis, University of Adgar.

[19] Gangemi, A. Ve Presutti, V., (2006). "The bourne identity of a web resource" , In Proceedings of Identity Reference and the Web Workshop (IRW), Laboratory for Applied Ontology, Roma, Italy.

[20] Sauermann, L., Cyganiak, R. ve Völkel, M., (2008). “Cool URIs for the semantic web working draft W3C”, 1–15.

[21] Kanne, C. C. ve Moerkotte, G., (1999). Efficient relational storage and retrieval of XML documents, Technical Report 8/99, University of Mannheim.

[22] W3 Schools Web Sitesi, www.w3schools.com/xsl/books.xml, 15 Nisan 2015.

[23] Schenkel, R. ve Theobald, M., (2006). “Structural Feedback for Keyword-Based XML Retrieval”, ECIR 2006, Springer, Heidelberg, 326–337.

[24] RDF About, http://www.rdfabout.com/intro/, 15 Nisan 2016.

[25] Jentzsch, A., Andersson, B., Hassanzadeh, O., Stephens, S. ve Bizer, C., (2009). “Enabling Tailored Therapeuticd with Linked Data”, Proceedings Linked Data on the Web” (LDOW2009), April 20th, 2009, Madrid, Spain.

[26] Allemang, D. ve Hendle, J., (2008). “Semantic Web for the Working Ontologist: Effective Modeling in RDFS and OWL”, Morgan Kaufmann Publishers Inc., San Francisco, CA.

[27] Christiaan, M. ve Klein, A., (2004). Change Management for Distributed Ontologies, Ph.D. dissertation, Vrije Universiteit.

[28] OWL, https://tr.wikipedia.org/wiki/OWL, 15 Nisan 2016.

[29] Gruber, T. R., (1995). “Toward principles for design of ontologies used for knowledge sharing”, International Journal of Human and Computer Studies, 43(5/6): 907-928.

[30] Chen, H., Finin, T. ve Joshi, A., (2003). “An ontology for context-aware pervasive computing environments”, The Knowledge Engineering Review, 18 (3):197–207.

[31] Wang, Y., Gong, J., ve Wu, X., (2007). “Geospatial semantic interoperability based on Ontology”, Geo-Spatial Information Science, 10 (3):204–207.

[32] Noy, N. F. ve McGuinness, D. L., (2001). Ontology Development 101: A Guide to Creating Your First Ontology, Technical Report SMI-2001-0880, Standford Medical Informatics.

[33] Fensel, D., (2002). “Ontology-based knowledge management”, Computer 35, (11):55-59.

[34] Gruber, T. R., (1991). Ontolingua: A Mechanism to Support Portable Ontologies, Technical Report KSL 91-66, Standford University Knowledge Systems Laboratory.

[35] RDF, http://www.w3.org/RDF/, 15 Nisan 2016.

[36] OWL Web Ontology Language Guide, http://www.w3.org/TR/204/REC-owl-guide, 15 Nisan 2016.

[37] Query Language for RDF W3C Recommendation, https://www.w3.org/TR/rdf-sparql-query, 15 Nisan 2016.

[38] Guarino, N., (1998). “Formal ontology and information systems”, Proceedings of the First International Conference, 46 (6):3–15.

[39] Miller, G. A., (1995). “WordNet: A lexical database for English”, Communications of the ACM, 38 (11): 39–41.

[40] Swartout, B., Patil, R., Knight, K. ve Russ, T., (1996). “Toward distributed use of large-scale Ontologies”, Proc. of the Tenth Workshop on Knowledge Acquisition for Knowledge-Based Systems, 138–148.

[41] Lenat, D., (1995). “CYC: A Large-Scale Investment in knowledge infrastructure”, Communications of the ACM, 38 (11): 33–38.

[42] Pease, A., Niles, I. ve Li, J., (2002). “The suggested upper merged ontology: A large ontology for the semantic web and its applications”, Working Notes of the AAAI-2002 Workshop on Ontologies and the Semantic Web, July 28-August 1, Edmonton, Canada .

[43] Giunchiglia, F., Marchese, M., Zaihrayeu, I., (2006). “Encoding classifications into lightweight ontologies”. In: Sure, Y., Domingue, J. (eds.) ESWC 2006,LNCS, Springer, Heidelberg, 4011:80–94.

[44] Winer, D., (2014). “Review of ontology based storytelling devices”. In: Dershowitz, N., Nissan, E. (eds.) Choueka Festschrift, Part II. LNCS, Springer, Heidelberg, 8002:394–405.

[45] Juhnyoung, L., (2004). “Frequently Asked Questions on Ontology Technology”, IBM T.J. Watson Research Center, 19 August, Hawthorne, NY.

[46] Gašević, D., Djurić, D. ve Devedžić, V., (2009). Model Driven Engineering and Ontology Development (2nd ed.), Springer, Berlin.

[47] Jena Ontology API, http://jena.apache.org/documentation/ontology/, 10 Nisan 2016.

[48] Sesame, https://www.w3.org/2001/sw/wiki/Sesame, 15 Nisan 2016.

[49] Segaran, T., Evans, C. ve Taylor, J., (2009). Programming the Semantic Web, Sebastopol, O'Reilly Media.

[50] Van Assem, M., Malais´e, V., Miles, A. ve Schreiber, G. (2006). "A method to convert thesauri to SKOS", In: Sure, Y., Domingue, J. (eds.) ESWC 2006. LNCS, Springer, Heidelberg, 4011:95–109.

[51] Eppler, M. J., (2008). "A process-based classification of knowledge maps and application examples", Knowledge and Process Management, 15 (1):59–71.

[52] Vail,E. F., (1999). "Knowledge Mapping: Getting Started with Knowledge Management", Information Systems Management, 16(4):16-23.

[53] Güngör, O., Güngör, T., (2007). "Türkçe Bir Sözlükteki Tanımlardan Kavramlar Arasındaki Üst-kavram İlişkilerinin Çıkarılması", Akademik Bilişim Konferansı, 1(1):1–13.

[54] TDK Sözlüğü, http://www.tdk.gov.tr, 10 Ekim 2015.

[55] Akın, A. A. ve Akın, M. D., (2007). "Zemberek, an open source nlp framework for Turkish languages", Structure, 10:1-5.

[56] Zemberek kaynak kod, http://code.google.com/p/zemberek, 15 Ekim 2015.

[57] Apache Jena, http://jena.apache.org/, 15 Ekim 2015.

[58] Yıldız Teknik Üniversitesi Öğrenci İşleri Yönetmelikler ve Yönergeler, http://www.ogi.yildiz.edu.tr/ogi/6/Yönetmelikler-ve-Yönergeler/36, 25 Ekim 2015.

[59] Resmi Gazete, http://www.resmigazete.gov.tr/, 30 Ekim 2015.

[60] Erling, O. ve Mikhailov, I. (2009). "RDF Support in the Virtuoso DBMS", Studies in Computional Intelligence, 221:7-24.

[61] D3js, https://d3js.org/, 10 Ocak 2016.

[62] Liao, Z. ve Cheung, M., T., (2002). "Internet-based e-banking and consumer attitudes: an empirical study", Information and Management, 39 (4):283–295.

EK-A

5N1K REFERANS ONTOLOJİ

```
<rdf:RDF
   xmlns:rdf="http://www.w3.org/1999/02/22-rdf-syntax-ns#"
   xmlns:owl="http://www.w3.org/2002/07/owl#"
   xmlns:skos="http://www.w3.org/2004/02/skos/core#"
   xmlns:rdfs="http://www.w3.org/2000/01/rdf-schema#"
   xmlns:okm="http://www.semanticweb.org/OKM_Ontology#"
   xmlns:foaf="http://xmlns.com/foaf/0.1/"
   xmlns:xsd="http://www.w3.org/2001/XMLSchema#" >
  <rdf:Description rdf:about="http://www.semanticweb.org/OKM_Ontology#Kime">
   <rdfs:subClassOf
rdf:resource="http://www.semanticweb.org/OKM_Ontology#Kim"/>
   <rdf:type rdf:resource="http://www.w3.org/2002/07/owl#Class"/>
  </rdf:Description>
  <rdf:Description rdf:about="http://www.semanticweb.org/OKM_Ontology#Kimi">
   <rdfs:subClassOf
rdf:resource="http://www.semanticweb.org/OKM_Ontology#Kim"/>
   <rdf:type rdf:resource="http://www.w3.org/2002/07/owl#Class"/>
  </rdf:Description>
  <rdf:Description rdf:about="http://xmlns.com/foaf/0.1/Document">
   <rdfs:subClassOf
rdf:resource="http://www.semanticweb.org/OKM_Ontology#Thing"/>
   <rdf:type rdf:resource="http://www.w3.org/2002/07/owl#Class"/>
  </rdf:Description>
```

<rdf:Description rdf:about="http://www.semanticweb.org/OKM_Ontology#HangiSiklikla">
<rdfs:subClassOf rdf:resource="http://www.semanticweb.org/OKM_Ontology#NeZaman"/>
<rdf:type rdf:resource="http://www.w3.org/2002/07/owl#Class"/>
</rdf:Description>
<rdf:Description rdf:about="http://www.semanticweb.org/OKM_Ontology#Neden">
<rdfs:subClassOf rdf:resource="http://www.semanticweb.org/OKM_Ontology#BesNbirK"/>
<rdf:type rdf:resource="http://www.w3.org/2002/07/owl#Class"/>
</rdf:Description>
<rdf:Description rdf:about="http://www.semanticweb.org/OKM_Ontology#HangiDeneyimle">
<rdfs:subClassOf rdf:resource="http://www.semanticweb.org/OKM_Ontology#Nasil"/>
<rdf:type rdf:resource="http://www.w3.org/2002/07/owl#Class"/>
</rdf:Description>
<rdf:Description rdf:about="http://www.semanticweb.org/OKM_Ontology#Nede">
<rdfs:subClassOf rdf:resource="http://www.semanticweb.org/OKM_Ontology#Ne"/>
<rdf:type rdf:resource="http://www.w3.org/2002/07/owl#Class"/>
</rdf:Description>
<rdf:Description rdf:about="http://www.semanticweb.org/OKM_Ontology#NeyeDayanarak">
<rdfs:subClassOf rdf:resource="http://www.semanticweb.org/OKM_Ontology#Neden"/>
<rdf:type rdf:resource="http://www.w3.org/2002/07/owl#Class"/>
</rdf:Description>
<rdf:Description rdf:about="http://www.semanticweb.org/OKM_Ontology#Neyden">
<rdfs:subClassOf rdf:resource="http://www.semanticweb.org/OKM_Ontology#Ne"/>
<rdf:type rdf:resource="http://www.w3.org/2002/07/owl#Class"/>
</rdf:Description>
<rdf:Description rdf:about="http://www.semanticweb.org/OKM_Ontology#Kim">
<rdfs:subClassOf rdf:resource="http://www.semanticweb.org/OKM_Ontology#BesNbirK"/>

<rdf:type rdf:resource="http://www.w3.org/2002/07/owl#Class"/>

</rdf:Description>

<rdf:Description rdf:about="http://www.semanticweb.org/OKM_Ontology#Thing">

<rdf:type rdf:resource="http://www.w3.org/2002/07/owl#Class"/>

</rdf:Description>

<rdf:Description rdf:about="http://www.semanticweb.org/OKM_Ontology#in">

<rdfs:range
rdf:resource="http://www.semanticweb.org/OKM_Ontology#Statement"/>

<rdfs:domain
rdf:resource="http://www.semanticweb.org/OKM_Ontology#BesNbirK"/>

<rdf:type rdf:resource="http://www.w3.org/2002/07/owl#ObjectProperty"/>

</rdf:Description>

<rdf:Description rdf:about="http://www.semanticweb.org/OKM_Ontology#Kimden">

<rdfs:subClassOf
rdf:resource="http://www.semanticweb.org/OKM_Ontology#Kim"/>

<rdf:type rdf:resource="http://www.w3.org/2002/07/owl#Class"/>

</rdf:Description>

<rdf:Description
rdf:about="http://www.semanticweb.org/OKM_Ontology#KimTarafindan">

<rdfs:subClassOf
rdf:resource="http://www.semanticweb.org/OKM_Ontology#Kim"/>

<rdf:type rdf:resource="http://www.w3.org/2002/07/owl#Class"/>

</rdf:Description>

<rdf:Description rdf:about="http://www.semanticweb.org/OKM_Ontology#Neyi">

<rdfs:subClassOf
rdf:resource="http://www.semanticweb.org/OKM_Ontology#Ne"/>

<rdf:type rdf:resource="http://www.w3.org/2002/07/owl#Class"/>

</rdf:Description>

<rdf:Description rdf:about="http://www.semanticweb.org/OKM_Ontology#Nereyi">

<rdfs:subClassOf
rdf:resource="http://www.semanticweb.org/OKM_Ontology#Nerede"/>

<rdf:type rdf:resource="http://www.w3.org/2002/07/owl#Class"/>

</rdf:Description>

<rdf:Description
rdf:about="http://www.semanticweb.org/OKM_Ontology#HangiTeknikle">

<rdfs:subClassOf rdf:resource="http://www.semanticweb.org/OKM_Ontology#Nasil"/>

<rdf:type rdf:resource="http://www.w3.org/2002/07/owl#Class"/>

</rdf:Description>

<rdf:Description rdf:about="http://www.semanticweb.org/OKM_Ontology#HangiGoreviGeregi">

<rdfs:subClassOf rdf:resource="http://www.semanticweb.org/OKM_Ontology#Neden"/>

<rdf:type rdf:resource="http://www.w3.org/2002/07/owl#Class"/>

</rdf:Description>

<rdf:Description rdf:about="http://www.semanticweb.org/OKM_Ontology#Statement">

<rdfs:subClassOf rdf:resource="http://www.semanticweb.org/OKM_Ontology#Thing"/>

<rdf:type rdf:resource="http://www.w3.org/2002/07/owl#Class"/>

</rdf:Description>

<rdf:Description rdf:about="http://www.semanticweb.org/OKM_Ontology#isContainedBy">

<rdfs:range rdf:resource="http://xmlns.com/foaf/0.1/Document"/>

<rdfs:domain rdf:resource="http://www.semanticweb.org/OKM_Ontology#Statement"/>

<rdf:type rdf:resource="http://www.w3.org/2002/07/owl#ObjectProperty"/>

</rdf:Description>

<rdf:Description rdf:about="http://www.semanticweb.org/OKM_Ontology#NeAmacla">

<rdfs:subClassOf rdf:resource="http://www.semanticweb.org/OKM_Ontology#Neden"/>

<rdf:type rdf:resource="http://www.w3.org/2002/07/owl#Class"/>

</rdf:Description>

<rdf:Description rdf:about="http://www.semanticweb.org/OKM_Ontology#HangiAracla">

<rdfs:subClassOf rdf:resource="http://www.semanticweb.org/OKM_Ontology#Nasil"/>

<rdf:type rdf:resource="http://www.w3.org/2002/07/owl#Class"/>

</rdf:Description>

<rdf:Description rdf:about="http://www.semanticweb.org/OKM_Ontology#hasText">

<rdfs:range rdf:resource="http://www.w3.org/2001/XMLSchema#string"/>
<rdfs:domain
rdf:resource="http://www.semanticweb.org/OKM_Ontology#Statement"/>
<rdf:type rdf:resource="http://www.w3.org/2002/07/owl#DatatypeProperty"/>
</rdf:Description>
<rdf:Description rdf:about="http://www.semanticweb.org/OKM_Ontology#Neye">
<rdfs:subClassOf
rdf:resource="http://www.semanticweb.org/OKM_Ontology#Ne"/>
<rdf:type rdf:resource="http://www.w3.org/2002/07/owl#Class"/>
</rdf:Description>
<rdf:Description rdf:about="http://www.semanticweb.org/OKM_Ontology#Nereye">
<rdfs:subClassOf
rdf:resource="http://www.semanticweb.org/OKM_Ontology#Nerede"/>
<rdf:type rdf:resource="http://www.w3.org/2002/07/owl#Class"/>
</rdf:Description>
<rdf:Description rdf:about="http://www.semanticweb.org/OKM_Ontology#Nasil">
<rdfs:subClassOf
rdf:resource="http://www.semanticweb.org/OKM_Ontology#BesNbirK"/>
<rdf:type rdf:resource="http://www.w3.org/2002/07/owl#Class"/>
</rdf:Description>
<rdf:Description rdf:about="http://www.semanticweb.org/OKM_Ontology#Neyin">
<rdfs:subClassOf
rdf:resource="http://www.semanticweb.org/OKM_Ontology#Ne"/>
<rdf:type rdf:resource="http://www.w3.org/2002/07/owl#Class"/>
</rdf:Description>
<rdf:Description rdf:about="http://www.semanticweb.org/OKM_Ontology">
<rdf:type rdf:resource="http://www.w3.org/2002/07/owl#Ontology"/>
</rdf:Description>
<rdf:Description rdf:about="http://www.semanticweb.org/OKM_Ontology#Nerede">
<rdfs:subClassOf
rdf:resource="http://www.semanticweb.org/OKM_Ontology#BesNbirK"/>
<rdf:type rdf:resource="http://www.w3.org/2002/07/owl#Class"/>
</rdf:Description>
<rdf:Description rdf:about="http://www.semanticweb.org/OKM_Ontology#Neresi">

<rdfs:subClassOf
rdf:resource="http://www.semanticweb.org/OKM_Ontology#Nerede"/>

<rdf:type rdf:resource="http://www.w3.org/2002/07/owl#Class"/>

</rdf:Description>

<rdf:Description
rdf:about="http://www.semanticweb.org/OKM_Ontology#Nereden">

<rdfs:subClassOf
rdf:resource="http://www.semanticweb.org/OKM_Ontology#Nerede"/>

<rdf:type rdf:resource="http://www.w3.org/2002/07/owl#Class"/>

</rdf:Description>

<rdf:Description
rdf:about="http://www.semanticweb.org/OKM_Ontology#HangiYontemle">

<rdfs:subClassOf
rdf:resource="http://www.semanticweb.org/OKM_Ontology#Nasil"/>

<rdf:type rdf:resource="http://www.w3.org/2002/07/owl#Class"/>

</rdf:Description>

<rdf:Description
rdf:about="http://www.semanticweb.org/OKM_Ontology#BesNbirK">

<rdfs:subClassOf
rdf:resource="http://www.semanticweb.org/OKM_Ontology#Thing"/>

<rdf:type rdf:resource="http://www.w3.org/2002/07/owl#Class"/>

</rdf:Description>

<rdf:Description
rdf:about="http://www.semanticweb.org/OKM_Ontology#NeZaman">

<rdfs:subClassOf
rdf:resource="http://www.semanticweb.org/OKM_Ontology#BesNbirK"/>

<rdf:type rdf:resource="http://www.w3.org/2002/07/owl#Class"/>

</rdf:Description>

<rdf:Description
rdf:about="http://www.semanticweb.org/OKM_Ontology#HangiEmrediciGerekle">

<rdfs:subClassOf
rdf:resource="http://www.semanticweb.org/OKM_Ontology#Neden"/>

<rdf:type rdf:resource="http://www.w3.org/2002/07/owl#Class"/>

</rdf:Description>

<rdf:Description rdf:about="http://www.semanticweb.org/OKM_Ontology#Ne">

<rdfs:subClassOf
rdf:resource="http://www.semanticweb.org/OKM_Ontology#BesNbirK"/>

```
    <rdf:type rdf:resource="http://www.w3.org/2002/07/owl#Class"/>
  </rdf:Description>
</rdf:RDF>
```

EK-B

SÖZLÜK ONTOLOJİSİ

```
<rdf:RDF
   xmlns:rdf="http://www.w3.org/1999/02/22-rdf-syntax-ns#"
   xmlns:owl="http://www.w3.org/2002/07/owl#"
   xmlns:skos="http://www.w3.org/2004/02/skos/core#"
   xmlns:rdfs="http://www.w3.org/2000/01/rdf-schema#"
   xmlns:okm="http://www.semanticweb.org/OKM_Ontology#"
   xmlns:xsd="http://www.w3.org/2001/XMLSchema#" >
  …
  <rdf:Description rdf:about="http://www.semanticweb.org/OKM_Ontology#fosgen">
    <skos:broader rdf:resource="http://www.semanticweb.org/OKM_Ontology#gaz"/>
  </rdf:Description>
<rdf:Description rdf:about="http://www.semanticweb.org/OKM_Ontology#mimar">
    <skos:broader rdf:resource="http://www.semanticweb.org/OKM_Ontology#kişi"/>
  </rdf:Description>
<rdf:Description
rdf:about="http://www.semanticweb.org/OKM_Ontology#sarnıç_vagon">
    <skos:broader
rdf:resource="http://www.semanticweb.org/OKM_Ontology#vagon"/>
  </rdf:Description>
<rdf:Description
rdf:about="http://www.semanticweb.org/OKM_Ontology#itibarsızlık">
    <skos:broader
rdf:resource="http://www.semanticweb.org/OKM_Ontology#durum"/>
```

```
 </rdf:Description>
<rdf:Description
rdf:about="http://www.semanticweb.org/OKM_Ontology#kontenjan">
   <skos:narrower
rdf:resource="http://www.semanticweb.org/OKM_Ontology#öğrenci_kontenjan"/>
   <skos:narrower
rdf:resource="http://www.semanticweb.org/OKM_Ontology#stajyer_kontenjan"/>
   <skos:broader rdf:resource="http://www.semanticweb.org/OKM_Ontology#oran"/>
   <skos:broader
rdf:resource="http://www.semanticweb.org/OKM_Ontology#topluluk"/>
   <skos:broader rdf:resource="http://www.semanticweb.org/OKM_Ontology#ölçü"/>
 </rdf:Description>
<rdf:Description
rdf:about="http://www.semanticweb.org/OKM_Ontology#piyanoculuk">
   <skos:broader rdf:resource="http://www.semanticweb.org/OKM_Ontology#iş"/>
 </rdf:Description>
<rdf:Description
rdf:about="http://www.semanticweb.org/OKM_Ontology#siyaset_bilim">
   <skos:broader rdf:resource="http://www.semanticweb.org/OKM_Ontology#dal"/>
   <skos:broader
rdf:resource="http://www.semanticweb.org/OKM_Ontology#bilim"/>
 </rdf:Description>
<rdf:Description rdf:about="http://www.semanticweb.org/OKM_Ontology#silikoz">
   <skos:broader
rdf:resource="http://www.semanticweb.org/OKM_Ontology#hastalık"/>
 </rdf:Description>
<rdf:Description
rdf:about="http://www.semanticweb.org/OKM_Ontology#kese_kağıt">
   <skos:broader
rdf:resource="http://www.semanticweb.org/OKM_Ontology#torba"/>
 </rdf:Description>
<rdf:Description
rdf:about="http://www.semanticweb.org/OKM_Ontology#özel_öğrenci">
   <skos:broader
rdf:resource="http://www.semanticweb.org/OKM_Ontology#öğrenci"/>
 </rdf:Description>      ….
</rdf:RDF>
```

EK-C

OKM SİSTEM SÜRDÜRÜLEBİLİRLİĞİ ANKETİ

Bu anket, sizin deneyimleriniz doğrultusunda gerekli görülmesi halinde sistemde değişiklikler yapmak ve sisteminin sürdürülebilirliğini ölçmek için hazırlanmıştır. Yaptığınız değerlendirmelerin araştırmaya önemli katkıları olacaktır.

Katılımınız için teşekkürler.

Aşağıdaki değerlendirme sorularını

1; Kesinlikle katılmıyorum

2; Katılmıyorum

3; Kararsızım

4; Katılıyorum

5; Kesinlikle katılıyorum

olacak şekilde size uygun şekilde yanıtlayınız.

	1	2	3	4	5
1) Dokümanlardaki bilgileri ontolojik gösterimle haritalayan metodoloji, aradığım bilgilere yeterli düzeyde yanıt verebilmektedir.					
2)Ontolojik bilgi gösterimi üzerinde yaptığım sorgulamalar, bilgiler ve bilgiler arasındaki ilişkileri ortaya koyabilmektedir.					
3)OKM sistemi bilgiye erişim kolaylığı sunmaktadır.					
4)İhtiyaç duyduğum bilgiye farklı yollarla kolaylıkla erişmekteyim.					
5)Sistemi kullanırken yardıma ihtiyaç duymaktayım.					
6)Sistem, ihtiyaç duyduğum bilginin kaynağına ulaşımımı kolaylaştırmaktadır.					
7)Sistem, ihtiyaç olunan bilgiye erişim ile yeni bilgi çıkarımı yapmama yardımcı olmaktadır.					
8)Sorgulamalar sonucunda çıkan bilgiler ve ilişkiler					

	1	2	3	4	5
kolayca anlaşılabilir durumdadır.					
9) Sistemde, görsel aramalarda ilişki akışı gösterimine önem verilmiştir.					
10) Sistem, farklı arama metotlarıyla sorgulama yapmama imkan sunmaktadır.					
11) Sistem, görsel ve SPARQL sorgulama metotlarıyla sorgulanan kavramla ilişkili daha fazla kavram ve ilişki göstermektedir.					
12) Sistem, ulaşmak istediğim bilginin başka yerlerde de geçmesi durumunu da göstermektedir.					
13) Aradığım bilginin başka dokümanlarla da olan farklı ilişkisini görebilmekteyim.					
14) Sorgulama sonucu çıkan bilgi ve ilişkiler tutarsızlıklar barındırmaktadır.					
15) Aradığım bilgiyle ilişkisi olamayacağını düşündüğüm bilgiler görmekteyim.					
16) Sorgulama sonucunda tutarsız olduğunu düşündüğüm bilgiler sistem güvenilirliğini düşürmektedir.					
17) Sorgulama sonucunda ilişki barındırmayan, tek bilgi ve bilgiler yer almaktadır.					
18) OKM sistemi sunduğu hizmetin gerekliliğini hissettirmektedir.					

Değerli zamanınızı ayırdığınız için teşekkür ederiz.

ÖZGEÇMİŞ

KİŞİSEL BİLGİLER

Adı Soyadı : Nurgül YÜZBAŞIOĞLU

Doğum Tarihi ve Yeri : 21.08.1987 – İstanbul

Yabancı Dili : İngilizce

E-posta : nurgul.yuzbashioglu@gmail.com

ÖĞRENİM DURUMU

Derece	Alan	Okul/Üniversite	Mezuniyet Yılı
Y. Lisans	Bilgisayar Mühendisliği	Yıldız Teknik Üniversitesi	2016
Lisans	Bilgisayar Mühendisliği	Kırgızistan-Türkiye Manas Üniversitesi	2011
Lise	Fen Bilimleri	İhsan Mermerci Süper Lisesi	2006

İŞ TECRÜBESİ

Yıl	Firma/Kurum	Görevi
2014-halen	Yıldız Teknik Üniversitesi	Araştırma Görevlisi
2012-2014	Bolu Abant İzzet Baysal Üniversitesi	Araştırma Görevlisi

YAYINLARI

Bildiri

1. Yüzbaşıoğlu, N., Aktaş, M. S., (2016). “Ontology Aided Knowledge Mapping Workflow Execution Environment for Semantic Information Extraction”, 6th World Conference on Innovation and Computer Sciences INSODE, 12-14 Mayıs 2016, Antalya.