

T.C.
YILDIZ TEKNİK ÜNİVERSİTESİ
FEN BİLİMLERİ ENSTİTÜSÜ

COĞRAFİ BİLGİ SİSTEMİ VE SEMANTİK WEB
ARACILIĞIYLA BİNA DETAYLARINA İLİŞKİN ÇOK
KAYNAKLI VE ÇOK ÖLÇEKLİ MEKANSAL VERİLERİN
ENTEGRASYONU VE ZENGİNLEŞTİRMESİ

Abdulkadir MEMDUHOĞLU

DOKTORA TEZİ
Harita Mühendisliği Anabilim Dalı
Uzaktan Algılama ve CBS Programı

Danışman
Prof. Dr. A. Melih BAŞARANER

Kasım, 2021

T.C.
YILDIZ TEKNİK ÜNİVERSİTESİ
FEN BİLİMLERİ ENSTİTÜSÜ

**COĞRAFİ BİLGİ SİSTEMİ VE SEMANTİK WEB ARACILIĞIYLA BİNA
DETAYLARINA İLİŞKİN ÇOK KAYNAKLI VE ÇOK ÖLÇEKLİ
MEKANSAL VERİLERİN ENTEGRASYONU VE ZENGİNLEŞTİRMESİ**

Abdulkadir MEMDUHOĞLU tarafından hazırlanan tez çalışması 08.11.2021 tarihinde aşağıdaki jüri tarafından Yıldız Teknik Üniversitesi Fen Bilimleri Enstitüsü Harita Mühendisliği Anabilim Dalı Uzaktan Algılama ve CBS Programı **DOKTORA TEZİ** olarak kabul edilmiştir.

Prof. Dr. A. Melih BAŞARANER
Yıldız Teknik Üniversitesi
Danışman

Jüri Üyeleri

Prof. Dr. A. Melih BAŞARANER, Danışman
Yıldız Teknik Üniversitesi

Doç. Dr. A. Melis UZAR, Üye
Yıldız Teknik Üniversitesi

Doç. Dr. Caner GÜNEY, Üye
İstanbul Teknik Üniversitesi

Prof. Dr. Türkay GÖKGÖZ, Üye
Yıldız Teknik Üniversitesi

Prof. Dr. N. Necla ULUĞTEKİN, Üye
İstanbul Teknik Üniversitesi

Danışmanım Prof. Dr. A. Melih BAŞARANER sorumluluğunda tarafımda hazırlanan COĞRAFİ BİLGİ SİSTEMİ VE SEMANTİK WEB ARACILIĞIYLA BİNA DETAYLARINA İLİŞKİN ÇOK KAYNAKLI VE ÇOK ÖLÇEKLİ MEKANSAL VERİLERİN ENTEGRASYONU VE ZENGİNLEŞTİRMESİ başlıklı çalışmada veri toplama ve veri kullanımında gerekli yasal izinleri aldığımı, diğer kaynaklardan aldığım bilgileri ana metin ve referanslarda eksiksiz gösterdiğimi, araştırma verilerine ve sonuçlarına ilişkin çarpıtma ve/veya sahtecilik yapmadığımı, çalışmam süresince bilimsel araştırma ve etik ilkelerine uygun davrandığımı beyan ederim. Beyanımın aksinin ispatı halinde her türlü yasal sonucu kabul ederim.

Abdulkadir MEMDUHOĞLU

İmza

Aileme

TEŞEKKÜR

Tez çalışması boyunca beni yönlendiren ve gelişmeye katkı sunan başta tez danışmanım kıymetli hocam Prof. Dr. A. Melih BAŞARANER'e, tez izleme komitesindeki değerli hocalarım Doç. Dr. Caner GÜNEY ve Dr. Öğr. Üyesi A. Melis UZAR'a ve eğitimim boyunca aldığım derslerden edindiğim mesleki bilgiler ve desteklerinden ötürü Yıldız Teknik Üniversitesi Harita Mühendisliği Bölümü öğretim üyelerine teşekkür ederim.

Araştırmalarım ve tez yazım süresince sürekli bilgi alışverişinde bulunduğum ve çalışmalarına katkı sunan çalışma arkadaşlarım Arş. Gör. Dr. Müslüm HACAR, Arş. Gör. Hüseyin DUMAN, Doç. Dr. Zeynel Abidin POLAT, Arş. Gör. Mert GÜRTÜRK, Arş. Gör. Batuhan KILIÇ, Arş. Gör. Kadir ŞAHBAZ, Doç. Dr. Mustafa ULUKAVAK, Arş. Gör. Halil İbrahim ŞENOL, Arş. Gör. Yunus KAYA ve bana zaman ayırıp fikirlerini esirgemeyen Dr. Öğr. Üyesi Nizar POLAT'a teşekkürlerimi sunarım.

Tez çalışmalarımda kullandığım veriler için İstanbul Büyükşehir Belediyesi ve Harita Genel Müdürlüğüne teşekkür ederim.

Son olarak, araştırmalarım süresince bana anlayış gösteren, desteklerini hiç bir zaman esirgemeyen ve bunu hissettiren aileme ve eşim Bahar'a en içten teşekkürlerimi sunarım.

Abdulkadir MEMDUHOĞLU

İÇİNDEKİLER

SİMGE LİSTESİ	vii
KISALTMA LİSTESİ	viii
ŞEKİL LİSTESİ	x
TABLO LİSTESİ	xii
ÖZET	xiii
ABSTRACT	xv
1 GİRİŞ	1
1.1 Literatür Özeti	1
1.2 Tezin Amacı	3
1.3 Hipotez	3
2 MEKANSAL VERİ ENTEGRASYONU VE ZENGİNLEŞTİRMESİ	5
2.1 Çok Temsilli Mekansal Veri Modelleme	5
2.2 Mekansal Veri Eşleştirme ve Birleştirme	8
2.2.1 Geometrik Boyut	9
2.2.2 Semantik Boyut	10
3 MEKANSAL SEMANTİK WEB	12
3.1 Semantik Web	12
3.1.1 Bağlantılı Veriler	12
3.1.2 Ontolojiler	16
3.1.3 SPARQL Protokolü ve RDF Sorgu Dili (SPARQL)	18
3.1.4 Semantik Web Kural Dili (SWRL)	19
3.2 Mekansal Veriler İçin Semantik Web	19
4 YÖNTEM VE UYGULAMA	23
4.1 Çalışma Alanı	23
4.2 Uygulama 1: Semantik Web Aracılığıyla Mekansal Veri Entegrasyonu .	23

4.2.1	Kullanılan Veriler	24
4.2.2	Yöntem	25
4.2.3	Bulgular	37
4.2.4	Tartışma	40
4.3	Uygulama 2: Semantik Web Aracılığıyla Mekansal Veri Zenginleştirilmesi ve Binaların Ontoloji Tabanlı Sınıflandırılması	44
4.3.1	Kullanılan Veriler	45
4.3.2	Yöntem	45
4.3.3	Bulgular	53
4.3.4	Tartışma	56
5	SONUÇ VE ÖNERİLER	60
	KAYNAKÇA	63
A	KAYNAK KODLAR	73
B	KESİŞİM ALANI ORANI PERFORMANS DEĞERLERİ	77
C	SINIF TABLOLARI	78
	TEZDEN ÜRETİLMİŞ YAYINLAR	89

SİMGE LİSTESİ

$hd f_i$	Hedef veri setindeki bina alanı
O_{KA-max}	Kesişen alanların büyük alana oranı
O_{KA}	Kesişen alanların küçük alana oranı
ref_i	Referans veri setindeki bina alanı

KISALTMA LİSTESİ

ATKIS	Authorative Topographic-Cartographic Information System
BÖHHBÜY	Büyük Ölçekli Harita ve Harita Bilgileri Üretim Yönetmeliği
CBS	Coğrafi Bilgi Sistemleri
CNN	Convolutional Neural Network
CRS	Coordinate Reference System
ÇTMVT	Çok Temsilli Mekansal Veri Tabanları
DC	Dublin Core
DL	Description Logic
DUL	DOLCE+DnS Ultralite
GML	Geography Markup Language
HTTP	HyperText Transfer Protocol
IMTOP	Information Model Topography
INSPIRE	Infrastructure for Spatial Information in the European Community
IRI	Internationalized Resource Identifier
JSON-LD	JavaScript Object Notation-Linked Data
KBS	Kent Bilgi Sistemi
LOD	Linked Open Data
MRSL	Multiple Representation Schema Language
OCL	Object Constraint Language
OGC	Open Geospatial Consortium
OSM	OpenStreetMap
OWL	Web Ontology Language
POI	Point Of Interest

RCC8	Region Connection Calculus 8
RDF	Resource Description Framework
RDFS	Resource Description Framework Schema
SHACL	Shape Constraint Language
SKOS	Simple Knowledge Organisation System
SPARQL	SPARQL Protocol And RDF Query Language
SPIN	SPARQL Inference Notation
SQL	Stuctured Query Language
SWRL	Semantic Web Rule Language
TUCBS	Türkiye Ulusal Coğrafi Bilgi Sistemi
UHK	Ulusal Haritacılık Kurumu
UML	Unified Modelling Language
URI	Uniform Resource Identifier
URL	Uniform Resource Locator
VGI	Volunteered Geographic Information
W3C	World Wide Web Consortium
WKT	Well-Known Text
WWW	World Wide Web
XML	Extensible Markup Language

ŞEKİL LİSTESİ

Şekil 2.1	ÇTMVT’de gerçek dünya nesneleriyle coğrafi ve kartografik veri tabanları ilişkisi (Basaraner (2016)’dan uyarlanmıştır)	7
Şekil 3.1	Örnek bir RDF yapısının grafik ve tablo şeklinde gösterimi Memduhoglu ve Basaraner (2017)’den uyarlanarak)	13
Şekil 3.2	URL, URI ve IRI arasındaki farklar (Gandon (2014)’ten düzenlenerek)	14
Şekil 3.3	Linked Open Data Cloud üzerinde yer alan bağlantılı veri ağı (LOD, 2021)	15
Şekil 3.4	OWL’nin sınıflar, özellikler ve bireylerden oluşan temel bileşenleri (Horridge vd. (2011)’den uyarlanmıştır)	16
Şekil 3.5	Bir SPARQL sorgusunun yapısı (Feigenbaum (2008)’dan düzenlenerek)	18
Şekil 3.6	Geçişken ve simetrik özelliklerin örnek gösterimi	22
Şekil 4.1	Çalışma alanı	24
Şekil 4.2	Uygulamada kullanılan veriler	25
Şekil 4.3	Genel metodoloji ve çalışmanın uygulama adımları	26
Şekil 4.4	Çalışma alanı üzerinde seçilen alt alan (a) ve veri setlerinin (b, c, d) alt alanlardaki gösterimi	30
Şekil 4.5	Alan kesişim oranı ile kesinlik, duyarlılık ve F_1 -skor değerleri	30
Şekil 4.6	Binalar arası ilişki türlerinin tanımlama kuralları (Wiemann ve Bernard (2016)’dan uyarlanmıştır)	32
Şekil 4.7	Binalar ve meskun alan arasındaki ilişki türlerinin tanımlama kuralları	32
Şekil 4.8	Binalar ve meskun alan arasındaki ilişki türlerinin tanımlama kuralları	33
Şekil 4.9	Mekansal ontoloji sınıfları ve ilişkileri	35
Şekil 4.10	Birinci sorgudan gelen cevabın web haritası üzerinde gösterimi	38
Şekil 4.11	İkinci sorgudan gelen cevabın web haritası üzerinde gösterimi	39
Şekil 4.12	Üçüncü sorgudan gelen cevabın web haritası üzerinde gösterimi	40
Şekil 4.13	Birinci soru için GeoSPARQL ve SPARQL ile yapılan sorgular arasındaki mekansal farklılıklar	42
Şekil 4.14	Seçilen bina detaylarının gösterimi	47
Şekil 4.15	Uygulamanın işlem adımları	48
Şekil 4.16	Mekansal ontolojinin yerleşim sınıfı altında oluşturulan ayrıntılı sınıflar	50
Şekil 4.17	Etiket değerlerinin ontolojideki gösterimi	52

Şekil 4.18 Çıkarım motoru çalıştırıldıktan sonra ortaya çıkan sınıflandırma örneği	54
Şekil 4.19 Resmi Bina sınıfı ayrıntılarını gösteren ontoloji (a) ve web haritası görüntüsü (b)	55
Şekil 4.20 KarmaBina tanımlaması ile elde edilen binaları gösteren ontoloji (a) ve web haritası görünümü (b)	57
Şekil 4.21 Ontolojideki üst sınıflar (b) ve buna göre oluşturulmuş tematik harita (a)	58
Şekil 4.22 TOPO1B detaylarının sınıf sayılarını ve dağılımlarını gösteren harita	59

TABLO LİSTESİ

Tablo 2.1	Farklı iki veri setinde yer alan bina detayları arasındaki olası ilişkiler (Fan, Zipf, Fu ve Neis (2014)’ten uyarlanmıştır)	11
Tablo 4.1	Çalışmada kullanılan veri setleri, kısaltmaları ve detay sayıları . . .	25
Tablo 4.2	Performans değerlendirmesi için kullanılan karışıklık matrisi	28
Tablo 4.3	Otomatik eşleştirmenin TP, FP ve FN değerleri	31
Tablo 4.4	Eşleştirme performans istatistikleri	31
Tablo 4.5	Kardinal ilişki türleri ve özellikleri	32
Tablo 4.6	İlişki türlerine göre oluşturulan ilişki sayıları	37
Tablo 4.7	Birinci soru için GeoSPARQL ve SPARQL ile yapılan sorguların ortalama cevap süreleri	42
Tablo 4.8	Seçilen bina detay sayıları	46
Tablo 4.9	Leisure anahtarı için oluşturulan kapsam ve sınıf eşleniği tablosu . .	52
Tablo B.1	O_{KA} aralıklarının Doğru Pozitif (TP), Yanlış Pozitif (FP) ve Yanlış Negatif (FN) değerleri	77
Tablo C.1	building anahtarı için oluşturulan kapsam ve sınıf eşleniği tablosu .	78
Tablo C.2	shop anahtarı için oluşturulan kapsam ve sınıf eşleniği tablosu . . .	80
Tablo C.3	amenity anahtarı için oluşturulan kapsam ve sınıf eşleniği tablosu .	82
Tablo C.4	leisure anahtarı için oluşturulan kapsam ve sınıf eşleniği tablosu . .	84
Tablo C.5	historic anahtarı için oluşturulan kapsam ve sınıf eşleniği tablosu . .	85
Tablo C.6	tourism anahtarı için oluşturulan kapsam ve sınıf eşleniği tablosu .	86
Tablo C.7	office anahtarı için oluşturulan kapsam ve sınıf eşleniği tablosu . . .	87
Tablo C.8	healthcare anahtarı için oluşturulan kapsam ve sınıf eşleniği tablosu	88

COĞRAFİ BİLGİ SİSTEMİ VE SEMANTİK WEB ARACILIĞIYLA BİNA DETAYLARINA İLİŞKİN ÇOK KAYNAKLI VE ÇOK ÖLÇEKLİ MEKANSAL VERİLERİN ENTEGRASYONU VE ZENGİNLEŞTİRMESİ

Abdulkadir MEMDUHOĞLU

Harita Mühendisliği Anabilim Dalı

Doktora Tezi

Danışman: Prof. Dr. A. Melih BAŞARANER

Web 2.0 uygulamalarının ortaya çıkışıyla birlikte kullanıcılar, web'e içerik sağlayabilir olmuşlardır. Bu durum web'te yer alan veri miktarının hızla artmasına yol açmıştır. Tüm alanlarda olduğu gibi mekansal bilgi alanı da bu durumdan etkilenmiştir. Özellikle gönüllü coğrafi bilgi (VGI) uygulamalarının ortaya çıkmasıyla genellikle standart olmayan mekansal veriler artış göstermiştir. Buna paralel olarak, ücretsiz olarak servis edilen bu mekansal verilerin resmi mekansal verilerle entegre edilerek kullanıldığı uygulamalar yaygınlaşmaya başlamıştır. Örneğin, OpenStreetMap (OSM) gibi VGI uygulamaları, kriz koşullarında ihtiyaç duyulan bölgelerin haritalarını ortaklaşa olarak (crowdsourcing) hızlıca oluşturma veya güncelleme olanağı sağlamaktadır. Bu mekansal veri çeşitliliğinden azami yarar sağlayabilmek ve mevcut mekansal verileri daha hızlı ve az maliyetle güncelleyebilmek için öncelikle verilerin entegrasyonu ve zenginleştirilmesi gerekmektedir. Entegrasyon sadece zenginleştirme amaçlı değil aynı zamanda farklı ayrıntı düzeylerinde yer alan mekansal verilerin birbirleriyle ilişkilendirilmesi için de gereklidir. Bu bağlamda, 2000'li yılların başında ortaya konmuş ve zamanla gelişerek mekansal bilgi alanında da yer edinmiş olan semantik web yaklaşımı ve teknikleri bir potansiyele sahiptir. Makinelerin olguları analiz ederek insanların yerine karar verebildiği bu çağda, mekansal verilerin anlamlarının makineler için de tanımlanabildiği, analiz ve çıkarım yapılabilen semantik web ortamında kullanılması bu alanda atılacak önemli bir

adımdır.

Bu tez kapsamında farklı kaynaklardan gelen farklı ölçeklerdeki mekansal veriler alan kesişimine göre geometrik olarak eşleştirilmiş ve semantik web yöntemleriyle entegrasyonu ve zenginleştirilmesi sağlanmıştır. Bağlantılı veri şeklinde ortaya konan bu entegre edilmiş ve zenginleştirilmiş verilerle farklı ihtiyaçlara aynı yerden cevap verebilme imkanı uygulamalarda gösterilmiştir. Bunun yanı sıra semantik web'in en güçlü yanlarından olan mantıksal çıkarımlar ile tanımlanmamış detay ilişkileri ve çeşitli sınıflar elde edilerek örneklerle gösterilmiştir. Elde edilen sonuçlar semantik web tekniklerinin mekansal bilgi alanında kullanılmasının makinelerin bilgi çıkarımı konusuna önemli katkılar sunabileceğini ortaya koymuştur.

Anahtar Kelimeler: mekansal veri entegrasyonu, mekansal semantik web, çoklu temsil, mekansal veri zenginleştirilmesi, geometrik eşleştirme

INTEGRATION AND ENRICHMENT OF MULTI-SOURCE AND MULTI-SCALE SPATIAL DATA OF BUILDING FEATURES THROUGH GEOGRAPHIC INFORMATION SYSTEM AND SEMANTIC WEB

Abdulkadir MEMDUHOĞLU

Department of Geomatic Engineering
Doctor of Philosophy Thesis

Supervisor: Prof. Dr. A. Melih BAŞARANER

With the emergence of Web 2.0 applications, users have been able to provide content to the web. This has led to a rapid increase in the amount of data on the web. As in all fields, the spatial information has also been affected by this situation. Especially with the advent of Voluntary Geographic Information (VGI) applications, a large amount of, but often non-standard, spatial data has increased. Parallel to this, applications in which these spatial data, which are served free of charge, are integrated with official spatial data have become widespread. For example, VGI applications such as OpenStreetMap (OSM) provide the opportunity to quickly create or update maps of regions needed in crisis conditions jointly (crowdsourcing). In order to get the maximum benefit from this spatial data diversity and update the existing spatial data faster and with less cost, the integration and enrichment of the data is required first. Integration is not only necessary for enrichment, but also for associating spatial data with each other at different levels of detail. In this context, the semantic web approach and techniques, which were introduced in the early 2000s and developed over time, have a potential in the field of spatial information. In this age, where machines can analyze facts and decide instead of people, using spatial data in a semantic web environment where machines can understand, analyze and make inferences is an important step to be taken in this field.

Within the scope of this thesis, spatial data at different scales from different sources are

geometrically matched based on the area overlap and their integration and enrichment with semantic web methods. The possibility of responding to different needs from the same data set with this integrated and enriched data presented in the form of linked data has been demonstrated in the applications. In addition, logical inferences, which are the strongest aspects of the semantic web, and undefined feature relations and various classes are obtained and shown with examples. The results revealed that the use of semantic web techniques in the field of spatial information can make significant contributions to the knowledge inferences by machines.

Keywords: geospatial data integration, geospatial semantic web, multiple representation, geospatial data enrichment, geometric matching

1

GİRİŞ

Bu bölümde, ilk olarak mekansal veri entegrasyonu ve zenginleştirilmesi kapsamındaki semantik web temelli çalışmalara ilişkin literatür özeti verilmiştir. Bu bağlamda, spesifik aşamalara yönelik literatür çalışmaları ilgili konu başlıkları içinde daha ayrıntılı olarak incelenmiştir. Ardından, tezin amacı açıklanmış ve bununla ilintili hipotezler ortaya konmuştur.

1.1 Literatür Özeti

Mekansal veriler, çeşitli amaçlar için geometrik ve semantik olarak farklı ayrıntı düzeylerinde (çözünürlük/ölçek) ve farklı kişi/kurumlar tarafından üretilmektedirler. Bununla birlikte mekanın zamansal değişimi bu mekansal verilerdeki farklılığı arttırabilmektedir. Bu durum, aynı mekansal nesneler için farklı temsillerin ortaya çıkmasına yani heterojenliklere yol açmaktadır (Basaraner, 2012, 2016). Örneğin, bir bina büyük ölçekli bir mekansal veri setinde çokgen geometriyle, küçük ölçekli bir mekansal veri setinde ise nokta geometriyle temsil edilebilir. Benzer biçimde, semantik açıdan bu iki farklı mekansal veri setindeki binaların sınıfı ve öznitelikleri de farklılık gösterebilir. Bununla birlikte, mekansal verilerin kullanım amacı veya uygulama alanı da verilerin farklı temsiline yol açabilir. Örneğin, bir nehir; su yönetimi alanında çokgen geometri ve doğal yaşamla ilgili öznitelikler ile temsil edilirken, hidroloji alanında çizgi geometri ve fiziki öznitelik bilgileriyle temsil edilebilmektedir (Yi, 2013). Aynı gerçek dünya nesnesinin farklı temsiliyle ortaya çıkan bu heterojenlik durumu, mekan bilimleri alanında birleştirme (conflation) başlığı altında incelenirken, semantik web alanında ise ortak referans (co-reference) olarak bilinmektedir (Ruiz-Lendínez vd., 2011). Birleştirme, mekansal veri setlerinin entegrasyonu için kullanılan yöntemlere verilen genel addır. Bu yöntemler genelde veri setleri arasındaki eşlenikleri bulmak için çeşitli ölçüler (çoğunlukla benzerlik) kullanırlar (Schwering, 2006; Xavier vd., 2016; Yan & Li, 2015; Yu vd., 2018). Öte yandan, ortak referans bir Tekbiçimli Kaynak Belirleyici (URI) ile birden çok kaynağın

veya bir kaynağın farklı URI'ler ile tanımlandığı durumlarda ortaya çıkar. İkinci olarak belirtilen ve daha sık meydana gelen ortak referans sorunu, bu URI'ler arasında owl:sameAs bağlantıları üretilerek çözülebilir (Glaser vd., 2009; Vilches-Blázquez vd., 2012). İfade edildiği gibi, mekansal verilerin entegrasyonu için bir problem olan yapısal ve semantik heterojenlikler semantik web teknikleriyle giderilebilir (Memduhoglu & Basaraner, 2018).

Mekansal verilerdeki zamansal değişimlerin bu verilere işlenebilmesi için ülkelerin mekansal veri üretiminden sorumlu Ulusal Haritacılık Kurumu (UHK) belirli aralıklarla verilerini güncellemektedir. Bunun yanında yerel yönetimler de farklı amaçlar için çeşitli mekansal veriler üretilip bunları sürekli güncellemektedir (Homburg vd., 2019). Örneğin, belediyeler kent bilgi sistemleri oluşturarak bu verileri haritalar üzerinden halka sunmaktadırlar. Bu güncelleme faaliyetleri sırasında popüler bir Gönüllü Coğrafi Bilgi (VGI) uygulaması olan OpenStreetMap (OSM) verilerinin mevcut verilere entegrasyonu bu kurumlara hız ve maliyet anlamında bir avantaj sağlayabilir (Çabuk vd., 2015). OSM'de gönüllü kullanıcılar projeye veri girişi sağlayarak haritanın güncelliğine katkı sağlamaktadırlar. Bu süreç, belirli aralıklarla haritalarını güncelleyen UHK'lardan farklı işleyerek mekansal gerçekliğin değiştiği bazı durumlarda haritanın çok kısa bir sürede güncellenmesine olanak sağlamaktadır (X. Zhang vd., 2018). Bu duruma geometrik verilerden çok öznitelik verilerinde daha sık rastlanmaktadır. Örneğin, bir binanın kullanım türü değişikliği OSM'deki yerel kullanıcılar tarafından daha hızlı bir şekilde tespit edilerek haritaya işlenebilir. Bunların dışında resmi veri setlerinde yer almayan ve OSM'de yer alan bazı öznitelikler de verilerin entegrasyonu ile kullanılabilir (Du vd., 2017; Du vd., 2012). Örneğin, OSM'de yer alan engelli bireyler için tekerlekli sandalyeye uygun binaların resmi verilere entegrasyonu sağlanabilir. İfade edilen bu veri entegrasyonu ve zenginleştirilmesi işlemleri için mekansal semantik web uygun bir araç olarak kullanılabilir.

Semantik web, Berners-Lee vd. (2001) tarafından mevcut web'in bir uzantısı olarak ortaya atılmıştır. Semantik web kavramı ortaya çıktığı ilk zamanlardan beri mekansal bilgi alanı dahil olmak üzere birçok bilim alanı tarafından kullanılmaktadır. Semantik web, çoğunlukla sadece insanların anlayabildiği dokümanlardan oluşan mevcut web'i, makineler için de verinin tanımlandığı ve bu verilerden mantıksal çıkarımlar yapılabilen veri web'ine dönüştürme fikrini temel almaktadır. Semantik web ile bağlantılı veri kavramı da sıklıkla birlikte kullanılmaktadır. Semantik web'de veriler Berners-Lee (2006) tarafından ortaya konan dört ilkeye uygun olduğunda bağlantılı veri adını almaktadır. Bu kurallar: (1) İsimler için URI kullanılması, (2) kaynaklar için Hiper Metin Aktarım Protokolü (HTTP) URI'lerinin kullanılması, (3) URI'ler için faydalı bilgilerin kullanılması, (4) keşif için başka kaynaklarla bağlantı kurulması

şeklinde sıralanabilir.

Farklı kaynaklardan ve farklı ayrıntı düzeylerinden gelen verileri içeren entegre edilmiş bir çok ölçekli mekansal veri seti oluşturmak, mekansal detaylar hakkında maksimum bilgi edinmeyi sağlar (Uitermark vd., 2005). Bu nedenle, aynı mekansal nesnenin farklı ölçeklerdeki geometrileri ve ilişkileri, çok ölçekli temsili destekleyen bir mekansal ontoloji oluşturmak için tanımlanmalı ve modellenmelidir. Bu entegre edilmiş ve zenginleştirilmiş çok ölçekli ve mekansal bağlantılı verilerle, tek bir veri setiyle elde edilemeyen bilgiler keşfedilebilir ve aynı gerçek dünya nesnesinin çoklu temsili ortaya konabilir.

Bu bağlamda, semantik web teknolojileri ile mekansal verilerin birlikte kullanımı, söz konusu entegrasyon sorunlarının üstesinden gelmek ve yapay zeka uygulamalarına desteğin yanı sıra akıllı uygulamalar geliştirerek gizli bilgilerin çıkarım yoluyla ortaya konması için yeni olanaklar sunmaktadır (Sun vd., 2019).

Tezin sonraki bölümünde mekansal veri entegrasyonu ve zenginleştirilmesi konusuna değinilecektir. Üçüncü bölümde, mekansal semantik web ve bağlantılı teknolojiler ele alınacaktır. Dördüncü bölümde tez kapsamında gerçekleştirilen uygulamalar ve yöntemleri sonuçlarıyla birlikte tartışılacaktır. Son bölümde ise sonuç ve önerilere yer verilecektir.

1.2 Tezin Amacı

Bu çalışmada farklı kaynaklardan gelen ve farklı ayrıntı düzeylerindeki bina verilerinin geometrik eşleştirme yoluyla semantik web ortamında entegre edilmesi, zenginleştirilmesi ve çok temsilli bir şekilde sunulması amaçlanmıştır.

1.3 Hipotez

“Çok kaynaklı ve çok ölçekli bina verileri entegre edilerek zenginleştirilebilir ve semantik web aracılığıyla sunulabilir” şeklindeki genel hipoteze dayanarak aşağıda sıralanan özel hipotezler bu tez kapsamında sınanmaktadır.

- Çok kaynaklı ve çok ölçekli bina verileri geometrik olarak birbiriyle eşleştirilebilir.
- Eşleştirilen bu bina verileri semantik web ortamında bağlantılı veri şeklinde çok temsilli olarak sunulabilir.

- Üç farklı ölçekte yer alan mekansal verilerden dolayı farklı ve karmaşık olan aynı binaların dikey kardinal ilişkileri semantik web’de mantıksal tanımlar ile ifade edilebilir ve buradan çeşitli çıkarımlar (yeni ilişkiler, yeni sınıflar vb.) yapılabilir.
- Birden çok bina veri setini içeren entegre edilmiş mekansal bağlantılı veride tek bir veri setiyle bulunamayacak cevaplar semantik web yardımıyla bulunabilir.
- OSM verileriyle zenginleştirilmiş bina verileri bina kullanım türlerine göre sınıflandırılabilir veya semantik web aracılığıyla farklı mantıksal tanımlamalarla çeşitli yeni sınıflar elde edilebilir.

MEKANSAL VERİ ENTEGRASYONU VE ZENGINLEŞTİRMESİ

Mekansal veriler üretilirken doğal olarak bir özetleme aşamasından geçerler. Bu aşamada aynı mekana ait nesnelerin temsili; üreticiye, amaca ve standarda göre farklılaşabilmektedir. Bunun yanında mekansal veriler, belirli bir çözünürlüğe/ölçeğe sahiptir. Dolayısıyla, aynı gerçek dünya nesnesinin farklı geometrik ve semantik ayrıntı düzeylerinde modellenmesi söz konusu olmaktadır. Ortaya çıkan bu farklı temsillerin farklı amaçlar doğrultusunda bir arada kullanılması için entegrasyon gereklidir. Ayrıca bir veri setinde yer almayan veya yeterli olmayan verilerin başka bir veri setiyle tamamlanması aynı zamanda verinin zenginleştirilmesi anlamına gelmektedir. Farklı ölçeklerden gelen verilerin bir arada kullanılabilmesi için ise çok temsilli bir modelleme gereklidir. Bununla birlikte entegrasyon ve zenginleştirme yapılabilmesi için öncelikle farklı veri setlerindeki aynı gerçek dünya nesnelerinin birbiriyle eşleştirilmesi ve ilişkilendirilmesi gerekmektedir. Bu bölümde mekansal veri entegrasyonu ve zenginleştirilmesi için gerekli olan çok temsilli mekansal veri modelleme, mekansal veri eşleştirme ve birleştirme konularına değinilecektir.

2.1 Çok Temsilli Mekansal Veri Modelleme

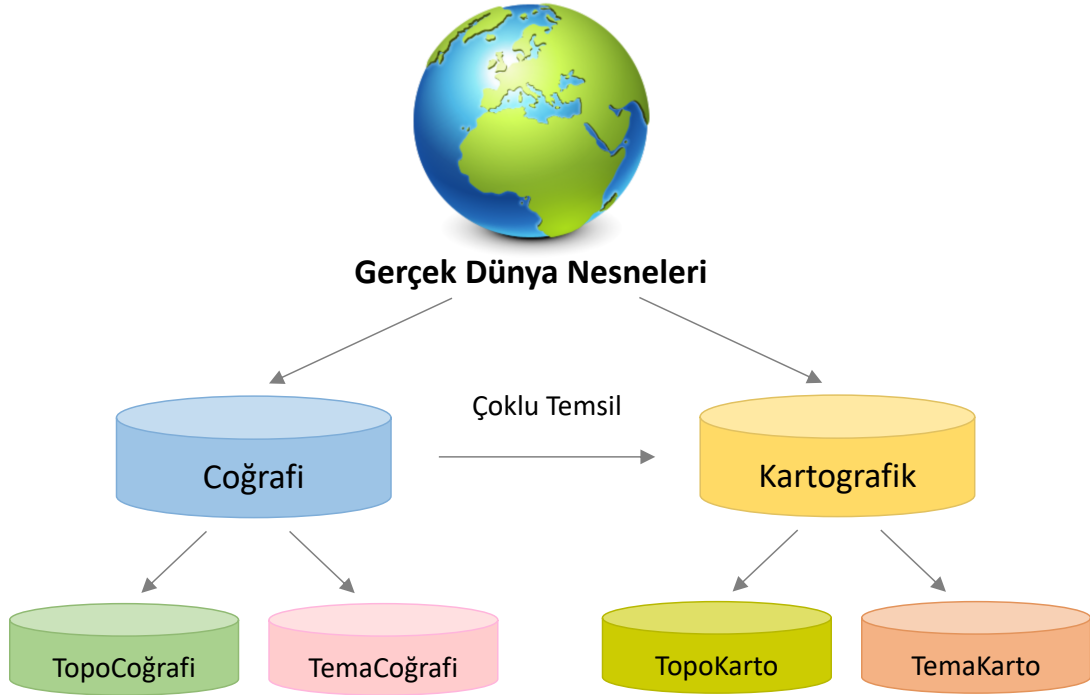
Mekansal veriler doğaları gereği belirli bir ölçekte üretilirler. UHK'lar genel olarak haritalarını gerekli olan en büyük ölçekte üretirken daha küçük ölçekleri bu temel haritalardan türeterek elde ederler. Farklı ölçeklerde detayların birbiriyle ilişkisi yoksa herhangi bir detay güncellendiğinde bu durum diğer ölçeklere aktarılamaz ve tutarsızlıklar ortaya çıkar. Bu tutarsızlıkların önüne geçebilmek adına detayların tüm ölçeklerde birbiriyle ilişkili olduğu Çok Temsilli Mekansal Veri Tabanları (ÇTMVT) yaklaşımı kullanılabilir. Bu yaklaşım, bazı araştırmacılar tarafından incelenmiş olsa da, (Basaraner, 2009; Bobzien vd., 2008; Devogele vd., 1996; Kilpeläinen, 2000; van Wijngaarden vd., 1997) verilerin büyük ölçekten kartografik genelleştirme ile küçük ölçeklere türetilmesi aşamasının karmaşıklığından dolayı UHK'lar tarafından

tam anlamıyla kullanılamamaktadır (Mustière & van Smaalen, 2007).

ÇTMVT, bir coğrafi alan hakkındaki çeşitli ilişkili mekansal veri setlerini yalnızca çok amaçlı coğrafi veri ve harita üretimi için değil, aynı zamanda farklı ayrıntı düzeylerinde (çözünürlük/ölçek) çok temsilli mekansal analiz ve görselleştirme için depolayarak daha karmaşık ve esnek bir ortam sunar. Başka bir ifadeyle, ÇTMVT paradigması, veri setlerinin üretimi, yönetimi ve bakımı gibi maliyetli ve zaman alıcı görevleri basitleştirir ve ilişkili veri setleri arasındaki tutarsızlıkları ortadan kaldırır.

Daha geniş bir perspektiften bakıldığında, bir ÇTMVT, farklı ayrıntı düzeylerinde birbirine bağlı topografik ve tematik coğrafi ve kartografik veri setlerini içerir (Basaraner, 2012, 2016). Mekansal veri setleri farklı amaçlara hizmet edecek şekilde kurulduğundan, aynı gerçek dünya nesnesi, bu veri setlerinde farklı detay sınıflarında ve/veya farklı semantik, geometrik ve/veya grafik tanımlarla temsil edilmektedir (Şekil 2.1). Daha net olmak gerekirse, bu heterojenlik esas olarak farklı semantik, geometrik ve/veya grafik çözünürlüklerden ve mekansal veri setlerinin topografya, navigasyon ve turizm gibi farklı temalar için oluşturulmasından kaynaklanmaktadır. Bir ÇTMVT oluşturmak için iki ana yöntem vardır: (1) Genelleştirme ve (2) entegrasyon. İlk yaklaşımda, mekansal veriler gereken en yüksek çözünürlükte veya en büyük ölçekte üretilir ve daha sonra genelleştirme yoluyla daha düşük çözünürlüklere veya daha küçük ölçeklere aktarılır. Genelleştirme işlemi sırasında, farklı ayrıntı düzeylerindeki mekansal detaylar arasındaki bağlantılar saklanır. İkinci yaklaşımda, ayrı fakat aynı gerçek dünya nesnelerini içeren veri setlerindeki mekansal detaylar, semantik ve geometrik eşleştirme kurallarına dayalı olarak birbirine bağlanır.

ÇTMVT, çok amaçlı mekansal veri modelleme, işleme, analiz, görselleştirme ve yayma/yayımlama için verimli bir ortam sağlar. Öte yandan, semantik web yöntem ve teknolojileri, otomatik entegrasyon ve çıkarım için umut vericidir. Bu anlamda, iki yaklaşımı birleştirmek, her iki paradigmanın güçlü yanlarından yararlanma olanağı sağlayacaktır. Literatürde, ÇTMVT'ye yönelik birkaç yaklaşım önerilmiştir. Bunlar arasında, bazı yaklaşımlar potansiyel olarak daha uygundur. Bunlardan ilki, “görünür öğelerin (geometri ve grafik işaretler) belirli bir semantikle benzersiz bir kombinasyonu” olarak tanımlanan VUEL isimli yaklaşımdır. VUEL, geometrik, semantik ve grafik boyutu birbirine bağlayan tabloyu sunmak için Birleşik Modelleme Dili (UML) notasyonunu kullanır (Bedard & Bernier, 2002). Diğer bir yaklaşım ise semantik web teknolojileriyle daha ilgili sayılabilecek damgalama (stamping) tekniğidir (Vangenot vd., 2002). Damgalama tekniği, nesneyi iki açıdan temsil etmek için ikili dosya biçimini benimser. Bu iki perspektiften, ilki bakış açısını ikincisi ise ayrıntı veya çözünürlük düzeyini temsil eder. Bu ikili tanımların her birine, (örneğin, “haritacılık”, 1 metre) bir damga denir ve damgalar, gerçek dünya



Şekil 2.1 ÇTMVT’de gerçek dünya nesneleriyle coğrafi ve kartografik veri tabanları ilişkisi (Basaraner (2016)’dan uyarlanmıştır)

nesnelerinin farklı temsillerini karakterize eder. Bu bağlamda, çoklu temsil için iki yaklaşım önerilmiştir: Özelleştirilmiş veri yapılarını tanımlamak için damgalama tekniğinin kullanıldığı bütünleşik yaklaşım ve temsillerin ilişkili bağlar aracılığıyla birbirine bağlandığı yaklaşım. Benslimane vd. (2003), damgalama tekniğini bir adım daha ileri götürmüş ve çoklu temsil amacıyla Tanımlama Mantığı (DL) tabanlı bir ontoloji dili oluşturmak için damgalama tekniğini DL ile birleştirmiştir. Bu yöntemle, DL ve damgalama tekniğine dayalı bir yapıcı (constructor) grubu oluşturulur ve MuRO (MultiRepresentation Ontologies) adı verilen çoklu temsil ontolojilerini desteklemek için kavramsal bir ontoloji dili sunulur. Kısacası MuRO, çeşitli bağlamlarda değişken bir dizi özellik veya nitelik ile belirli bir alanı ifade eden bir ontolojidir. Friis-Christensen vd. (2005), çoklu temsilleri yönetmek ve modellemek için bir çoklu temsil yönetim sistemi geliştirmiştir. Yazarlar, mekansal verilerde çoklu temsil bağlamında tutarlılık gereksinimlerini tanımlamak için kullanılan kavramsal bir şema dili olan Çoklu Temsil Şema Dili’ni (MRSL) tanıtmışlardır. MRSL’nin temel unsurları eşleştirme, tutarlılık ve restorasyon kurallarıdır. MRSL’nin grafik gösterimi, UML ve Nesne Kısıtlama Dili’nin (OCL) bir uzantısına dayanır. Ek açıklama ile bir UML meta sınıfının sanal bir alt sınıfını tanımlamanın bir yolunu sağlayan stereotipleri kullanırlar. Bu konuyla ilgili daha yeni bir çalışma Stoter vd. (2011) tarafından gerçekleştirilmiştir. Hem çok ölçekli veri içeriğinde hem de ölçek geçişlerinde zengin semantik bilgi içeren çok ölçekli bir topografik veri tabanı için bir bilgi modeli sunmuşlardır. Topografya Bilgi Modeli (IMTOP) adı verilen bu model, UML kullanılarak ifade edilmiş ve OCL ile

tamamlanmıştır. IMTOP, bir sınıfı belirli ölçeklerde modellemek için kalıtım, türetme ilişkileri ve türetilmiş öznitelikleri kullanır. Bu yaklaşımın avantajı, belirli bir harita ölçeği için sadece o ölçekle ilgili sınıfları gösteren okunması ve türetmesi basit bir model sunmasıdır. Dezavantajı ise, Kaynak Tanımlama Çerçevesi (RDF) veya Web Ontoloji Dili (OWL) gibi semantik teknolojiler yerine UML ve OCL ile ifade edilmesidir. Ayrıca, IMTOP'un RDF veya OWL gibi bir ontoloji diline aktarılmasının, verilerin geniş çapta yeniden kullanımı için gerekli olan makine düzeyinde daha fazla semantik bilgi sağlayacağı belirtilmiştir.

ÇTMVT ile ilgili incelenen çalışmalar, yöntemler ve araçların büyük bir çoğunluğu, ÇTMVT oluşturma sürecinin karmaşıklığı nedeniyle genel kabul görmemiştir. Özellikle, OCL ifadelerinin tanımını sağlayabilecek yazılım araçlarının olmaması ve çok uzun OCL ifadelerini içeren kafa karıştırıcı şemaların kullanılması bu yöntemlerin yaygın olarak kullanılmasını zorlaştırmaktadır. Bu teknikleri OWL veya RDF ile kullanmak, onlara güçlü bir standart desteği ve coğrafi bilgi topluluğu tarafından benimsenme kolaylığı sağlayacaktır. Ayrıca, OWL gibi DL içeren bir yaklaşımın uygulanması, ÇTMVT'deki veri setlerinden çıkarımlar yapılmasına izin verecektir.

Ek olarak, literatürde RDF/OWL gibi semantik web ve ÇTMVT yaklaşımlarının ilişkilendirildiği çalışmalar da yer almaktadır. Hahmann ve Burghardt (2010) semantik web ve ÇTMVT yaklaşımlarının benzerlik ve farklılıklarını inceleyerek semantik web yaklaşımıyla çoklu temsilin mümkün olup olmadığını araştırmıştır. Eiðsson (2018) kartografik ölçek kavramını semantik web ortamında tanımlayan bir çalışma ortaya koymuştur. Huang vd. (2018) semantik web'de tanımlanmış olan bu ölçek kavramını kullanarak farklı ölçeklerdeki geometrik temsillerin senkronizasyonunu sağlamıştır. Stoter vd. (2006) çoklu temsil için semantik veri entegrasyonuna başvurmuş ve Hollanda'nın topografik verilerinin farklı verilerle semantik olarak nasıl entegre edileceğini araştırmıştır. Bunların dışında X. X. Zhang vd. (2018) OSM güncellemelerini resmi veri setlerine ÇTMVT olarak aktaran bir yaklaşım ortaya koymuştur fakat bu yaklaşım semantik web yerine klasik Coğrafi Bilgi Sistemleri (CBS) yöntemlerini kullanmaktadır.

2.2 Mekansal Veri Eşleştirme ve Birleştirme

Mekansal veri entegrasyonu için ilk yapılması gereken işlem veri setlerindeki aynı gerçek dünya nesnelerinin eşleştirilmesidir. Mekansal veri eşleştirme işlemi semantik ve geometrik olarak iki farklı boyutta incelenebilir (Memduhoglu & Basaraner, 2018).

2.2.1 Geometrik Boyut

Mekansal verilerin eşleştirme bağlamında semantik verilere göre en büyük avantajı, aynı gerçek dünya nesnelerinin yeryüzündeki geometrik konumlarına göre birbiriyle ilişkilendirilebilir olmasıdır. Örneğin, birbirine yakın ölçekteki farklı veri setlerinde yer alan bir binanın farklı geometrilerinin (aynı koordinat sisteminde) belirli bir oranda örtüşmesi beklenir. Farklı veri setlerindeki aynı gerçek dünya nesnelerinin geometrileri temel alınarak yapılan bu tür bir eşleştirmeye geometrik eşleştirme denir. Mekansal verilerin geometrik eşleştirilmesi geometri türüne göre farklılık gösterebilmektedir. Temelde nokta, çizgi ve alan geometrilerine sahip mekansal detayların eşleştirilmesi için farklı yöntemler kullanılabilir. Sonraki bölümlerde geometri türüne göre nokta, çizgi, çokgen (alan) ve farklı geometrili olmak üzere dört farklı eşleştirme incelenecektir.

2.2.1.1 Nokta Geometri İle Temsil Edilen Detaylar

Mekansal nesnelerin bir kısmı, özetleme kuralları veya ölçek kısıtlamaları nedeniyle nokta geometrili olarak temsil edilirler. Bunlar; direk, rögar veya ağaç gibi araziden doğrudan nokta olarak toplanan detaylar olabildiği gibi ölçeğin küçültülmesine bağlı olarak genelleştirmeyle noktaya dönüştürülmüş bina gibi detaylar da olabilir. Ayrıca, ilgi noktası (POI) olarak adlandırılan özel yerler de nokta geometrili olarak temsil edilmektedir. Nokta geometrili detaylar birbirleriyle belirli bir eşik mesafesi değeri kullanılarak eşleştirilebilir. Örneğin, her ikisi de nokta geometriye sahip ilgi noktası ve bina eşleşmesi için, binalara uzaklığı 5 m eşik mesafesinden küçük olan ilgi noktaları bu binalarla eşleştirilebilir (Athanasίου vd., 2019; Hacı, 2019; Ruiz-Lendínez vd., 2017).

2.2.1.2 Çizgi Geometri İle Temsil Edilen Detaylar

Yol, demiryolu, akarsu ve idari sınır gibi birçok nesne, sıklıkla çizgi geometrili olarak temsil edilir. Bunların birbiri ile eşleştirilmesi için eşik mesafesi değerinin yanında doğrultu, kıvrımlılık veya uzunluk gibi ölçüler de kullanılmaktadır (Hacı & Gökğöz, 2019; Ruiz-Lendínez vd., 2011). Bu ölçüler kullanılarak örneğin iki farklı veri setinde yer alan yollar eşleştirilebilir veya birinde yer alan fakat diğerinde eksik olan bazı bilgiler tamamlanarak zenginleştirme yapılabilir. Bunun dışında, çizgilerin tampon bölgelerine dayalı çeşitli eşleştirme yöntemleri de vardır (Fu & Wu, 2008).

2.2.1.3 Çokgen Geometri İle Temsil Edilen Detaylar

Bina, göl, parsel veya yeşil alan gibi gerçek dünya nesneleri, genel olarak çokgen geometrili olarak temsil edilirler. Farklı veri setlerindeki çokgen geometrili detayların eşleştirilmesi için en çok kullanılan yöntemlerden biri alan kesişimidir (Fan vd., 2016). İki detay geometrik olarak belirli bir oranda kesişiyorsa bu detaylar birbirleriyle eşleştirilebilir. Bunun yanında benzerlik gibi çeşitli ölçülerin kullanıldığı farklı çokgen eşleştirme yöntemleri de literatürde yer almaktadır (Ai vd., 2013; Kim vd., 2010; Mitchell, 2010; Xu vd., 2017).

2.2.1.4 Farklı Geometriler İle Temsil Edilen Detaylar

Farklı iki geometriye sahip detayların birbirleriyle eşleştirilmesi durumudur. Aynı bina nesnesine ilişkin büyük ölçekteki çokgen geometrili detay ile daha küçük ölçekteki nokta geometrili detayın eşleştirilmesi bu eşleştirme türüne bir örnektir. Bu eşleştirme için çokgen içindeki nokta veya çokgenden belirli bir eşik mesafesinde yer alan nokta şeklinde ölçüler kullanılabilir. Bunun yanında nokta geometrili detaylar çizgi ve alan geometrili detayların eşlenmesi için yardımcı olarak da (Örneğin, çizgilerin düğüm noktalarında veya çokgenlerin ağırlık merkezlerinde) kullanılabilir.

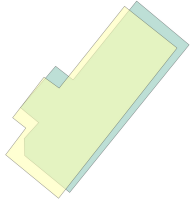
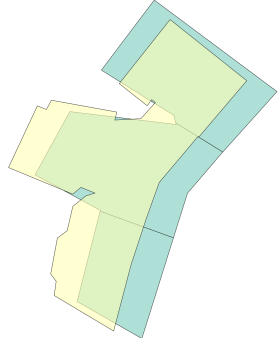
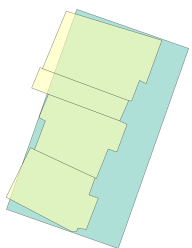
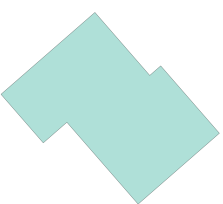
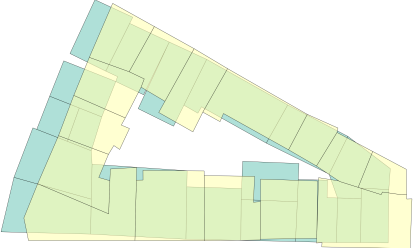
2.2.2 Semantik Boyut

Mekansal nesnelerin sınıfı ve özneteliği gibi anlam ifade eden kısımları o verilerin semantik boyutunu ortaya koyar. Verilerin bu anlam ifade eden kısımlarına göre eşlenmesi işlemine de semantik eşleştirme denir. Verilerin sadece isimlerinin kullanılarak sözdisimsel olarak eşleştirildiği yöntem de bu gruba dahil edilebilir. Genelde semantik benzerlik ölçülerinin kullanıldığı semantik eşleştirme için literatürde çeşitli çalışmalar yer almaktadır (Guarino, 1997; Li vd., 2012; Renteria-Agualimpia & Levashkin, 2011; Ulutaş vd., 2016). Ontolojiler verilerin anlamsal olarak tanımlandığı veri yapıları olduğundan ontoloji eşleştirme çalışmaları da yine semantik boyut altında incelenebilir (Chaabane & Jaziri, 2018; Du, 2015; Varanka & Cheatham, 2016; Volz, 2005).

Bununla birlikte, mekansal verilerin entegrasyonu için farklı veri setlerindeki ilişki türlerinin her iki boyutta da düşünülmesi gerekmektedir. Farklı ölçeklerdeki mekansal detaylar arasındaki ilişkiler, bu verilerin doğası gereği karmaşık olabilmektedir. Büyük ölçeklerde, bir detay geometrik olarak mümkün olduğunca ayrıntılı ve gerçeğine daha benzer şekilde temsil edilir. Ölçek küçüldüğünde ise, uyumsuzlukların ortaya çıkması halinde genelleştirme yapılması gerekir. Bunun sonucunda, daha az ayrıntıya sahip temsiller (örneğin, çokgen geometrili (boşluksuz) bir binanın dış sınır çizgisi

basitleştirilebilir veya bu bina diğer binalarla birleştirilebilir) veya tamamen farklı geometriler (örneğin, çokgen olarak temsil edilen bir bina noktaya dönüştürülebilir) ortaya çıkabilir. Bazen de detayın elenmesi (modelden tamamen çıkarılması) gerekebilir. Sonuç olarak, farklı ölçeklerdeki bina detayları arasında altı tür ilişki ortaya çıkabilir. İki farklı veri setinde yer alan binalar arasındaki olası ilişkiler Tablo 2.1’de gösterilmektedir. Eşleşen binalar bire-bir (1:1), bire-çok (1:n), çok-bir (n:1) ve çok-bir (n:m) ilişki türlerine sahip olabilir. Bunun dışında, veri setlerinden birinde bir bina için diğer veri setinde herhangi bir karşılık olmaması durumunda bire-hiç (1:0) veya hiç-bir (0:1) ilişki türleri de ortaya çıkmaktadır. Veri setleri arasında ortaya çıkan bu tür ilişkiler veri modelleme alanında kardinalite (ilişki derecesi veya ilişki sayısı) olarak adlandırılır (A. Stadler vd., 2009).

Tablo 2.1 Farklı iki veri setinde yer alan bina detayları arasındaki olası ilişkiler (Fan, Zipf, Fu ve Neis (2014)’ten uyarlanmıştır)

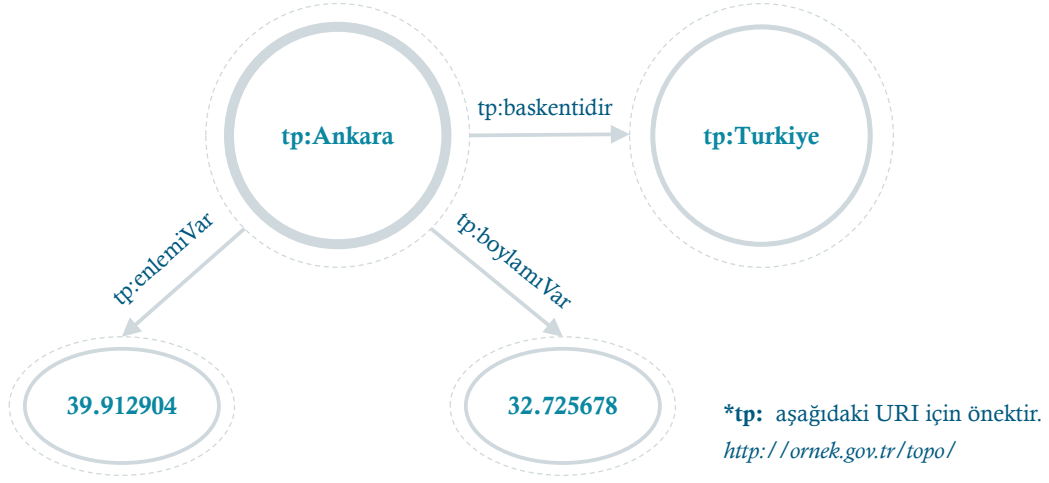
İlişki Türü	1:1	1:0	1:n
Gösterimi			
İlişki Türü	n:1	0:1	n:m
Gösterimi			

3.1 Semantik Web

Semantik web 2000’li yılların başında World Wide Web (WWW)’in kurucusu olan Tim Berners-Lee tarafından mevcut web’in bir uzantısı olarak ortaya konmuştur (Berners-Lee vd., 2001). Dokümanlar web’inden veri web’ine geçiş olarak da adlandırılan bu süreçte mekansal bilimler de dahil olmak üzere birçok bilim dalı verilerini semantik web’e uygun olarak sunmaya başlamıştır. Bölüm 1.1’de belirtildiği üzere semantik web’te veriler, belirli kurallar çerçevesinde tanımlandığında bağlantılı veri adını almaktadır.

3.1.1 Bağlantılı Veriler

Temel prensibi makineler için de verinin anlamının tanımlandığı ve buradan mantıksal çıkarımların yapılabildiği bir ortam sunmak olan semantik web, yapıtaşı olarak özne-yüklem-nesne şeklindeki veri yapısını benimseyen RDF’yi kullanır (Cyganiak vd., 2014). Sözü geçen bu özne-yüklem-nesne yapısına üçlü (triple) de denir. Örneğin, “Ankara Türkiye’nin başkentidir.” cümlesinde Ankara özne, Türkiye nesne, başkentidir ifadesi ise özne ile nesne arasındaki ilişkiyi ifade eden yüklem olmaktadır (Şekil 3.1). Yukarıda yer alan cümle, tam olarak semantik web’de tanımlandığında ifade edilmemiş arka plan bilgilerini de makineler için erişilebilir kılacaktır. Örneğin, ‘başkent’ mantıksal tanımlamasının ilişkileriyle beraber semantik web’te yapılmasıyla birlikte makineler; Ankara’nın bir şehir olduğunu, Türkiye’nin bir ülke olduğunu, her ülkenin sadece bir başkente sahip olduğunu ve herhangi başka bir şehrin Türkiye’nin başkenti olmadığı bilgilerini kolaylıkla çıkarım yoluyla bulabilir.



Özne	Yüklem	Nesne
http://ornek.gov.tr/topo/ Ankara	http://ornek.gov.tr/topo/ baskentidir	tp: Türkiye
http://ornek.gov.tr/topo/ Ankara	http://ornek.gov.tr/topo/ enlemiVar	39.912904
http://ornek.gov.tr/topo/ Ankara	http://ornek.gov.tr/topo/ boylamıVar	32.725678

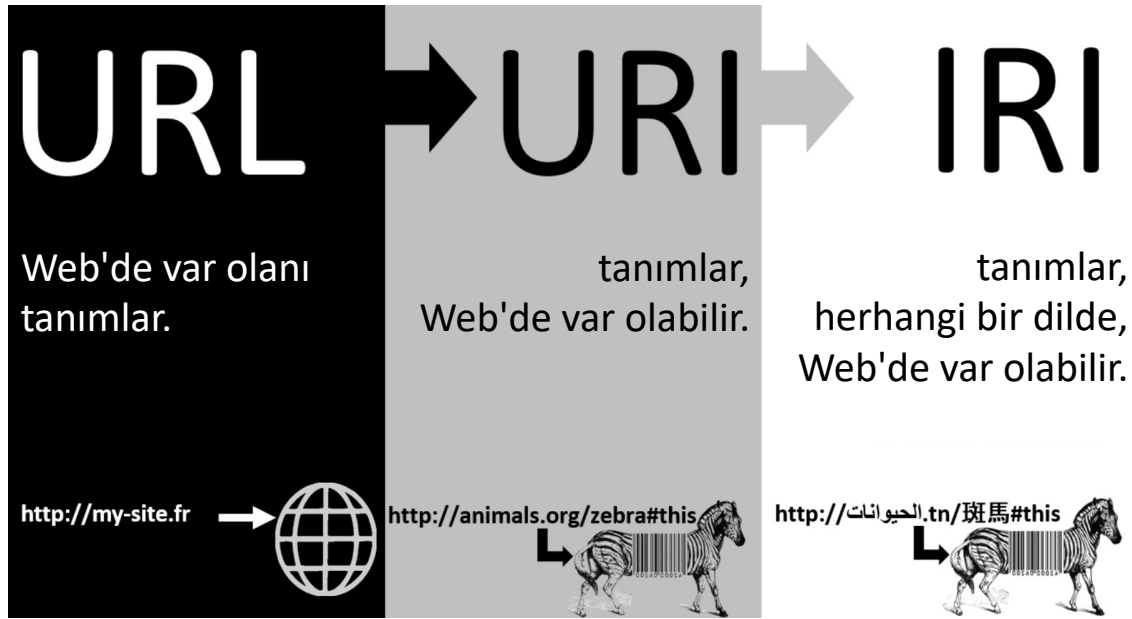
Şekil 3.1 Örnek bir RDF yapısının grafik ve tablo şeklinde gösterimi Memduhoglu ve Basaraner (2017)’den uyarlanarak)

Genişletilebilir İşaretleme Dili (XML), Turtle, Ntriples veya Bağlantılı Veri için JavaScript Object Notasyonu (JSON-LD) gibi metin temelli formatlarla tanımlanan RDF, bir alana ait sınıf ve ilişkileri sözlükler veya ontolojiler şeklinde modellemek için kullanılır. Aşağıdaki kod bloğunda RDF/XML formatında bir üçlü örneği gösterilmiştir.

```
<rdf:Description rdf:about="ozne">
  <predicate rdf:resource="yuklem" />
  <predicate> nesne veya deger </predicate>
</rdf:Description>
```

XML standardı gibi RDF modeli de tek başına kaynakları tanımlamak için bir kısıt mekanizması ortaya koymaz (Decker vd., 2000). Bu eksiklik, RDF Şeması (RDFS) olarak adlandırılan, RDF tanımlayıcı sözlük kullanılarak giderilir. RDFS, RDF kaynaklarını tanımlamak için bir dizi sınıf ve özellikten oluşur. Örneğin, rdfs:Class yeni bir sınıf tanımlamak için kullanılırken, rdfs:subClassOf hiyerarşik olarak alt sınıf tanımlamak için kullanılır (Kettouch, 2017). Bu şekilde yapılan tanımlamalarla belirli alanlara ait taksonomiler, sözlükler veya ontolojiler ortaya konabilir. Bununla birlikte, RDF veri modeli dilin veya sabit değerlerin veri türünün belirtilmesine de izin verir.

Şekil 3.1’de görüldüğü üzere RDF’de kaynakları tanımlamak için URI’ler kullanılmaktadır. URI, soyut veya fiziksel bir kaynağı tanımlamak için kompakt bir karakter dizisidir (Berners-Lee vd., 1998). Bir URI bir ismi, ağ üzerindeki bir konumu veya ikisini birden tanımlayabilir. Şekil 3.2 Tekbiçimli Kaynak Konumlayıcı (URL), URI ve Uluslararasılaştırılmış Kaynak Tanımlayıcı (IRI) arasındaki farklılığı göstermektedir. URL kaynakların web’de yer alan ağ konumlarına olan erişime öncelik sağlayarak isimlerini tanımlar. URI de URL gibi web’te yer alan kaynakların isimlerini tanımlamaya yarar fakat ağ konumları öncelik değildir. IRI ise URI’nin Unicode karakterleri ile genişletilerek farklı dillerde kullanılmasını sağlar.



Şekil 3.2 URL, URI ve IRI arasındaki farklar (Gandon (2014)’ten düzenlenerek)

Semantik web’te amaç verileri makineler tarafından da anlaşılır kılmaktır. Aynı zamanda verilerin insanlar tarafından da anlaşılması gerekmektedir. Bu tür bir problemin ortaya çıkmaması için başvurulabilir (dereferenceable) URI teknolojisi kullanılmaktadır. Bu yöntem, tanımlanan bir kaynağın kopyasını veya temsilini Internet protokolleriyle (Örn. HTTP) elde etmek için kullanılan bir kaynak alma mekanizması sağlamaktadır (Kettouch, 2017; Yaghouti vd., 2015).

RDF modeli, URI’leri daha kısa bir biçimde yazmak ve yeniden kullanılabilir örnekleri tanımlamak için CURIE (veya Kompakt URI) kısaltma stilini kullanır. Kullanılacak URI’ler ve örnekleri -dokümanın formatına göre değişmekle birlikte- genel olarak dosyanın başında tanımlanır. Dokümanın geri kalan kısmında URI öneki ve benzersiz kısımlar kullanılarak daha kompakt bir doküman ortaya konur. Aşağıdaki kod bloğunda yukarıdaki stile uygun Turtle formatında örnek bir RDF parçası yer almaktadır.

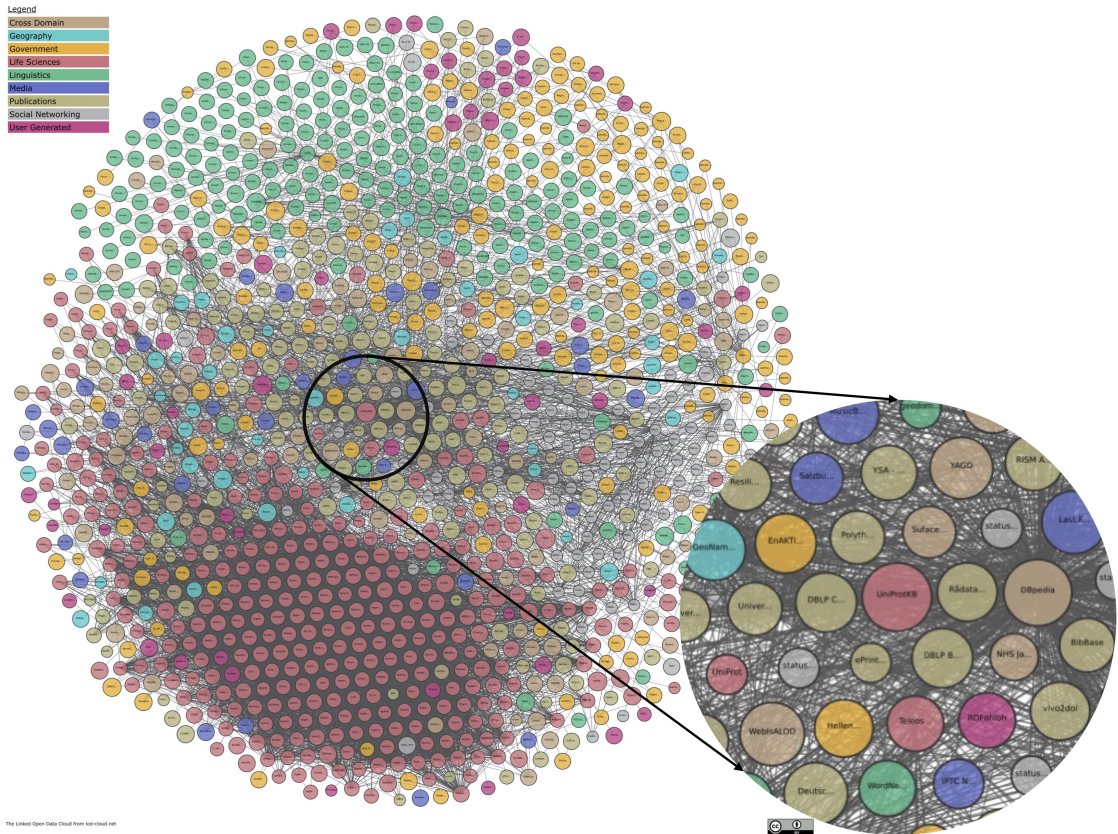
```

@prefix owl: <http://www.w3.org/2002/07/owl#> .
@prefix rdf: <http://www.w3.org/1999/02/22-rdf-syntax-ns#> .
@prefix rdfs: <http://www.w3.org/2000/01/rdf-schema#> .
@prefix uho: <http://ulusalharita.tr/ontology#> .

uho:Hydrography rdf:type owl:Class ;
                rdfs:subClassOf uho:TopographicFeature ;
                rdfs:label "Hidrografiya"@tr .

```

Web'te yer alan bağlantılı veriler, ontolojiler ve sözlüklerden oluşan veri ağı, Bağlantılı Açık Veri (LOD) Bulutu (LOD, 2021) üzerinden izlenebildiği kadarıyla her geçen gün büyümektedir. 2007'de LOD Bulutunda 12 veri seti varken 2020'de bu sayı 1269'a çıkmıştır. LOD Bulutundaki alanlarına göre renklendirilmiş güncel bağlantılı veriler Şekil 3.3'te görülebilir. Örneğin, "yararlı bilgi tabanı (knowledge base)" olarak da tanımlanan ve web üzerinde en çok kullanılan kapsamlı bağlantılı veri kümesinden biri olan DBpedia, Wikipedia'dan yapılandırılmış bilgileri çıkararak yayınlamaktadır (DBP, 2021).

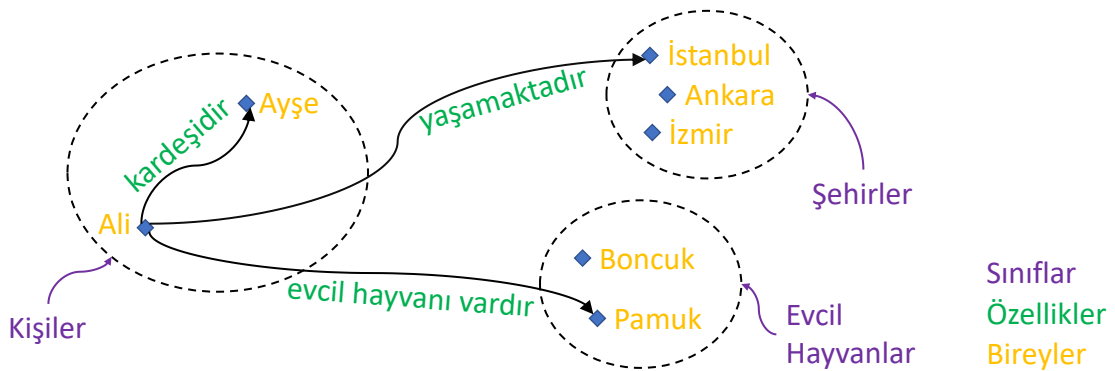


Şekil 3.3 Linked Open Data Cloud üzerinde yer alan bağlantılı veri ağı (LOD, 2021)

3.1.2 Ontolojiler

Kelime anlamı olarak ontoloji varlık bilimi demektir. Eskiden beri felsefe alanında kullanılan bu kavram son yıllarda bilgisayar bilimlerinde de kullanılmaya başlanmıştır. RDF'yi temel olarak kullanan ontolojiler belirli bir alanla ilgili sınıflar, kavramlar ve ilişkileri tanımlamak için kullanılır. Gruber (1993) ontolojiyi, “paylaşılan bir kavramsallaştırmanın biçimsel, açık bir belirtimi” olarak tanımlar. Burada “kavramsallaştırma”, gerçek dünya nesnesinin soyut modelini ve belirli bir alana ait olan ilgili ilişkileri ifade eder. “Açık”, ontolojide kullanılan ve açıkça tanımlanan bilgi anlamına gelir. “Biçimsel”, ontolojinin makinelerin anlayabileceği belirli kurallar altında tanımlanması gerektiğini ifade eder. Son olarak, “Paylaşılan”, ontolojide kullanılan bilgilerin alan uzmanları tarafından bir fikir birliğiyle oluşturulması gerektiği anlamına gelir (Laurini, 2017).

Her ne kadar RDFS yapısı ile bir ontoloji ortaya konabilir olsa da kapsamlı sınıf ve ilişkileri tanımlamak için RDFS yeterli değildir. Bu yüzden ontolojilerin büyük bir çoğunluğu OWL ile tanımlanmaktadır (Bechhofer vd., 2004). RDF'ye dayalı olmasına rağmen, OWL ile daha gelişmiş sınıf tanımları ve kısıtlamaları yapılabilir. Örneğin, Ankara'nın Türkiye'nin tek başkenti olduğu ve başka hiçbir şehrin başkent olamayacağı OWL'nin Ters İşlevsel Özellik (Inverse Functional Property) adlı özelliği kullanılarak tanımlanabilir. OWL, DL'yi kullanır ve bu mantıksal modele dayalı kararlar verir. OWL ile oluşturulan belgeler, web üzerinden yayınlanabilen ontolojiler olarak da bilinir. OWL'nin üç alt dili vardır: OWL Lite, OWL DL ve OWL Full. OWL Lite, bir sınıflandırma hiyerarşisi ve basit kısıtlamalar için dil yapıları sağlar. OWL DL, hesaplamanın eksiksizliğini ve karar verilebilirliğini korurken maksimum ifadeyi destekler. OWL Full, hesaplama garantisi olmaksızın RDF için maksimum ifade ve daha fazla sözdizimsel özgürlük sağlar. Tanımdan da anlaşılacağı gibi OWL DL, ontolojilerin oluşturulması için en yaygın alt dildir. Bu dil en temel biçimde sınıflar, özellikler ve bireyler olarak tanımlanabilir (Şekil 3.4).



Şekil 3.4 OWL'nin sınıflar, özellikler ve bireylerden oluşan temel bileşenleri (Horridge vd. (2011)'den uyarlanmıştır)

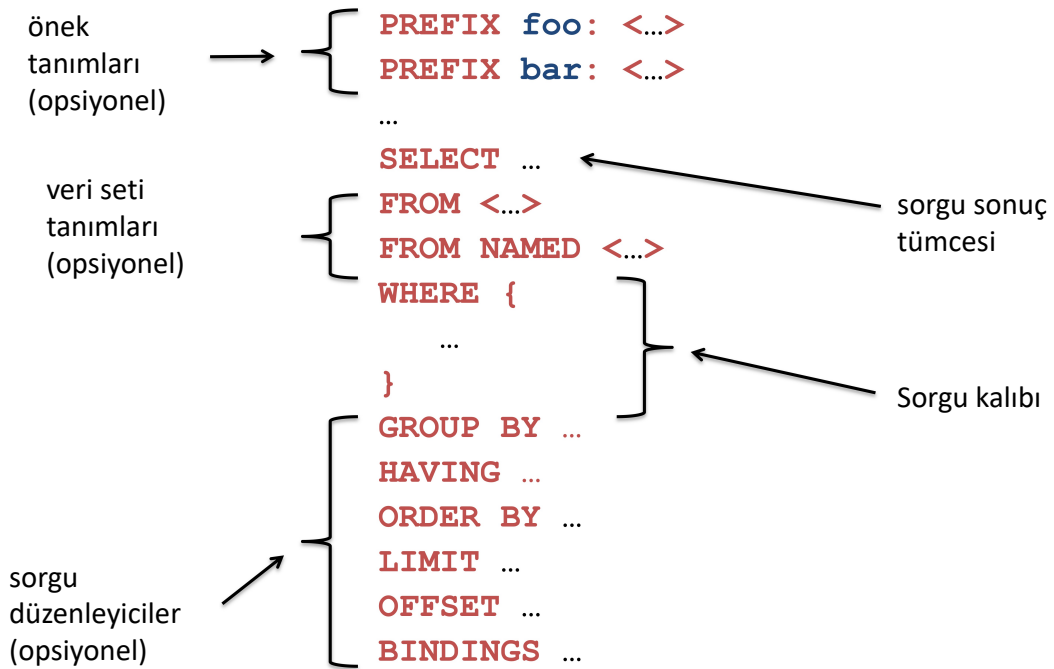
Sınıf, ortak özellikleri paylaşan bireyler kümesidir. Özellikler, bireyler veya veri değerleri arasındaki ilişkileri tanımlar. Bireyler, sınıfların ve özelliklerin örnekleridir. Örneğin, Şekil 3.4'te görüldüğü gibi, İngiltere, Türkiye ve ABD, ülke sınıfına ait bireylerdir ve “yaşamaktadır”, ilişkiyi ifade eden özelliğe karşılık gelir. OWL, programlama dilinden ziyade bilgiyi temsil eden tanımsal bir dildir. Sınıflar ve sınıflarla ilgili ifadeler kümesi, Terminolojik Kutu veya Tbox olarak bilinir. Bireyler ve bunlarla ilgili gerçekler kümesi ise, İddia (Assertion) Kutusu veya Abox olarak bilinir. OWL ile ontoloji oluşturmak veya oluşturulmuş mevcut ontolojileri anlamak için OWL bileşenlerine biraz daha ayrıntılı bakmak gerekir. Temel olarak OWL'nin sınıflar, özellikler (properties), bireyler (individuals), veri tipleri ve açıklamalar (annotations) olmak üzere beş farklı bileşenden oluştuğu söylenebilir. Sınıflar, yukarıda da açıklandığı üzere belirli ortak özelliklere sahip bireylerin kümелendiği tanımlardır. İlişkiler ikiye ayrılır: Bireyler arasındaki özellikleri tanımlayan nesne özellikleri (object properties) ve bireylerin veri tipleriyle olan ilişkilerini tanımlayan veri özellikleri (datatype properties). Şekil 3.1'deki örneğe bakacak olursak “tp:başkentidir” bir nesne özelliği iken, “tp:enlemiVar” bir veri özelliğini temsil etmektedir. Veri tipleri de bireyler için tanımlanan verilerin karakter (string), tam sayı (integer) vb. tipleri ifade eder. Örneğin, “tp:enlemiVar” veri ilişkisi için “ondalıklı sayı” (double) veri tipi tanımlanabilir. Son olarak ontolojinin adı, kim tarafından oluşturulduğu, versiyonu, oluşturulma yılı vb. tanımlar, açıklamalar olarak ifade edilmektedir (Bechhofer vd., 2004). Kendinden tanımlı bileşenlere sahip çok sayıda sistemden oluşan OWL ile ontolojilerin veri entegrasyonunda kullanımının otomatikleştirilmesi mümkündür. Bu bağlamda OWL, semantik düzeyde mekansal bilginin birlikte çalışabilirliğini destekler.

Ontolojilerin oluşturulması, düzenlenmesi ve diğer pek çok işlemin yapılabilmesine olanak sağlayan Protégé, yaygın olarak kullanılan bir açık kaynaklı yazılımdır (Musen, 2015). Protégé, OWL ve RDF ontolojilerinin yüklenmesini, düzenlenmesini ve görselleştirilmesini destekler. Ayrıca, Semantik Web Kural Dili (SWRL) dosyalarının içe aktarılmasına ve çalıştırılmasına izin verir. Protégé, HermiT, FaCT++ ve Pellet gibi çeşitli çıkarım motorlarını ve eklentileri de içerir. Bu çıkarım motorlarıyla, ontolojiler biçiminde var olan verilerden bilgi çıkarımı yapılabilir. Masaüstü Protégé yazılımının WebProtege adlı bir web sürümü de vardır. Bu sürüm ile projeler, WebProtege'ye yüklenebilir ve çevrimiçi olarak kullanılabilir. Ontoloji düzenleyicilerine ek olarak, semantik web'i desteklemek için birçok araç, kitaplık ve yazılım paketi bulunmaktadır. Örneğin Apache Jena, semantik web uygulamaları için ücretsiz ve açık kaynaklı bir Java geliştirme platformu sunar (Jena, 2021). Apache Jena içerisinde, SPARQL sunucusu olarak Fuseki ve RDF'leri depolayan üçlü deposu olarak TDB isimli yazılım bileşenlerini içerir. Yukarıda bahsedilen yazılımların tümü açık kaynaklı olduğundan, güçlü topluluk desteğine ve web'te ücretsiz olarak bulunabilen iyi yapılandırılmış

eğitim ve bilgilendirme dokümanlarına sahiptirler.

3.1.3 SPARQL Protokolü ve RDF Sorgu Dili (SPARQL)

SPARQL Protokolü ve RDF Sorgu Dili (SPARQL) World Wide Web Consortium (W3C) tarafından geliştirilen semantik web'in sorgulama dilidir (Aranda vd., 2013). İlişkisel veri tabanlarını sorgulamak için kullanılan Yapısal Sorgu Dili'nin (SQL) semantik web'teki muadilidir. SELECT, FROM, WHERE gibi ifadeleri kullanarak SQL benzeri bir yapı sunar (Şekil 3.5). Fakat, RDF ile tanımlanan üçlülere sorgulamak için kullanıldığından ilişkisel veri modeli yerine bir çizge (düğüm-kenar-düğüm) modeli üzerinde çalışır.



Şekil 3.5 Bir SPARQL sorgusunun yapısı (Feigenbaum (2008)'dan düzenlenerek)

Aşağıdaki kod bloğunda Şekil 3.1'de gösterilen RDF için oluşturulmuş örnek bir SPARQL sorgusu yer almaktadır.

```
@Prefix tp: <http://ornek.gov.tr/topo/>
```

```
SELECT ?detay
WHERE {
  ?detay tp:baskentidir tp:Turkiye . }
```

Yukarıdaki sorguda “?detay” değişkeniyle Türkiye'nin başkenti sorgulanmıştır ve “tp:Ankara” cevabının alınması beklenir.

3.1.4 Semantik Web Kural Dili (SWRL)

Güçlü bir çıkarım yeteneğine sahip olmasına rağmen OWL, karmaşık çıkarımları gerçekleştirme konusunda hala sınırlıdır. Bu sınırlılığı örneklemek amacıyla yaygın olarak amca örneği kullanılır. Bu örnekte, eğer A, B'nin erkek kardeşiysen ve B, C'nin babasıysa, o zaman A, C'nin amcasıdır. A ile C arasındaki ilişkiyi içeren böyle bir bilgi çıkarımını OWL tek başına gerçekleştiremez. Bu nedenle, ontoloji dillerinin ifade gücünü belirli kurallarla yükseltmek adına OWL ve RuleML isimli kural dilini birleştiren SWRL geliştirilmiştir (Horrocks vd., 2004; Huang, 2020). SWRL kuralları öncül (antecedent) ve netice (consequent) olmak üzere iki kısımdan oluşur.

öncül (antecedent) $\Rightarrow$ netice (consequent)

Öncül ve netice kısımları da atom denilen daha küçük parçalardan oluşurlar. Öncül kısımdaki atomların hepsinin mantıksal olarak geçerli olduğu durumda netice kısmı da geçerli olur. Yukarıda sözü geçen amca örneğini gerçekleştirebilmek için aşağıda yer alan SWRL kuralının uygulanması gerekmektedir.

erkekKardesidir(?A,?B) $\wedge$ babasidir(?B,?C) $\Rightarrow$ amcasidir(?A,?C)

SWRL açık dünya varsayımını benimser ve sadece tekdüze (monotonik) çıkarımları destekler. Açık dünya varsayımı, doğruluğu veya yanlışlığı henüz bilinmeyen bir bilgi için kesin hüküm vermez (kapalı dünya varsayımında, bir bilgi veride yer almıyorsa yanlış olarak varsayılır). Kısacası açık dünya varsayımında doğru bilgi de yanlış bilgi de açıkça tanımlanmalıdır. Bu yüzden SWRL olumsuzluk ifadelerini desteklemez. Örneğin, bina türü belirtilmeyen detayları farklı bir işaretle gösterme durumu SWRL ile gerçekleştirilemez. Bu tür bazı kısıtları belirli bir dereceye kadar ifade etmek için daha sonradan ortaya çıkan SPARQL Çıkarım Notasyonu (SPIN) ve Şekil Kısıtlama Dili (SHACL) yaklaşımları kullanılabilir.

3.2 Mekansal Veriler İçin Semantik Web

Mekansal bilgi topluluğunda, semantik web kavramının ortaya çıkmasından bu yana özellikle veri entegrasyonu, uyumlaştırma, bilgi paylaşımı ve yeniden kullanım için ontolojiler ve diğer semantik web teknolojileri kullanılmaktadır (Hakimpour, 2003; Kashyap & Sheth, 1997; Kavouras & Kokla, 2002; Kuhn, 2001; Smith & Mark, 1998; Uitermark vd., 1999; van den Brink vd., 2017). Semantik web'den yararlanmak için, her geçen gün daha fazla UHK mekansal verilerini bağlantılı veriler olarak yayınlamaktadır (Regalia vd., 2018). Bu bağlamda, Büyük Britanya'nın ulusal haritacılık kurumu olan Ordnance Survey, verilerini bağlantılı veri olarak yayınlamak için, on yıldan fazla bir süredir semantik web teknolojilerini

araştırmakta ve kullanmaktadır (Goodwin vd., 2008). Ulusal mekansal veri tedarikçisi olarak Amerikan Jeolojik Araştırmalar (U.S. Geological Survey) kurumu da semantik web teknolojilerine odaklanmış ve kendi ulusal haritalarından bağlantılı veriler geliştirmiştir (Usery & Varanka, 2012; Varanka vd., 2015). Ordnance Survey İrlanda da kendi coğrafi verilerini bir bağlantılı veri platformu aracılığıyla yayınlamıştır (Debruyne, Meehan, vd., 2017). Ayrıca, İspanyol ulusal veri setleri İdari birimler, Hidrografya, İstatistiksel birimler ve Meteoroloji gibi Avrupa Birliği Mekansal Bilgi Altyapısı (INSPIRE) temaları, veri kaynağı, çözünürlük ve dil çeşitliliği gibi nedenlerden kaynaklanan farklı heterojenlik sorunlarının üstesinden gelmek için bağlantılı veriler olarak oluşturulmuş, entegre edilmiş ve yayınlanmıştır (Vilches-Blázquez vd., 2014). Ek olarak, Avrupa'dan dört UHK (Finlandiya, Hollanda, Norveç ve İspanya), ulusal mekansal veri altyapıları için veri uyumlaştırma ve entegrasyonda bağlantılı veri teknolojisini kullanan Avrupa Açık Konum Servisleri adlı bir proje yürütmüştür (Ronzhin vd., 2019). Ulusal kurumların yanında VGI tarafında da mekansal verilerin bağlantılı olarak sunulması konusunda gelişmeler yaşanmıştır. GeoNames (GN, 2021) ve LinkedGeoData (C. Stadler vd., 2012) bu gelişmelere ilişkin en önemli iki örnektir. GeoNames, tüm dünyada konum bilgileriyle birlikte 6 milyondan fazla yer ismi ve eş anlamlı isimleri içermekte ve web'de bir bilgi tabanı olarak hizmet etmektedir. LinkedGeoData ise, OSM'den toplanan bilgileri kullanmakta ve bunları web'de bağlantılı veriler olarak sunmaktadır. Semantik web'de ontolojiler aracılığıyla mekansal verileri yayınlamanın ana hedeflerinden biri, makineler tarafından yorumlanabilen mantıksal çıkarımları mümkün kılmak için mekansal detaylar arasındaki kapsamlı ilişkilerin modellenmesidir. Bu durum, kullanıcıların verilerini çeşitli mekansal verilerle sorunsuz bir şekilde entegre etmelerine ve daha akıllı uygulamalar ortaya koymalarına olanak tanır.

Mekansal ontolojiler, mekansal verileri bağlantılı veriler olarak yayınlamanın önemli bir parçasıdır. Literatürde, araştırmacılar mekansal ontolojileri; topografik detayların modellenmesi (Varanka & Usery, 2015), entegrasyon (Wiemann & Bernard, 2016), zenginleştirme (Debruyne, McGlinn, vd., 2017; Hu & Janowicz, 2016), paylaşım (Claramunt, 2020) ve mekansal veriler üzerinde muhakeme (Thomson, 2009) gibi konularda kullanmıştır. Bir mekansal ontoloji oluşturma'nın ilk adımı, ontolojide hangi kavramların, ilişkilerin ve özelliklerin tanımlanacağına karar vermektir (Goodwin vd., 2008). Semantik web'te bu ilişkileri tanımlamak için iki yol vardır: (1) İlişkilerin manuel olarak yeniden tanımlanması veya (2) Friend Of A Friend (FOAF, 2021), Simple Knowledge Organisation System (SKOS, 2021), Dublin Core (DC, 2021) veya DOLCE+DnS Ultralite (DUL, 2021) gibi mevcut sözlüklerden veya ontolojilerden yararlanılması. İkinci durumda sözü geçen ontoloji veya sözlükler sıkça kullanılan genel sınıf ve ilişkileri tanımlarlar. Mevcut bir ontolojiyi veya sözlüğü sıfırdan

oluşturmak yerine yeniden kullanmak semantik web'in temel kavramlarından biridir. Mevcut ve topluluk destekli ontolojileri kullanmak, kullanıcılar için her seferinde ontolojileri yeniden keşfetmek zorunda olmadıkları için avantajlar sağlar. Ayrıca bir ontoloji tasarlamak ve diğer disiplinler tarafından kullanılmasını sağlamak zor bir iştir. Bu nedenle, Ontoloji Tasarım Kalıpları, çeşitli disiplinler tarafından kullanılmak üzere basit ve küçük ontolojiler oluşturmak için kullanılır (Gangemi & Presutti, 2009). Bu bağlamda mekansal ontolojilerin oluşturulması için de çeşitli tasarım kalıpları ortaya konmuştur (Carral vd., 2013; Hu vd., 2013; Sinha vd., 2014).

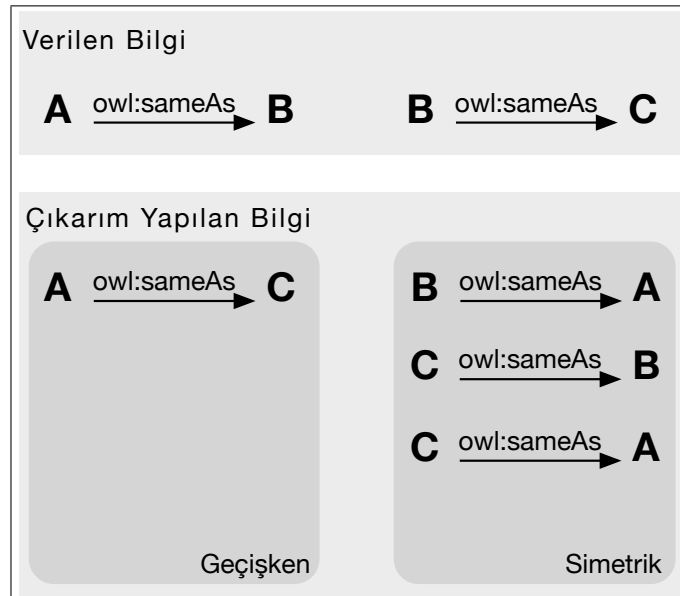
Mekansal bağlantılı veri uygulamalarının artan kullanımıyla, mekansal semantik web standartlarına yönelik doğal bir talep ortaya çıkmıştır. Bu bağlamda OGC, mekansal bağlantılı verileri sorgulamak için GeoSPARQL adlı bir standart oluşturmak için çaba sarf etmiştir (Perry & Herring, 2012). GeoSPARQL, SPARQL'nin genişletilmiş bir sürümü olarak mekansal bağlantılı veri setlerini sorgulamak, güncellemek ve işlemek için kullanılır. Bu standardın yanı sıra, mekansal ontolojiler için bir temel olarak kullanılmak üzere aynı adı taşıyan bir sözlük de yayınlanmıştır. Bu bağlamda, popüler bir mekansal sözlük olan GeoSPARQL 8 Bölge Bağlantı Hesabı (RCC8) ve Açık Mekansal Konsorsiyum (OGC)'un Basit Detaylar (Simple Features) standartlarını topolojik ve mereolojik (parça-bütün) ilişkileri tanımlamak için kullanmaktadır. Ayrıca, geometrilerin kodlanması için İyi-Bilinen Metin (WKT)/Coğrafi İşaretleme Dili (GML) desteğini de içermektedir. GeoSPARQL sözlüğü yayınlandığı günden beri yaygın olarak kullanılmasına rağmen, çok ölçekli verileri veya ölçek tanımlarını desteklememektedir (Eiðsson, 2018; Ronzhin vd., 2019). Ancak, Huang vd. (2018) ve Eiðsson (2018), GeoSPARQL standardına farklı ölçeklerdeki mekansal nesnelerin birbiriyle ilişkilendirilebildiği bir ölçek sözlüğü ekleyerek genişletmişlerdir.

Pratikte, bağlantılı verilerin semantik sorgular (SPARQL) aracılığıyla RDF verilerini almak ve saklamak için RDF depoları (triplestore) kullanılarak yönetilmesi gerekir. Mekansal ontolojiler ve bağlantılı verilerin günden güne artmasıyla birlikte bu mekansal ontolojiler için geliştirilen RDF depoları da çoğalmıştır. Bunların performanslarının test edilmesi ve hangi durumlarda hangi RDF depolarının kullanımının uygun olacağını tespit edilmesi amacıyla RDF depoları üzerine birçok değerlendirme ve kıyaslama çalışması yapılmıştır (Battle & Kolas, 2012; Bellini & Nesi, 2018; Garbis vd., 2013; Huang vd., 2019). Bu çalışmalar sonucunda, mekansal sorguların istemci süresi anlamında nispeten yeterli olduğu ve gün geçtikçe bu yazılımların gelişim gösterdiği ortaya çıkmıştır.

Önemli bir sınırlama olarak, mekansal semantik web henüz CBS kadar olgunlaşmamıştır ve geleneksel coğrafi verileri bağlantılı verilere dönüştürmek ve bunları görüntülemek için yaygın olarak benimsenen araçlara sahip değildir

(Vilches-Blázquez & Saavedra, 2019). Mekansal semantik web özelinde geliştirilmiş araçlar mevcut olmasına rağmen, bunlar çoğunlukla üniversite araştırma projelerine dayandıkları için uzun vadeli bakım ve destek için pek uygun değildir (Regalia vd., 2018).

Diğer taraftan, semantik web, klasik CBS ve ÇTMVT yöntemleriyle kıyaslandığında kaynakların birbirleriyle ilişkilendirilmesi konusunda doğal bir avantaja sahiptir. Semantik web'te aynı kaynakları birbirine bağlamak için en çok kullanılan bağlantı türlerinden biri owl:sameAs ilişkisidir. Yaygın olarak kullanılan diğer bir bağlantı türü, benzer ancak tamamen aynı olmayan kaynakları ilişkilendirmek için kullanılan skos:closeMatch'tir. Bu ikisi aynı zamanda mekansal detaylar için en sık kullanılan bağlantı türleridir. Geçişken ve simetrik adındaki mantıksal karakteristik özellikleri ile birbirlerinden farklılaşırlar. owl:sameAs hem geçişken hem de simetrik özelliklere sahipken, skos:closeMatch yalnızca simetrik özelliğe sahiptir. Geçişken ve simetrik karakteristik özelliklere örnek vermek gerekirse, A, B ve C'nin kaynaklar (veya detaylar) olduğunu düşünelim. A, owl:sameAs ilişkisi ile B'ye bağlandığında ve B, owl:sameAs ilişkisi ile C'ye bağlandığında OWL destekli bir çıkarım motoru, owl:sameAs'ın geçişken özelliğinden dolayı A'nın da C ile ilişkili olduğunu ve simetrik özelliğinden dolayı da B'nin A ile, C'nin B ve A ile ilişkili olduğu çıkarımını yapar (Şekil 3.6).



Şekil 3.6 Geçişken ve simetrik özelliklerin örnek gösterimi

Semantik web'in mantıksal karakteristik özellikleri gibi avantajları kullanılarak ve mekansal detayların kardinal ilişkileri tanımlanarak bunların çoklu temsili, entegrasyonu ve zenginleştirilmesi gerçekleştirilebilir. Sonraki bölümde bununla ilgili yöntem ve uygulamalara yer verilmiştir.

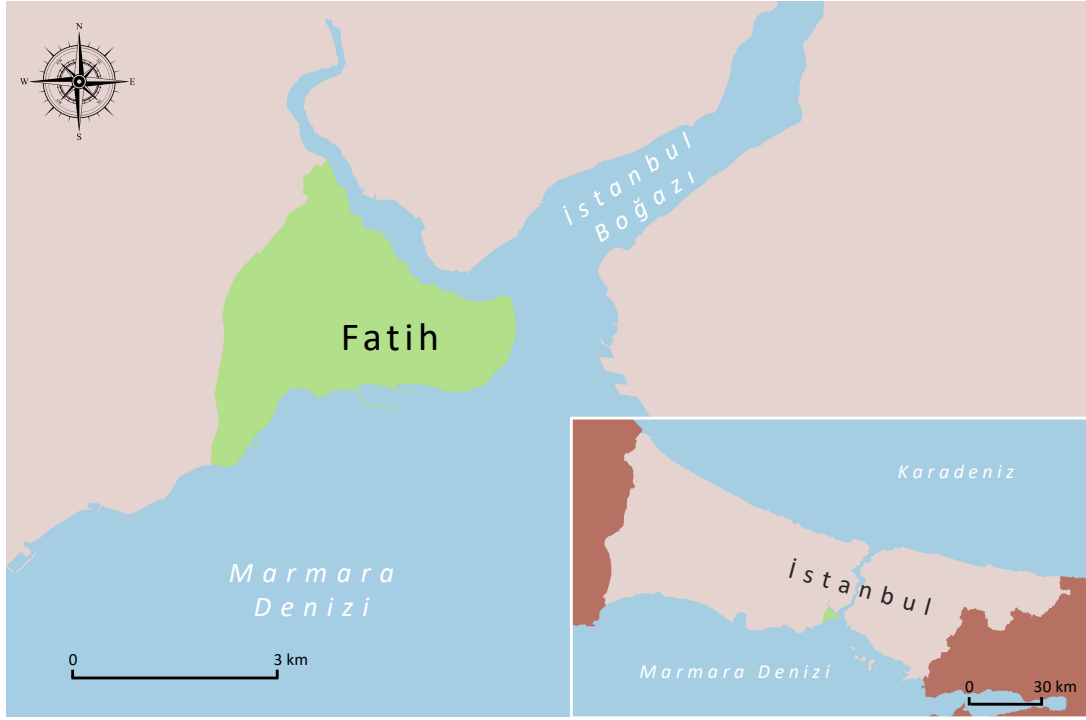
Semantik web'in mekansal veri entegrasyonu ve zenginleştirmesinde sağlayabileceği katkıları ortaya koymak için mevcut teknikleri uygun biçimde bir araya getiren yaklaşımlar geliştirilmiş ve üç farklı bina veri setine dayalı iki uygulama ile bu yaklaşımların işlevsellikleri gösterilmiştir.

4.1 Çalışma Alanı

Çalışma alanı İstanbul'un Fatih ilçesi olarak seçilmiştir (Şekil 4.1). Tarihi yarımadayı içeren bu alan tarihi ve turistik binaların yanı sıra işyerleri ve konut alanlarını da içerdiğinden bina kullanım türleri açısından zengin çeşitliliğe sahiptir. Bunun yanında OSM verilerinin tamlığı (öznitelik) bağlamında da ülkenin diğer bölgelerine kıyasla bir avantaja sahiptir. Yapılan her iki uygulamanın da çalışma alanı aynıdır. Fakat kullanılan veriler farklılık göstermektedir.

4.2 Uygulama 1: Semantik Web Aracılığıyla Mekansal Veri Entegrasyonu

Çok kaynaklı mekansal verilerin artan kullanılabilirliğiyle birlikte, tek bir veri kaynağı kullanılarak cevaplanamayan karmaşık soruları yanıtlamak için maksimum bilgiyi toplamayı amaçlayan mekansal veri entegrasyonu çabaları hız kazanmıştır. Bu amaçla kullanılan ve farklı derecelerde başarı gösteren çeşitli yaklaşımlar olmasına rağmen, özellikle çok ölçekli bina verilerinin entegrasyonu ve zenginleştirilmesi kapsamında semantik web yöntemi ve araçları yeterince test edilmemiştir. Bu bağlamda, uygulamada, çok kaynaklı ve çok ölçekli kentsel bina verileri, alan kesişimine dayalı bir geometrik eşleştirme yöntemi ile eşleştirilmiş, ardından bina detaylarının çok ölçekli temsillerini ve ayrıntılı ilişkilerini tanımlamak için bir mekansal ontoloji geliştirilmiştir. Son olarak, mekansal ontolojideki bazı (uygun) detaylar DBpedia ve YAGO (2021) gibi popüler bilgi tabanlarıyla ilişkilendirilmiştir.

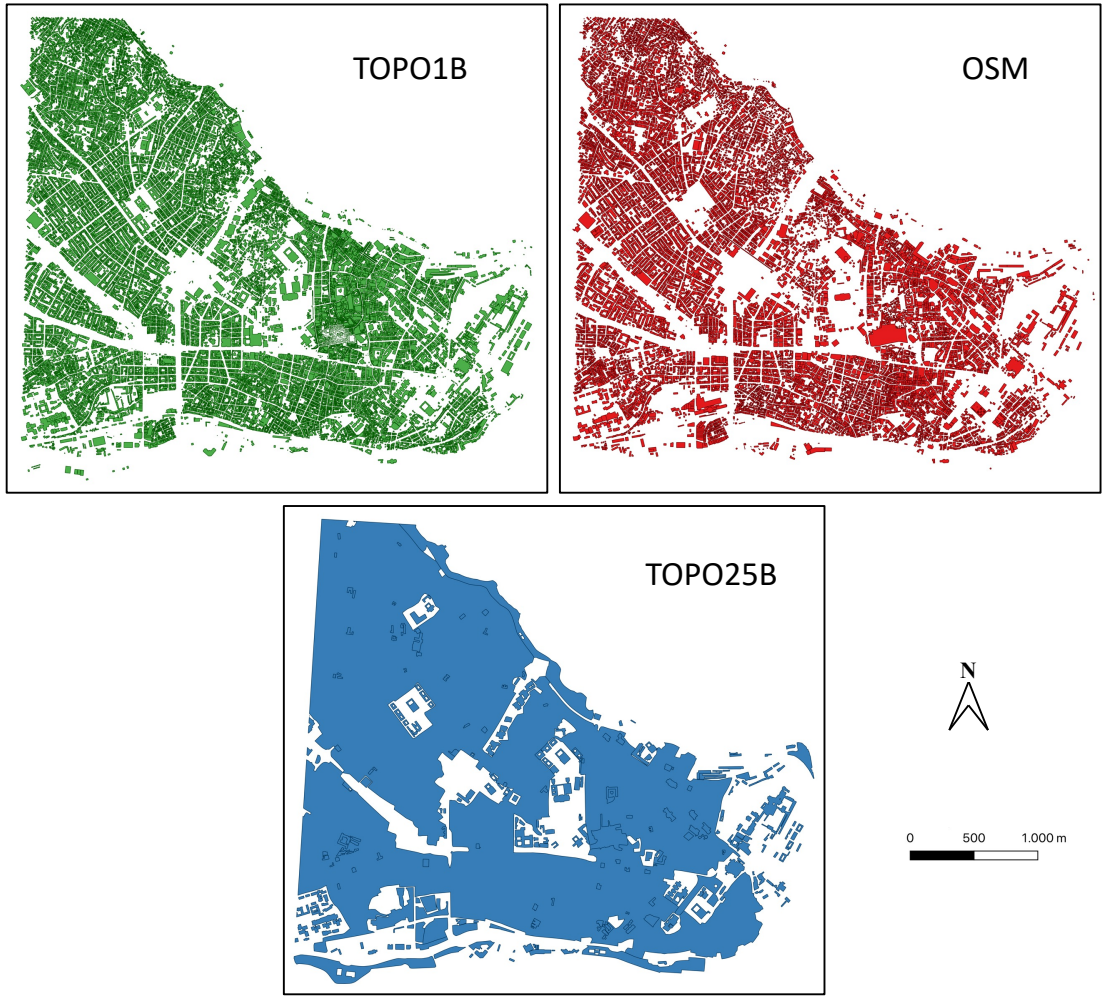


Şekil 4.1 Çalışma alanı

Sonuçların, web haritası üzerinden sunumu için örnek sorular kullanılarak sorgulama ve görselleştirme süreçleri gösterilmiştir. Çalışma ayrıca üç farklı coğrafi veri seti arasındaki karmaşık ilişkileri modellemek için çıkarım yapma gibi semantik web yöntemlerinin avantajlarını da göstermektedir.

4.2.1 Kullanılan Veriler

Bu uygulamada farklı kaynaklardan elde edilen ve farklı ölçeklerde yer alan üç bina veri seti kullanılmıştır (Tablo 4.1). Birincisi İstanbul Büyükşehir Belediyesi sorumluluğunda üretilen Gauss-Krüger Orta Meridyen 30 (ITRF96) projeksiyon koordinat sistemindeki 1:1 000 ölçekli topografik verilerdir (TOPO1B). İkincisi, Coğrafi (WGS84) koordinat sistemindeki OSM verileridir ve homojen bir ölçeğe sahip olmamakla birlikte bu bölge için yaklaşık 1:10 000 veya biraz daha küçük bir ölçeğe sahip olduğu ifade edilebilir (Basaraner, 2020). Bu bağlamda, Olteanu-Raimond vd. (2017) ve Zhou vd. (2018) çalışmalarında OSM'nin yaklaşık 1:10 000 ölçeğine sahip olduğunu belirtmişlerdir. Üçüncüsü, Türkiye'nin UHK'sı olan Harita Genel Müdürlüğü (HGM) tarafından üretilen Universal Transversal Merkator Dilim 35 (WGS84) projeksiyon koordinat sistemindeki 1:25 000 ölçekli topografik verilerdir (TOPO25B). TOPO25B, tüm ülkeyi kapsayan en büyük ölçekli üründür. Bu haritanın üretiminde kullanılan geometrik veriler, hava fotogrametrisi yoluyla toplanır. Bu nedenle, TOPO1B ve TOPO25B arasında doğrudan bir genelleştirme ilişkisi yoktur.



Şekil 4.2 Uygulamada kullanılan veriler

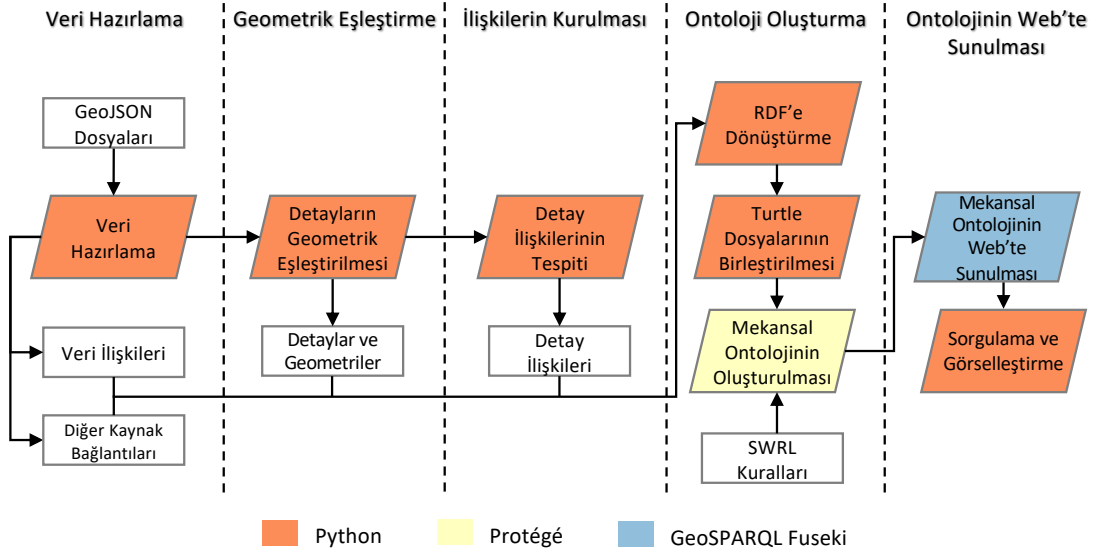
Çalışma alanı kentleşme açısından oldukça yoğun olduğundan, TOPO25B’de çoğu bina, tekli temsile ölçeğin izin vermemesi nedeniyle, geometrilerinin birleştirilmesiyle oluşturulmuş meskun alan çokgenleriyle temsil edilmektedir. Kullanılan veriler Şekil 4.2’de gösterilmiştir. Veri setlerinin isimleri buradan sonra Tablo 4.1’de yer alan kısaltmalarla anılacaktır.

Tablo 4.1 Çalışmada kullanılan veri setleri, kısaltmaları ve detay sayıları

Veri Seti	Veri Seti Kısaltması	Detay Sayıları
1:1 000 ölçekli topografik (resmi)	TOPO1B	33580
~1:10 000 ölçekli OpenStreetMap	OSM	20503
1:25 000 ölçekli topografik (resmi)	TOPO25B	281

4.2.2 Yöntem

Mekansal verilerin entegrasyonu ve zenginleştirilmesi için geometrik eşleştirme ve semantik web temelli bir yaklaşım ortaya koyan bu çalışma beş aşamadan oluşmaktadır: i. veri hazırlama, ii. binaların geometrik eşleştirilmesi, iii.



Şekil 4.3 Genel metodoloji ve çalışmanın uygulama adımları

binalar arasındaki ilişkilerin belirlenmesi, iv. mekansal ontoloji oluşturulması ve v. ontolojinin web üzerinden sunulması. Bu kapsamda ağırlıklı olarak üç farklı yazılım paketi/ortamı kullanılmıştır (Şekil 4.3). İlk üç adım açık kaynaklı paketler aracılığıyla Python ortamında (Python kodları için Ek A'ya bakınız), dördüncü aşama açık kaynaklı Protégé ontoloji düzenleyicisinde ve son aşama ise açık kaynaklı GeoSPARQL Fuseki web sunucusu ve Python paketlerinin birlikte kullanımıyla gerçekleştirilmiştir.

İlk olarak, özel dosya formatlarındaki (shapefile ve geodatabase) mekansal veriler, mekansal yazılım topluluğu tarafından sıklıkla kullanıldığından ve metin tabanlı açık kaynaklı bir format olduğundan seçilen, GeoJSON formatına dönüştürülmüştür. Ardından bu üç veri seti, popüler GeoPandas ve Shapely paketleri aracılığıyla veri hazırlama aşaması için Python ortamına aktarılmıştır. Bu aşamada Türkçe'ye özgü karakterler Latin karşılıklarıyla değiştirilmiş ve detaylara INSPIRE direktifinde önerildiği şekilde benzersiz tanımlayıcı (Unique ID) URI'ler eklenmiştir (INSPIRE, 2014). Uygulamadaki URI'ler aşağıda örneği sunulan şekilde önerilen formata uygun olarak oluşturulmuştur.

`http://{subdomain}.{domain}/{type}/{metadata}/{authority}/{localid}/{version}`

Gönüllüler, etiketleri OSM'de anahtar:değer çiftleri olarak tanımlamakta özgürdür ve bu etiketlerin büyük bir kısmı orijinal olarak “diğer etiketler” adlı bir sütunda tutulmaktadır. Gönüllülerden mümkün olduğunca fazla bilgi korumak ve “diğer etiketler” sütununu ayrı sütunlara bölmek için, OSM verileri overpass-turbo web

sitesinden her etiket için ayrı bir sütun olacak şekilde aşağıda verilen kod aracılığıyla GeoJSON formatında elde edilmiştir. Ancak, bu yöntemin çok sayıda sütun oluşturması gibi bir dezavantajı vardır. Bu dezavantaj, daha sonra ilişkisel modeldeki verilerin RDF'ye dönüştürülmesi aşamasında giderilecektir.

```
way({{bbox}})(if:count_tags() > 0);
out geom; out meta;
```

Bölüm 3.2'de belirtilen LinkedGeoData projesindeki bağlantılı veriler uygulamanın gerçekleştirildiği süreçte SPARQL uç noktasında erişilebilir olmadığından bu çalışmada kullanılamamıştır. Proje web sayfasında indirilebilir statik dosyalar yer alsa da bunlar güncel değildir. Ayrıca, bu dosyalar kısıtlı sayıda etiket içerdiğinden gereksinimleri tam olarak karşılamamaktadır. İfade edilen bu sebeplerden dolayı bağlantılı veri olarak LinkedGeoData projesinde yer alan OSM verileri kullanılamamış, bunun yerine GeoJSON olarak indirilen veriler ilerleyen süreçte RDF'e dönüştürülmüştür.

Aynı mekansal nesneyi temsil eden farklı kaynaklardan gelen mekansal detayları birbirleriyle ilişkilendirmek için bir eşleştirme işleminin gerçekleştirilmesi gerekir. Bu bağlamda, coğrafi mekanda ortak konumu paylaşan mekansal detaylar potansiyel olarak aynı mekansal nesneyi temsil eder. Ancak çeşitli nedenlerle veri setleri arasında geometrik tutarsızlıklar (konumsal kayma, morfolojik fark vb.) olabilir. Bu, esas olarak veri hatalarından (örneğin, uydu görüntüsü hataları) veya gönüllülerin çoğunluğu profesyonel haritacı olmadığı için kullanıcı hatalarından kaynaklanmaktadır. Bu bağlamda, Rutzing vd. (2009) uydu görüntülerinde bina geometrilerinin bozulmasının yaklaşık %30 olduğunu tespit etmiştir. Fan, Zipf ve Fu (2014), OSM ile Alman Resmi Topografik-Kartografik Bilgi Sistemi (ATKIS) arasındaki karşılık gelen binaları tanımlamak için Denklem 4.1'deki kesişen alan ile daha küçük bina alanı arasındaki oran (O_{KA}) değerini kullanmıştır. Yazarlar, O_{KA} için geometrik bozulmalardan kaynaklı %30'luk değeri eşik değer olarak kullanmış ve O_{KA} değeri %30'dan büyükse, bu iki binayı eş binalar olarak tanımlamışlardır.

$$O_{KA} = \frac{Alan_{kesisen}}{\min(Alan_{ref_i}, Alan_{hdf_i})} \quad (4.1)$$

Burada $Alan_{kesisen}$ iki binanın kesişen alanının, ref_i referans veri setinde yer alan bina alanının ve hdf_i ise hedef veri setinde yer alan bina alanının büyüklüğünü göstermektedir ($i : 1...n$).

Literatürde bazı çalışmalar, binaları geometrilerle birlikte öznitelikleri ve topolojik ilişkileri de kullanarak eşleştirir (Regalia vd., 2018; Revell & Antoine, 2009; X. Zhang

vd., 2014; Y. Zhang vd., 2019). Ayrıca, bazı çalışmalar öznitelik tabanlı eşleştirmeyi geometrik eşleştirme ile birleştirir, ancak bunun için iyi organize edilmiş ve eksiksiz bir veri seti gerekir. OSM'den gelen öznitelik verileri çoğunlukla tam değildir ve iyi organize edilmemiştir (Baglioni vd., 2007). Bazı alanlarda ise, OSM verileri daha yüksek kalite ve tamlığa (öznitelik bağlamında) sahiptir, ancak bu durum veri setinin tamamında homojen değildir (Fan, Zipf, Fu, & Neis, 2014). Bu, OSM'nin öznitelik verilerinin resmi mekansal verilerle eşleştirilmesini zorlaştırır. Öte yandan, bazı alanlarda yeterli bilgi varsa, OSM binalarının öznitelikleri geometrik eşleştirmenin doğrulanması için kullanılabilir. Ancak, OSM ile ilgili yukarıda söz edilen sorunlardan dolayı, öznitelik temelli detay eşleştirmesinin genel olarak daha az güvenilir olduğu ifade edilebilir.

Bina detayları, geometrik olarak sıklıkla çokgen şeklinde temsil edilir ve farklı ölçeklerde aynı bina nesnesine karşılık gelen bina geometrileri belirli bir oranda örtüşmelidir. Fan, Zipf ve Fu (2014), OSM ile resmi ATKIS veri setindeki binalar arasında 4 m'lik bir kayma tespit etmiş ve bunun esas olarak %30 oranında bina geometrilerinin bozulmasından kaynaklandığı sonucuna varmıştır. Daha sonra, bu bozulma değeri eşleşen binaları belirlemek amacıyla alan kesişimi oranı için eşik değeri olarak kullanılmıştır. Basaraner (2020) bu uygulamanın çalışma alanını da kapsayan bir bölgede, resmi büyük ölçekli topografik veriler ile OSM arasında ortalama 2,41 m'lik bir konumsal kayma olduğunu saptamıştır. Uygun kesişim oranını tespit etmek amacıyla çalışma alanının küçük bir bölgesinde alan kesişim oranı (O_{KA}) için bir eşik değeri belirlenmiş ve bu değeri manuel bir eşleştirme ile karşılaştırılarak değerlendirilmiştir. Literatürde farklı eşik değerleri ile çeşitli çalışmalarda alan kesişimi yaklaşımı kullanılmışken (Fan, Zipf, Fu, & Neis, 2014; Liu vd., 2018; Rutzing vd., 2009), eşik değerinin belirlenmesi kısmi benzersizdir. Bir kıyaslama olarak manuel eşleştirmeye karşı otomatik eşleştirmenin performansını değerlendirmek için kesinlik (precision), duyarlılık (recall) ve F_1 -skor (F_1 -score) parametreleri kullanılmıştır. Bu parametrelerin hesaplanmasında kullanılan karışıklık matrisi Tablo 4.2'de verilmiştir.

Tablo 4.2 Performans değerlendirmesi için kullanılan karışıklık matrisi

		Tahmin Edilen Değer (Algoritma)	
		Pozitif	Negatif
Gerçek Değer	Pozitif	Doğru Pozitif (TP)	Yanlış Negatif (FN)
	Negatif	Yanlış Pozitif (FP)	Doğru Negatif (TN)

Kesinlik, doğru tahmin edilen gerçek değerlerin tüm doğru tahmin edilenlere oranıdır. Başka bir ifadeyle, doğru şekilde tahmin edilen değerlerin gerçekte hangi oranda

doğru olduğunu gösterir. Duyarlılık, doğru tahmin edilen gerçek değerlerin tüm gerçek değerlere oranıdır. Yani, doğru olarak tahmin edilmesi gereken değerlerin ne oranda doğru tahmin edildiğini gösterir. F_1 -skoru ise kesinlik ve duyarlılık değerlerinin harmonik ortalamasıdır ve sınıflandırmanın performansı hakkında bilgi verir.

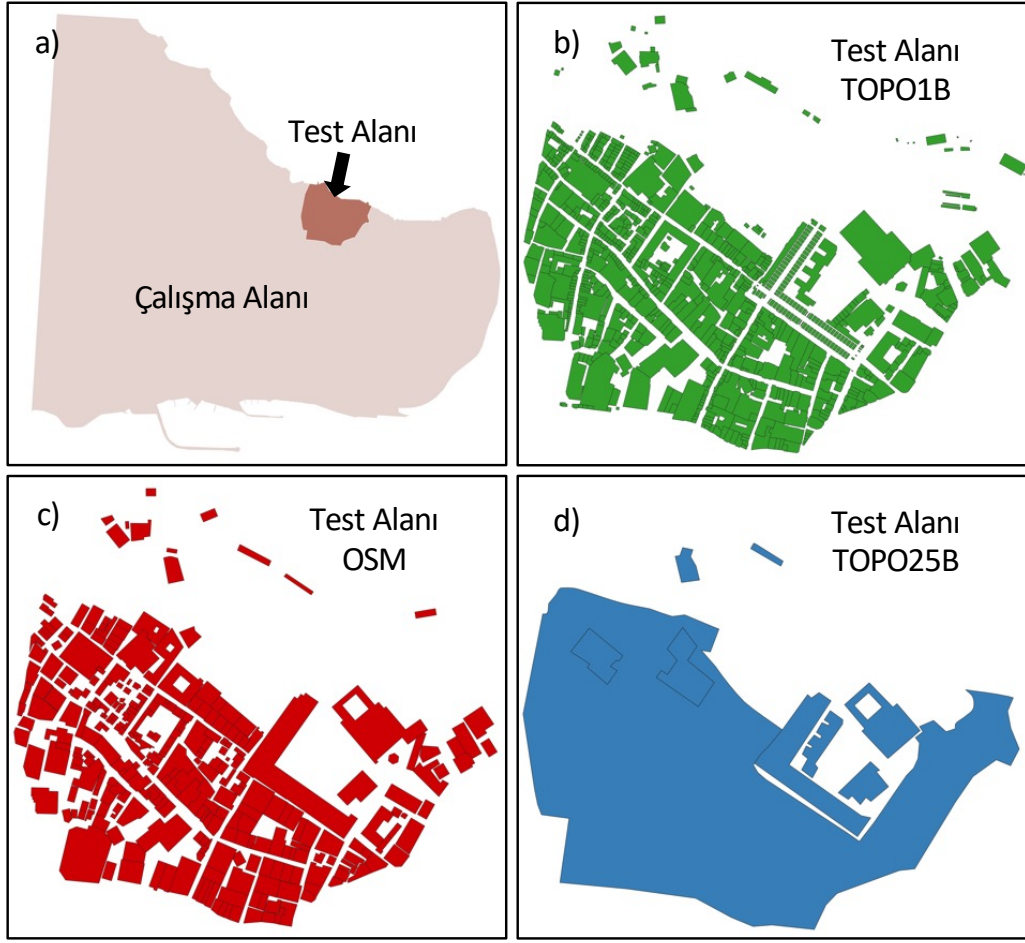
Bu parametreler, uzun süredir sınıflandırma veya kaynakları eşleştirme için kullanılmakta ve etkinliğin ölçülmesi için yeterli kabul edilmektedir (Savoy & Gaussier, 2010). Bu parametreler, sırasıyla Denklem 4.2, 4.3 ve 4.4 ile hesaplanır.

$$Kesinlik = \frac{Dogru Pozitifler (TP)}{Dogru Pozitifler (TP)+Yanlis Pozitifler (FP)} \times 100 \quad (4.2)$$

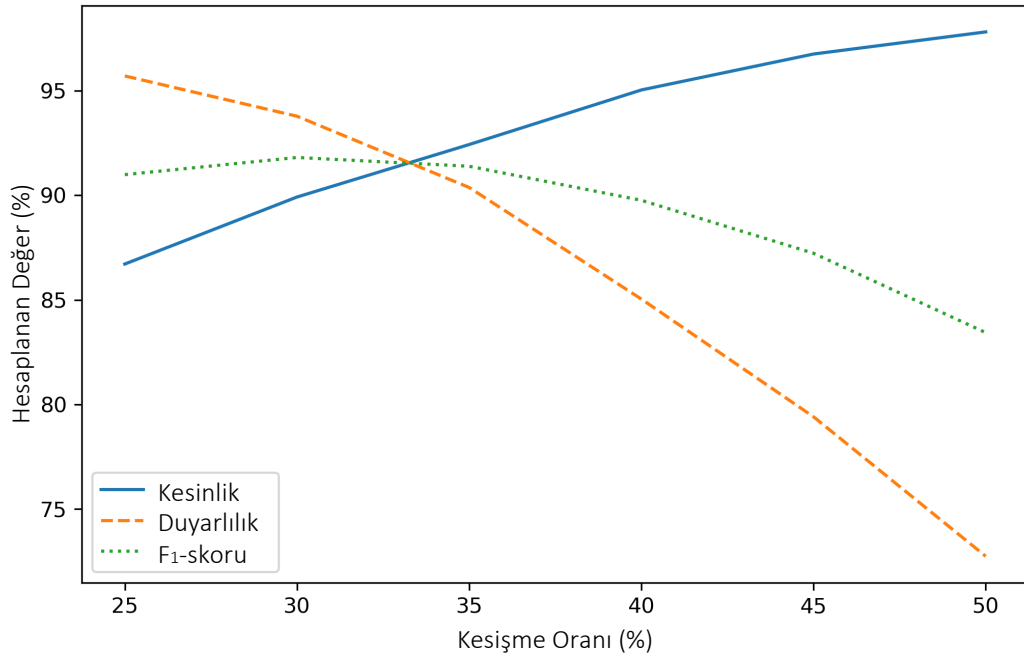
$$Duyarlilik = \frac{Dogru Pozitifler (TP)}{Dogru Pozitifler (TP)+Yanlis Negatifler (FN)} \times 100 \quad (4.3)$$

$$F_1 - skor = \frac{2 \times kesinlik \times duyarlilik}{kesinlik + duyarlilik} \times 100 \quad (4.4)$$

Uygun alan kesişimi (O_{KA}) eşik değerini belirlemek ve geometrik eşleştirmeyi doğrulamak için tüm çalışma alanı ile benzer özellikler gösteren küçük bir alt alan test alanı olarak seçilmiştir (Şekil 4.4). Bu test alanında, TOPO1B ve OSM arasında karşılık gelen binaları belirlemek için önce manuel eşleştirme yapılarak kıyaslama verileri hazırlanmıştır. Test alanında TOPO1B veri setindeki toplam 698 binadan 596'sı, OSM veri setindeki toplam 321 binadan 315'i ile manuel olarak eşleştirilmiş ve toplam 675 eşleşme çifti belirlenmiştir. İkinci aşamada, bir girdi parametresi olarak Denklem 1'e göre %25-%50 arasında değişen farklı O_{KA} değerleri kullanılarak aynı veri setleri üzerinde bir Python betiği yardımıyla birkaç otomatik eşleştirme gerçekleştirilmiştir. Son olarak, O_{KA} 'nın en uygun eşik değerini belirlemek için otomatik eşleştirme sonuçlarının kesinlik, duyarlılık ve F_1 -skor değerleri kıyaslama verilerine göre hesaplanmıştır (Şekil 4.5).



Şekil 4.4 Çalışma alanı üzerinde seçilen alt alan (a) ve veri setlerinin (b, c, d) alt alanlardaki gösterimi



Şekil 4.5 Alan kesişim oranı ile kesinlik, duyarlılık ve F_1 -skor değerleri

Şekil 4.5'teki sonuçlar karşılaştırıldığında tüm değerlerin %35 O_{KA} 'e yakın bir değerde (~ 34.1) kesiştiği (kesinlik ve duyarlılık arasında bir denge bulduğu) görülmektedir. Bu nedenle, çalışma alanı için %35 O_{KA} değeri eşik değer olarak belirlenmiştir (farklı O_{KA} aralıklarının tüm değerleri için Ek B'ye bakınız). Belirlenen eşik değer için test alanındaki otomatik eşleştirmenin Doğru Pozitif (TP), Yanlış Pozitif (FP) ve Yanlış Negatif (FN) değerleri Tablo 4.3'te gösterilmiştir. Bu bağlamda, Tablo 4.3 incelendiğinde, kıyas kabul edilen 675 manuel bina eşleştirmesinden 610'u belirlenen O_{KA} değeri kullanılarak otomatik olarak tespit edilmiştir.

Tablo 4.3 Otomatik eşleştirmenin TP, FP ve FN değerleri

	TP	FP	FN
Otomatik Eşleştirme	610	50	65

Sonuç olarak, belirlenen O_{KA} değeriyle üç veri seti arasında iki eşleştirme işlemi gerçekleştirilmiştir. Birincisi TOPO1B (referans veri seti) ve OSM (hedef veri seti) arasındadır. İkincisi ise, OSM (referans veri seti) ve TOPO25B (hedef veri seti) arasındadır. Bu eşleştirme işlemleri bir Python betiği ile gerçekleştirilmiştir. O_{KA} değeri orijinal olarak TOPO1B ve OSM veri setlerinin geometrik eşleştirilmesinden elde edilmiş olsa da, ampirik olarak gerçekleştirilen ön testler aynı değer için OSM binalarını TOPO25B binaları ile eşleştirmek için de uygun olduğunu göstermiştir. Bu durum, TOPO1B ve TOPO25B binalarındaki mutlak konumsal doğruluğun birbirine büyük ölçüde benzer olmasından kaynaklanmaktadır. Bu bağlamda, kıyaslama veri setini hazırlamak amacıyla bu iki veri seti için de bir öncekiyle aynı manuel eşleştirme prosedürü uygulanmıştır. Bu amaçla sırasıyla 590 adet OSM binası, 224 adet TOPO25B binası ile ve 18592 adet OSM binası da 1 adet TOPO25B meskun alanı ile eşleştirilmiştir. Sonuç olarak, performans istatistikleri %90'ın üzerinde hesaplanmıştır (Tablo 4.4). Fakat burada TOPO25B meskun alan istatistiklerinin sadece bir çokgene dayandığı göz önünde bulundurulmalıdır.

Tablo 4.4 Eşleştirme performans istatistikleri

Eşleştirme	Kesinlik	Duyarlılık	F1-skor
TOPO1B-OSM	92.4	90.4	91.4
OSM-TOPO25B Binaları	91.2	95.7	93.4
OSM-TOPO25B Meskun Alanları	99.7	99.9	99.8

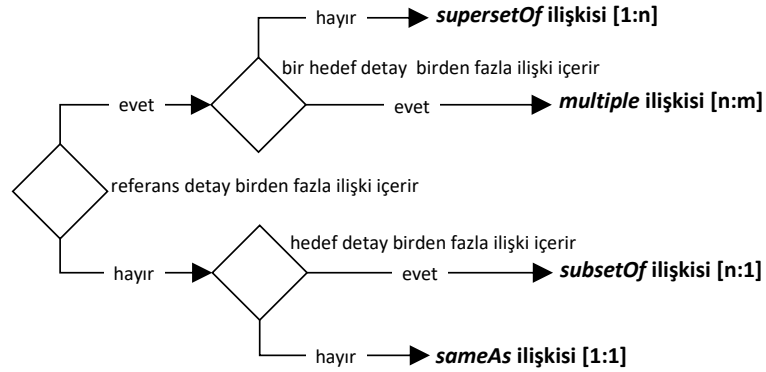
Farklı ölçeklerdeki bina detaylarının kardinal ilişkilerinin modellenmesi, kartografik genelleştirme süreçlerinin karmaşıklığı nedeniyle zor olabilir. Veri setlerinin farklı kökenleri ve üreticilerin farklı bakış açıları bu sorunu daha da zorlaştırmaktadır. Örneğin, bir veri setinde beş bina detayı ile temsil edilen bir bina grubu, OSM veri setinde tutarsız bir şekilde üç bina detayı ile temsil edilebilir. Bu sorun, esas olarak, OSM'ye katkıda bulunanlar tarafından uydu görüntülerinden bina detaylarının ortak

sınırlarını belirleme zorluğundan kaynaklanmaktadır. Bu tür bir tutarsızlık sorunu, farklı ölçeklerde veya kalitedeki veri setlerinin entegrasyonunu zorlaştırmaktadır. Bu soruna bir çözüm olarak, mantıksal karakteristik özellikler kullanılarak mekansal kardinal ilişkileri için bir sözlük oluşturulmuş ve Tablo 4.5'teki gibi tanımlanmıştır.

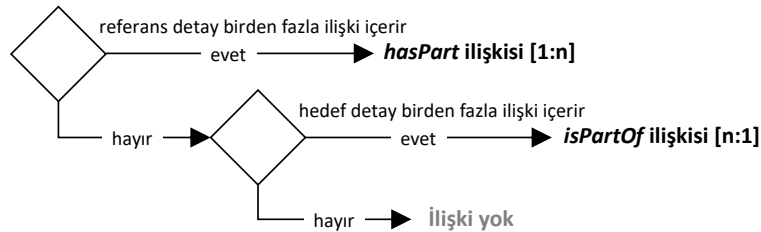
Tablo 4.5 Kardinal ilişki türleri ve özellikleri

İlişki Türü	Detay Türü	Mantıksal Karakteristik Özelliği
sameAs (aynı) [1:1]	Bina - Bina	Geçişken ve Simetrik
subsetOf (altKüme) [n:1]	Bina - Bina	Geçişken
supersetOf (üstKüme) [1:n]	Bina - Bina	Geçişken
multiple (çoklu) [n:m]	Bina - Bina	Simetrik
isPartOf (parça) [n:1]	Bina - Meskun Alan	Geçişken
hasPart (bütün) [1:n]	Bina - Meskun Alan	Geçişken

Referans ve hedef veri setleri arasında eşleşen binaların ilişki sayıları (dereceleri) incelenmiş ve Şekil 4.6'daki gibi Wiemann ve Bernard (2016)'dan uyarlanan tanımlama kuralları ile binalar arasındaki ilişki türleri tespit edilmiştir. Aynı şekilde binalar ve meskun alan arasındaki ilişki türleri de Şekil 4.7'de gösterilen tanımlama kuralları ile belirlenmiştir.



Şekil 4.6 Binalar arası ilişki türlerinin tanımlama kuralları (Wiemann ve Bernard (2016)'dan uyarlanmıştır)



Şekil 4.7 Binalar ve meskun alan arasındaki ilişki türlerinin tanımlama kuralları

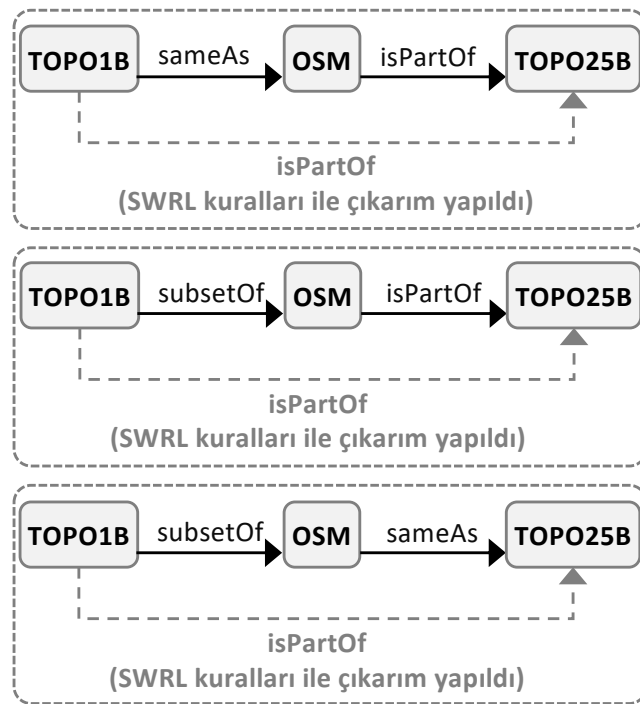
Binalar arasındaki ilişki türleri belirlendikten sonra bu bilgiler mekansal ontolojiye aktarılmıştır. Öte yandan, doğrudan ontolojide tanımlanmamış olan TOPO1B

ve TOPO25B veri setleri arasındaki bina ilişkileri, semantik web'in önemli bir özelliği olan çıkarım özelliğiyle belirlenebilir. Şekil 3.6'da gösterildiği üzere her iki eşleştirmede aynı ilişki türü varsa ve bu ilişki türü geçişken özelliğe sahip ise TOPO1B ve TOPO25B arasında bu ilişkiler çıkarım yoluyla otomatik olarak tespit edilebilir. Fakat tüm ilişkiler bu şekilde kurulamayacağından bir geliştirmeye ihtiyaç duyulmaktadır. Bu işlem için aşağıda sunulan SWRL kuralları oluşturulmuş ve bu kurallar Protégé yazılımında mekansal ontolojiye uygulanmıştır (Şekil 4.8). Böylece geometrik eşleştirme yapılmamış olan TOPO1B ve TOPO25B detayları arasında da ilişki kurulmuştur. Bu bağlamda, mekansal nesneler arasında SWRL kuralları yardımıyla ek ilişkilerin tespit edilmesi özgün bir yöntemdir.

@prefix spo: <http://ulusalharita.tr/ont/spo/>.

@prefix dul: <http://www.ontologydesignpatterns.org/ont/dul/DUL.owl/>.

$spo:sameAs(?t1k,?osm) \wedge dul:isPartOf(?osm,?t25k) \Rightarrow dul:isPartOf(?t1k,?t25k)$
 $spo:subsetOf(?t1k,?osm) \wedge dul:isPartOf(?osm,?t25k) \Rightarrow dul:isPartOf(?t1k,?t25k)$
 $spo:subsetOf(?t1k,?osm) \wedge spo:sameAs(?osm,?t25k) \Rightarrow dul:isPartOf(?t1k,?t25k)$



Şekil 4.8 Binalar ve meskun alan arasındaki ilişki türlerinin tanımlama kuralları

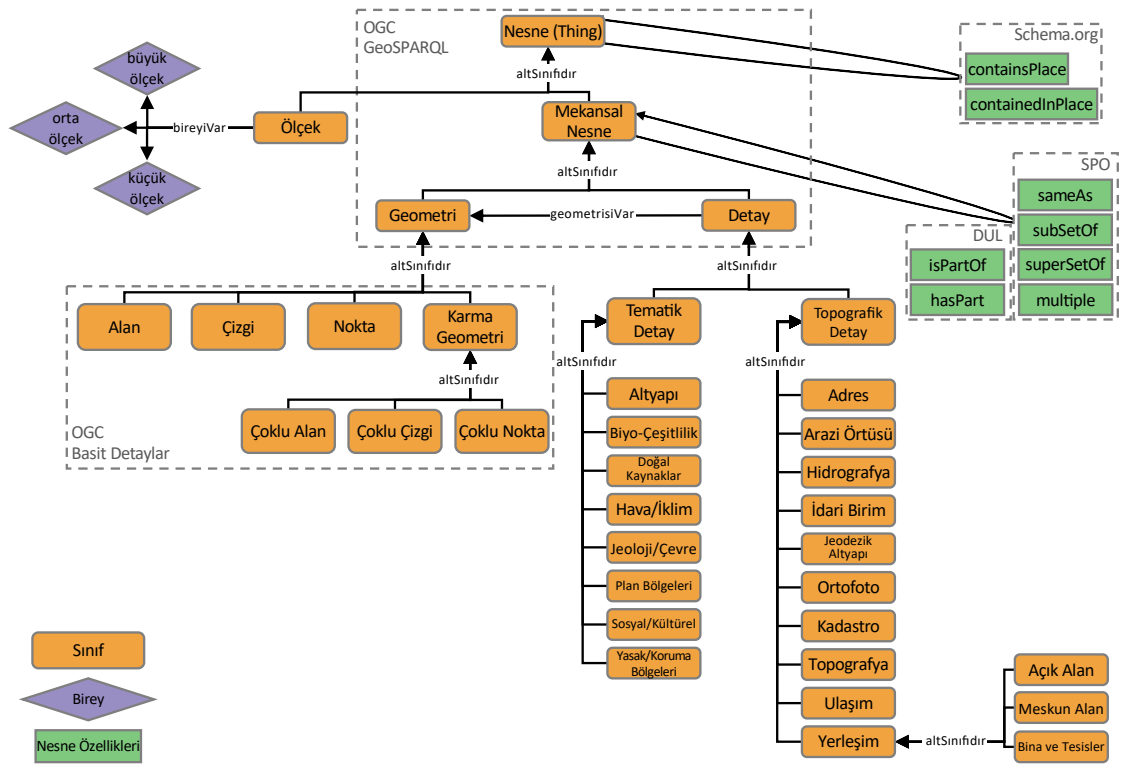
Mekansal ontoloji oluşturmak için ilk adım olarak, açık kaynak ontoloji editörü Protégé ile OWL'de mekansal kavramlar, geometri türleri, kartografik ölçek ve binalar arasındaki ilişkiler tanımlanmıştır. Mekansal ontoloji, esas olarak GeoSPARQL sözlüğüne ve Huang vd. (2018) ve Eiðsson (2018) tarafından tanımlanan ölçek sözlüğündeki ölçek tanımlarına dayalı olarak tasarlanmıştır. Bununla birlikte mevcut

ontolojileri/sözlükleri yeniden kullanmak için SKOS, DC, DUL ve schema.org gibi diğer yaygın olarak kullanılan ontolojilerden/sözlüklerden bazı bölümleri de içermektedir. Örneğin, binaların ilişkilendirilmesinde kullanılan parça-bütün (isPartOf - hasPart) ilişkileri DUL ontolojisinden alınmıştır. Ayrıca, bir idari birim içindeki binaları tanımlamak için schema.org'dan schema:containedInPlace özelliği kullanılmıştır. Bununla birlikte, mekansal ontolojinin detay sınıfları, binalarla birlikte diğer detay sınıflarını da içeren Türkiye Ulusal Mekansal Veri Altyapısı'ndan alınmıştır.

Mekansal ontoloji oluşturulduktan sonra, ilişkisel veri formatlarındaki bina detaylarının RDF'ye dönüştürülmesi ve mekansal ontolojiye bireyler olarak eklenmesi gerekmektedir. İlişkisel modeldeki mekansal verileri RDF'ye dönüştürmek için üretilmiş araçlar olmasına rağmen, bunların hiçbiri genel kabul görmemiş veya uzun vadeli bakım desteği almamıştır. GeoTriples yazılımı, aktif bir web sitesi ve bakım desteğine sahip olduğundan burada diğerlerine göre öne çıksa da veri setlerimizi GeoTriples ile dönüştürmek çok fazla zaman aldığından dolayı mümkün olmamıştır. Protégé yazılımındaki entegre bir araç olan Cellfie eklentisi de geometrilerin WKT olarak bir sütunda yer aldığı elektronik tablolardan verileri dönüştürmek için kullanılabilir. Ancak bu araç da, bu çalışmada kullanılan veri miktarını işleyememiştir. Bu araçların yetersizliğinin ana nedeninin, OSM'de gönüllülerin sağladığı tüm bilgileri kullanmak uğruna yapılan işlemle ortaya çıkan seyrek satırlı 169 sütunun olduğu düşünülmektedir. OSM veri setindeki sütunların içinde satırların sadece %5'i doluyken, geri kalanı boştur. Bu durumda, RDF'ye dönüşüm için sözü geçen yazılımlar kullanılamadığından bina detayları, RdfPandas, RDFLib ve GeoPandas gibi Python paketleri kullanılarak dönüştürülmüştür.

Detayların ve ilişkilerin mekansal ontolojiye rahatlıkla eklenebilmesi için Turtle formatı tercih edilmiştir. Çünkü Turtle, dosyalarının sonuna ek bilgiler eklenerek kolayca genişletilebilir bir format sunmaktadır (XML formatı gibi başlangıç veya bitiş etiketi içermemektedir). Bu nedenle, farklı işlem adımlarının her biri için (bina detayları, veri özellikleri, diğer kaynaklarla ilişkiler vb.) bir Turtle dosyası oluşturulmuş ve ardından hepsi mekansal ontoloji şeklinde tek bir dosya elde etmek için birleştirilmiştir. Oluşturulan mekansal ontolojinin sınıfları ve ilişkileri Şekil 4.9'da görülmektedir.

Geometriler ayrı ayrı adlandırılmış bireyler olarak kullanılmış ve WKT formatında detaylarla ilişkilendirilmiştir. WKT için varsayılan Koordinat Referans Sistemi (CRS) WGS84 olduğundan, diğer CRS'leri desteklemeyen yazılım paketleri arasındaki olası uyumsuzlukları önlemek için bu koordinat sistemi seçilmiştir. Bu nedenle orijinal veri setlerinin koordinatları WGS84 sistemine dönüştürülmüştür. İlişkisel coğrafi verilerin RDF'ye dönüştürülmesi işlemi, 2,7 GHz Intel Core i5-5257U CPU'ya, 8 GB



Şekil 4.9 Mekansal ontoloji sınıfları ve ilişkileri

belleğe (RAM), bir katı hal sürücüsüne (SSD) ve 64 bit macOS (11.2.2) işletim sistemine sahip olan bir bilgisayarda tek bir iş parçacığı (single thread) çalıştırılarak yaklaşık 8 dakika sürmüştür. Bu süre diğer yazılımlarla (GeoTriple ve Cellfie plugin) karşılaştırıldığında son derece avantajlıdır. RDF dönüşümünün ardından, sonraki işlemlerde verilerin işlenmesini kolaylaştırmak için özellikle OSM verisinde fazlaca yer alan boş satırlar nedeniyle oluşan boş üçlüler silinmiştir. İlişkisel modelde yer alan verideki sütun başlıkları da mekansal ontoloji için veri özellikleri oluşturmak amacıyla kullanılmış ve bir Python betiği ile Turtle formatına dönüştürülmüştür.

OSM veri setinde, bazı detaylar için Wikipedia bilgileri bulunmaktadır. Bu bilgiler bir Python betiği aracılığıyla uygun bir kaynak adı oluşturmak ve DBpedia ile arasında bağlantılar kurmak için kullanılmıştır. Ayrıca tüm TOPO1B binalarının Fatih ilçesinde olması ve Fatih ilçesinin DBpedia ve YAGO gibi popüler bilgi tabanlarında kaynak olarak yer alması nedeniyle, bu kaynaklar arasında kurulan owl:sameAs ilişkileri ile ortak referans sorunu çözülmüştür. Bu bağlantılar, oluşturulmalarının ardından bir Python betiği ile Turtle formatına dönüştürülmüştür.

Ölçekler arasındaki detay ilişkileri de Turtle formatına dönüştürülmüş ve tüm bilgiler bir araya getirilerek bir mekansal ontoloji oluşturulmuştur. Mekansal ontolojiden bir kaynağın örnek bir RDF'si aşağıda verilmiştir. Sonuç olarak, bu mekansal ontoloji, çok ölçekli bir ortamda birbiriyle ilişkili üç farklı veri setini barındırmaktadır.

```

@[prefix definitions]
  [TOP01K_resource]    rdf:type owl:NamedIndividual ,
                        spo:BuildingAndFacility ;
                        schema:containedInPlace idt:Fatih ;
                        spo:subSetOf [OSM_resource] ;
                        dul:isPartOf [TOP025K_resource] ;
                        geo:hasGeometry [TOP01K_geometry] ;
                        ...

  ...

  [TOP01K_geometry]    rdf:type owl:NamedIndividual ,
                        sf:Polygon ;
                        spo:hasScale spo:largeScale ;
                        geo:asWKT "POLYGON ((28.94179428
                        41.00279794, ...))" ;

  ...

  idt:Fatih owl:sameAs <http://dbpedia.org/resource/Fatih> .
  idt:Fatih owl:sameAs <http://yago-knowledge.org/resource/Fatih> .

```

İlgili tüm bağlantılı bilgileri içeren bir veri ağı, semantik web'in temel amaçlarından biridir. Bu nedenle bağlantılı veriler, Bölüm 1.1'de ifade edilen ilkelere göre HTTP standardı kullanılarak web üzerinde erişilebilir olmalıdır. Şu anda web'te bulunan bağlantılı veri setlerinin çoğuna SPARQL uç noktaları aracılığıyla erişilmektedir. Bu çalışmada, mekansal veriler için daha uygun olan GeoSPARQL standardıyla uyumlu bir uç nokta oluşturulmalıdır. Bu gereksinimleri karşılamak için Apache Jena, Parliament (2021) ve Strabon (2021) gibi açık kaynaklı yazılım alternatifleri bulunmaktadır. Bu çalışmada, Apache Jena projesinden GeoSPARQL uyumlu bir HTTP sunucusu olan GeoSPARQL Fuseki kullanılmıştır. Java tabanlı basit bir araç olmasına rağmen, web'de verileri sunmak ve HTTP üzerinden sorguları yanıtlamak için yeterince güçlüdür.

Web üzerinden sunumu gerçekleştirilen bağlantılı verilerin gösterimini sağlamak amacıyla HTTP üzerinden örnek sorgular gerçekleştirilmiş ve cevapları yerel (localhost) ortamda web tabanlı bir harita uygulamasında gösterilmiştir. Mekansal ontolojiyi sorgulamak için bir SPARQL uç nokta arabirimi olan SPARQLWrapper adlı bir Python paketi kullanılmıştır. Sorgulardan elde edilen sonuçlar Python ortamında hazırlanmış ve mekansal veriler JavaScript tabanlı etkileşimli bir web harita kütüphanesi olan Leaflet'i kullanan Folium paketi ile görselleştirilmiştir.

4.2.3 Bulgular

Sonuç olarak, üç farklı veri seti geometrik eşleştirme kullanılarak entegre edilmiş ve Tablo 4.6'da ilişki türlerine göre sayıları özetlenen bina detayları arasındaki çeşitli mantıksal özelliklere sahip kardinal ilişkileri ortaya çıkarmak için ontolojilerle tanımlanmıştır. Bu bağlamda, 35263 isPartOf ilişkisinden 16518'ini belirlemek için SWRL kuralları kullanılırken, geri kalanı geometrik eşleştirme ile belirlenmiştir. Tüm bu ilişkiler daha sonra çoklu temsilleri kolaylaştırmak ve farklı kullanıcıların çeşitli gereksinimlerine yönelik soruları yanıtlamak amacıyla kullanılmıştır.

Tablo 4.6 İlişki türlerine göre oluşturulan ilişki sayıları

İlişki Türü	İlişki Sayısı
sameAs [1:1]	8053
subsetOf [n:1]	11819
supersetOf [1:n]	2084
multiple [n:m]	10064
isPartOf [n:1]	35263

Mekansal ontolojide hasPart ilişkisi tanımlanmış olmasına rağmen, OSM veya TOPO1B veri setlerindeki detaylar arasında meskun alan yer almadığı için hasPart ilişkisi doğrudan tanımlanmamıştır. Ancak bu ilişki türü isPartOf ilişkisinin tersi olduğu için bu detaylar arasında çıkarım yapıldığında otomatik olarak kurulabilmektedir.

Gösterim için, mekansal ontoloji sorgulanarak cevaplanabilecek üç soru hazırlanmıştır. Sorular aşağıdaki gibidir:

1. *Tarihi Kapalıçarşı'da hangi dükkanlar/yerler vardır?*
2. *Tüm alanda tekerlekli sandalye kullanımına uygun yapılar hangileridir?*
3. *Hangi binalar her üç ölçekte de bireysel binalar olarak tasvir edilmiştir?*

Bu sorularla, farklı hedef kitlelerin farklı ihtiyaçlarına cevap verilmesi amaçlanmaktadır. Örneğin, ilk sorunun cevabı bölgeyi keşfetmek isteyen bir turist için, ikincisi engelli bir kişi için ve sonuncusu bir haritacının ihtiyaç duyacağı bir cevap olabilir. Bu sorular, farklı ölçeklerde farklı veri setlerini temsil eden mekansal ontoloji ile yanıtlanabilir. Sırasıyla soruların sorgu ve cevapları aşağıda sunulmuştur.

Birinci soruyu cevaplamak için aşağıda verilen SPARQL sorgusu hazırlanmış ve HTTP üzerinden istenmiştir. Sorgunun sonuçları Şekil 4.10'da bir web haritası üzerinde görselleştirilmiştir.

```
PREFIX dul: <http://www.ontologydesignpatterns.org/ont/dul/DUL.owl#>
PREFIX geo: <http://www.opengis.net/ont/geosparql#>
PREFIX spo: <http://ulusalharita.tr/ont/spo#>
```

```
SELECT *
WHERE {
    ?indg spo:hasName_en "Grand Bazaar" ; geo:hasGeometry ?indgg .
    ?indgg geo:asWKT ?indgwkt .
    ?gbsub spo:subsetOf ?indg ; geo:hasGeometry ?gbsubg .
    ?gbsubg geo:asWKT ?gbsubwkt . }
```



Şekil 4.10 Birinci sorgudan gelen cevabın web haritası üzerinde gösterimi

İkinci soruyu cevaplamak için aşağıda verilen SPARQL sorgusu hazırlanmış ve HTTP üzerinden istenmiştir. Sorgunun sonuçları Şekil 4.11’de bir web haritası üzerinde görselleştirilmiştir.

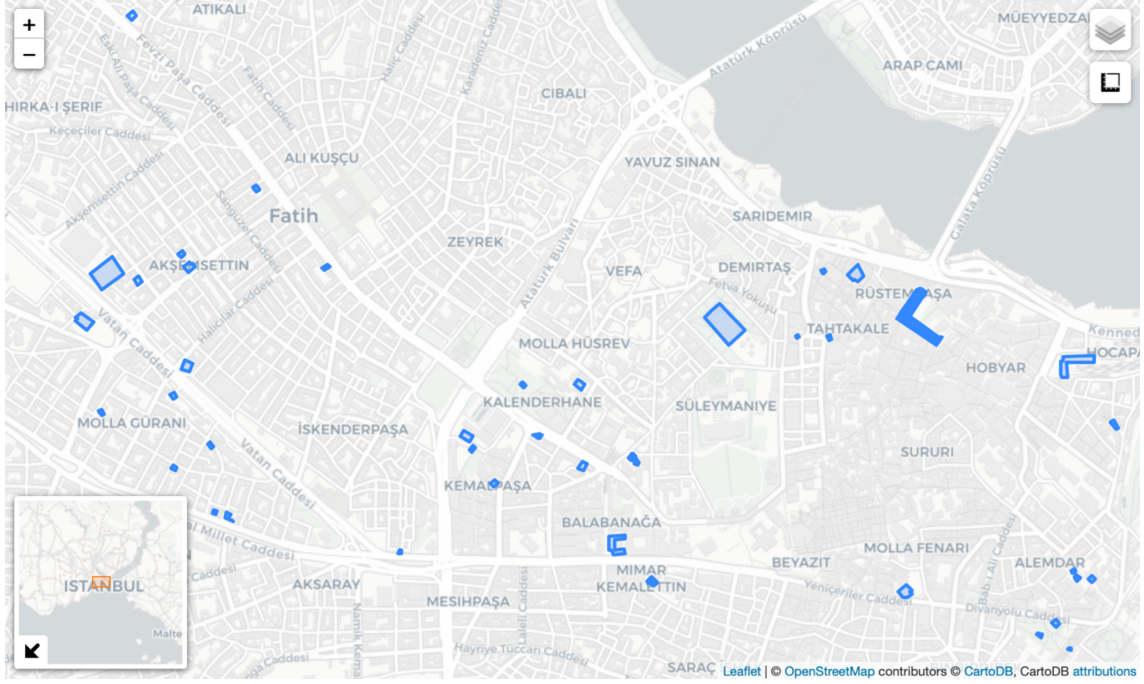
```
PREFIX dul: <http://www.ontologydesignpatterns.org/ont/dul/DUL.owl#>
PREFIX geo: <http://www.opengis.net/ont/geosparql#>
PREFIX spo: <http://ulusalharita.tr/ont/spo#>
PREFIX xsd: <http://www.w3.org/2001/XMLSchema#>
```

```
SELECT *
```

```

WHERE {
    ?ind spo:hasWheelchair 'yes' ; geo:hasGeometry ?indg .
    ?indg geo:asWKT ?indwkt .
    ?indw spo:hasWheelchair 'yes' .
    ?indr spo:subSetOf ?indw ; geo:hasGeometry ?indrg .
    ?indrg geo:asWKT ?indrwt .}

```



Şekil 4.11 İkinci sorgudan gelen cevabın web haritası üzerinde gösterimi

Üçüncü soruyu cevaplamak için aşağıda verilen SPARQL sorgusu hazırlanmış ve HTTP üzerinden istenmiştir. Sorgunun sonuçları Şekil 4.12’de bir web haritası üzerinde görselleştirilmiştir.

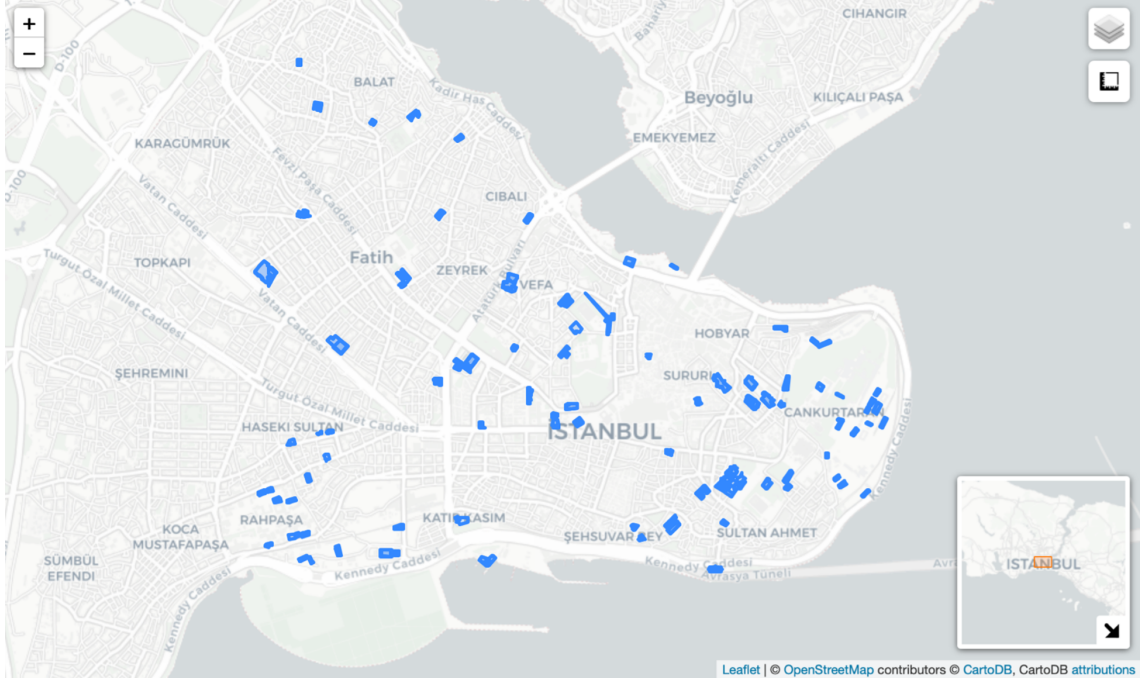
```

PREFIX geo: <http://www.opengis.net/ont/geosparql#>
PREFIX spo: <http://ulusalharita.tr/ont/spo#>
PREFIX skos: <http://www.w3.org/2004/02/skos/core#>
PREFIX xsd: <http://www.w3.org/2001/XMLSchema#>

SELECT *
WHERE {
    ?indl geo:hasGeometry ?indlg ; spo:sameAs ?indm .
    ?indlg spo:hasScale spo:largeScale ; geo:asWKT ?indlwkt .
    ?indm geo:hasGeometry ?indmg ; spo:sameAs ?inds .

```

```
?indmg spo:hasScale spo:mediumScale ; geo:asWKT ?indmwkt .  
?inds geo:hasGeometry ?indsg .  
?indsg spo:hasScale spo:smallScale ; geo:asWKT ?indswkt .}
```



Şekil 4.12        sorgudan gelen cevabın web haritası   zerinde g  sterimi

4.2.4 Tartıřma

Bu b  l  mde, uygulamanın sonucu olarak ortaya konan mekansal ontolojinin oluřturulması, kullanılması, sorgulanması ve g  rselleřtirilmesi ařamalarında   ne   ıkan noktalar ve zorluklar tartıřılmıřtır.

Veri setleri entegre edilmeden yukarıda belirtilen sorulara cevap verilemez.   rneğın TOPO1B veri setinde Kapalı arşı'nın sınırı yoktur, sadece m  nferit d  kkanlar/yerler vardır. Aksine, OSM veri setinde sadece Kapalı arşı'nın dıř sınırlarını belirten bir   okgen mevcuttur ancak m  nferit d  kkanlar/yerler yoktur. Bu nedenle, ilk sorgu, yanıt i  in OSM'den gelen sınırları ve TOPO1B'deki d  kkanları/yerleri kullanmaktadır. Bu durum diğ  r   rnek sorular i  in de ge  erlidir.   rneğın hem TOPO1B hem de TOPO25B veri setinde (resmi veri setleri) binaların tekerlekli sandalye uygunluğ   hakkında bir bilgi yoktur, ancak bu bilgi katkıda bulunan g  n  ll  lerin bir   abası olarak OSM'de bulunmaktadır. Fakat burada, TOPO1B ve OSM arasındaki doğ  ruluk farklılıklarının mekansal analiz sonu  larında bazı belirsizliklere yol a  abileceğ  i unutulmamalıdır.

Geometrik eřleřtirme, veri setlerini entegre etmek i  in kritik bir adımdır. Bu   alıřmada,

geometrilerine dayalı olarak eşleşen çok kaynaklı ve çok ölçekli veri setlerinde %90'ın üzerinde kesinlik ve duyarlılık değerlerine ulaşılmıştır. Burada manuel eşleştirmenin yapan kişiye göre değişebileceği ve bu durumun eşik değeri, kesinlik ve duyarlılık değerlerini belirli bir ölçüde etkileyebileceği unutulmamalıdır. Bu açıdan tecrübe önemlidir. Eşleştirme performansı, Bölüm 4.2.2'de bahsedildiği gibi öznitelikler veya topolojik ilişkiler kullanılarak geliştirilebilse de bu çalışmanın kapsamında değildir. Öte yandan, veri setlerinin tam özniteliklerinin olmaması, resmi verilerin OSM ile daha doğru eşleşmesinin önünde bir engeldir. Ayrıca, OSM, katı kalite kısıtlamaları olmaksızın gönüllülerin çabası ile oluşturulduğundan öznitelik değerlerinin tamlığı açısından homojen değildir. Katkıda bulunanların sayısının artmasıyla birlikte kalite yeterli bir seviyeye ulaştığında bu sorun zamanla ortadan kalkabilir. O_{KA} yaklaşımına dayalı detay eşleştirmesi bu uygulamaya özel değil, geneldir ve diğer çokgen veri setlerine uygulanabilir. Ancak, eşik değerin veri setleri arasındaki kalite ve ölçek farklılıklarına göre ayarlanması gerekebilir. Bu çalışmada ise, ikinci eşleştirme işlemi için de aynı eşik değeri kullanılmıştır.

TOPO1B binalarının yaklaşık %15'i ve OSM binalarının %2'si herhangi bir bina ile eşleşmemiştir. Bunun başlıca nedeni, bazı küçük binalar ve müstemilat gibi yapıların, gönüllüler tarafından tespit edilemedikleri için OSM veri setinde yer almamış olmasıdır. Ayrıca, veri setlerinin zamansal farklılıkları da binalar arasında uyumsuzluklara neden olabilir. Örneğin aynı konumda bir veri setinde yeni inşa edilmiş bir bina yer alırken diğer bir veri setinde eski bir bina yer alıyor olabilir. Bu durumda binalar geometrilerine göre eşleştirilebilir ancak bu eşleştirme semantik olarak yanlış olacaktır. Bu durum, 'yapım yılı' gibi bir öznitelik ile önlenabilir, ancak veri setlerinde bu bilgi yer almamaktadır.

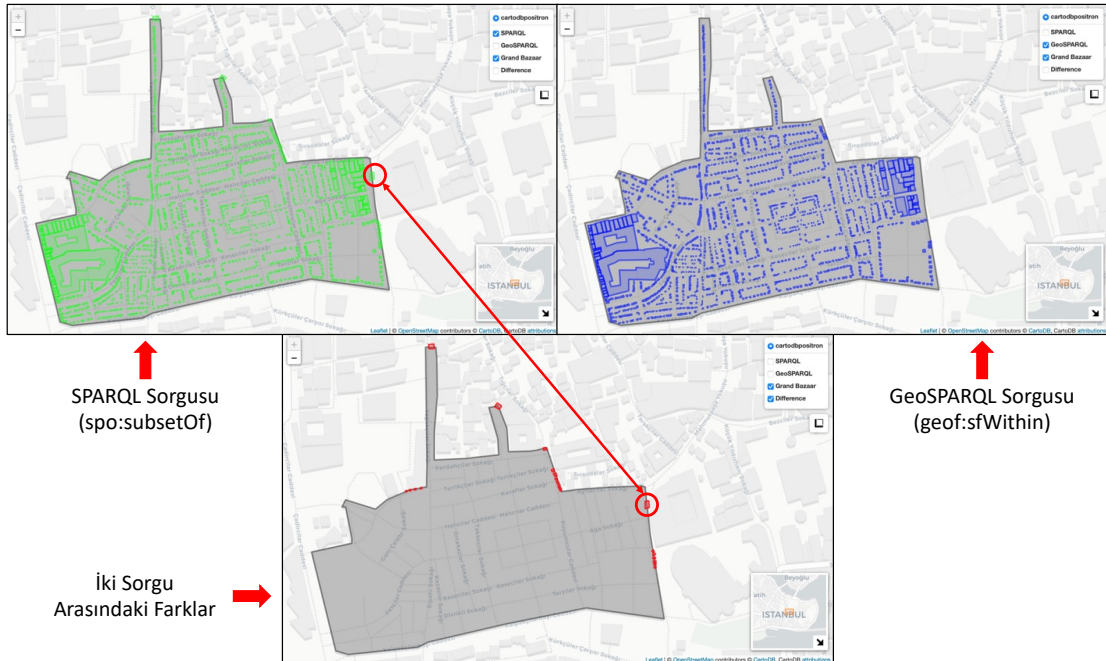
Burada, üçten fazla veri seti kullanıldığında, ilişkileri tanımlamanın ve çıkarım için mantıksal kuralların uygulanmasının daha karmaşık olacağı belirtilmelidir. Bu durumda veri setleri, tutarsızlıklar göz önünde bulundurularak kademeli olarak işlenebilir. Veri setlerini ayrıştırma düzeylerine göre sıralamak, bunların kademeli entegrasyonunu kolaylaştırabilir.

Daha önce ifade edildiği gibi, mekansal semantik web teknolojileri CBS'ye kıyasla yenidir ve henüz onun kadar olgunlaşmamıştır. Örneğin CBS ortamında topolojik bir analiz veya sorgulama saniyeler içinde yapılabilirken, semantik web'de özellikle topolojik ilişkiler kullanıldığında (kesişme, içerme vb.) GeoSPARQL sorgu işleme süresi daha uzun olabilmektedir. Bu bağlamda, çalışmanın en büyük katkısı, ontolojilerde yer alan mekansal detaylar arasındaki ilişkilerin önceden belirtilmesini sağlamak amacıyla detayların uygun mantıksal özellikleri ile kardinal ilişkilerinin tanımlanmasıdır. Bu süreç, sorguları yanıtlamak için daha düşük işlem süreleri

(yaklaşık 8 kat) sağlamıştır (Tablo 4.7). Örneğin, kardinal ilişkileri tanımlamışsa uygulamanın ilk örnek sorusu, topolojik ilişkilerin (ör., içerme) hesaplanması yoluyla GeoSPARQL kullanılarak sorgulanmalıdır. Ancak, kardinal ilişkilerin tanımı ile sözü geçen sorgular için topolojik ilişkilerin tanımlanmasına da GeoSPARQL standardının kullanılmasına da ihtiyaç yoktur. İlk sorgu, örnekte gösterildiği gibi, SPARQL kullanılarak gerçekleştirilebilir. Ancak bu durumda mekansal fonksiyonlarla (içerme, kesişme vb.) oluşturulan GeoSPARQL sorgu sonuçları, özellikle çokgen sınırlarında, eşik değerine bağlı olarak SPARQL sonuçlarından farklılık gösterebilir (Şekil 4.13). Bunun sebebi SPARQL sorgusunda kullanılan `spo:subsetOf` ilişkisinin %35 O_{KA} değeriyle kullanılmış olmasıdır. GeoSPARQL sorgusu ile yapılan `sf:within` işlemi sınırın tamamen içine giren (veya ağırlık merkezini içeren) detayları seçerken SPARQL sorgusu %35 kesişen detayları da seçmektedir. Diğer taraftan, bu sorgular komut dosyalarına sabit olarak kodlanmıştır, bu nedenle kullanıcı tarafından seçilemezler. Gelecekteki çalışmalarda, kullanıcıların HTTP üzerinden sorgu seçmesi veya oluşturması için bir web arayüzü oluşturulabilir.

Tablo 4.7 Birinci soru için GeoSPARQL ve SPARQL ile yapılan sorguların ortalama cevap süreleri

	SPARQL	GeoSPARQL
Ortalama Cevap Süresi (ms)	168.6	1406.2



Şekil 4.13 Birinci soru için GeoSPARQL ve SPARQL ile yapılan sorgular arasındaki mekansal farklılıklar

Koordinat sistemi tanımı ve dönüşümü, coğrafi veriler için esastır. Bu işlem, CBS’de kolayca yapılabilirken, mekansal semantik web’de zorlayıcı olabilir. Mekansal

verileri destekleyen semantik web yazılımlarının çoğu, yalnızca WKT için varsayılan olan coğrafi koordinat sistemini (WGS84) yönetebilmektedir. Mekansal işlemlerde, ölçü birimi olarak “metre” yerine varsayılan koordinat sistemi birimi olan “derece” kullanılması, durumu daha karmaşık hale getirmektedir. Örneğin, derece ölçü birimindeki koordinat sisteminde 50 m’lik bir tampon analizini semantik web araçlarında uygulamak zordur. WKT formatı farklı koordinat sistemlerini destekleyebilmesine rağmen, şu anda kullanılan semantik web yazılım paketlerinin çoğu yalnızca varsayılan koordinat sistemini desteklemektedir.

Mekansal detayları, halihazırda kullanılan ilişkisel modeldeki mekansal veri formatlarından bireyler olarak RDF’lere dönüştürmek yoğun bir çaba gerektirmektedir. Çalışmada, bu süreç büyük ölçüde Python betiklerine dayanmaktadır. Çünkü mevcut dönüştürme araçları çoğunlukla üniversite projeleri tarafından minimum desteğe sahip olarak geliştirilmektedir ve genellikle belirli bir miktardan fazlasını (uygulamada öznitelikleriyle birlikte yaklaşık 30000 detay yer almaktadır) işleyemez. Çalışmanın ana sınırlaması, uygulama için kullanılan Python betiklerinin verilere bağımlı olması ve verilerle ilgili arka plan bilgisine ihtiyaç duymasısıdır.

Sextant (Nikolaou vd., 2015) gibi mekansal bağlantılı verileri görselleştirmek için bazı yazılım paketleri ve araçlar olmasına rağmen, bunlar mekansal bilgi topluluğu tarafından henüz genel kabul görmemiştir. Burada kullanılan yaklaşım, sorgu sonuçlarını yaygın olarak kullanılan JSON formatında almak ve bunu Python ortamında bir mekansal veri çerçevesi olarak hazırlayarak web harita uygulaması için hazır hale getirmektir. Bu bağlamda, çok ölçekli temsil ayarının ya sorguda ya da web haritası uygulamasının katman bölmesinde manuel olarak yapılması gerekir. Bu nedenle, kullanıcının tercih ettiği yakınlaştırma düzeyine bağlı olarak, detayların ilgili ölçeklerdeki temsillerinin otomatik olarak harita üzerinde görüntüleneceği çok ölçekli bir ortam gereklidir.

UHK’ların bir ÇTMVT’yi sürdürme niyetinin ana nedenlerinden biri, güncellemeleri ölçekler arasında yaymaktır. Çalışmada, GeoSPARQL sorguları ile çok ölçekli mekansal bağlantılı veriler görselleştirilebilirken, güncelleme yayılımı hakkında herhangi bir uygulama yapılmamıştır. Mekansal semantik web ortamında belirli periyotlarda otomatik veya manuel olarak VGI tabanlı yetkili verilerin güncellenmesi olasılığı araştırmacılara açık bir konudur.

Ontolojilerdeki mekansal detayların kardinal ilişkilerinin modellenmesi, Egenhofer ve Mark (1995)’ın “topoloji önemlidir, metrik iyileştirir” (“topology matters, metric refines”) ifadesiyle örtüşmektedir. Bu bağlamda, mekansal verilere daha fazla anlam eklemek, makineler tarafından daha iyi çıkarım yapmak adına yeni olanaklar sunar.

4.3 Uygulama 2: Semantik Web Aracılığıyla Mekansal Veri Zenginleştirilmesi ve Binaların Ontoloji Tabanlı Sınıflandırılması

Artan mekansal verilerle birlikte mevcut verilerin entegrasyon yoluyla zenginleştirilmesi konuları da önem kazanmış ve araştırmacıların ilgisini çekmiştir. Burada popüler bir VGI uygulaması olan OSM öne çıkarak çeşitli uygulamalar geliştirilmesine katkıda bulunmuştur. Bu bağlamda, Codescu vd. (2012) OSM verilerini kullanarak ontoloji tabanlı bir rota planlaması uygulaması geliştirmiştir. Bu çalışma, klasik navigasyon uygulamalarından, rota üzerinde ihtiyaç duyulabilecek yerleri önermesi özelliğiyle farklılaşmaktadır. Fan, Zipf ve Fu (2014) OSM verilerini kullanarak morfolojik analiz yöntemleriyle bina türlerini tespit etme üzerine bir uygulama geliştirmiştir. Yazarlar, çalışma alanlarında %90'ın üzerinde bir doğruluk oranıyla konut alanlarını tespit etmişlerdir. Kunze ve Hecht (2015) bina verilerini OSM ile semantik olarak zenginleştirerek konut alanlarının nüfusunun daha hassas bir şekilde hesaplanmasını sağlamışlardır. Davidovic vd. (2016), 40 farklı şehirde OSM'de gönüllülerin oluşturduğu etiketleri OSM Wiki Harita Detayları kılavuzunda yapılan önerilere kıyasla incelemiş ve buradaki önerilere uygunluğun genellikle ortalama veya zayıf olduğunu tespit etmiştir. Fonte vd. (2018) bina kullanım türlerinin tespiti için OSM'nin yanında Facebook ve Foursquare uygulamalarını da kullanarak binaların %80'den fazlasının kullanım türünü tespit etmiştir. Bununla birlikte, literatürde OSM'nin veri kalitesinin artırılması ve kalite analizi ile ilgili çalışmalar da yer almaktadır. Vargas-Muñoz vd. (2019) kırsal bölgelerdeki bina etiketlerinin düzeltilmesi için Evrişimli Sinir Ağları (CNN) temelli bir yöntem önermiştir. Önerilen bu yöntemin, Tanzania ve Zimbabve'nin iki bölgesi için OSM etiket ve açıklama doğruluğunda önemli bir gelişme gösterdiğini tespit etmişlerdir. Nowak Da Costa (2016) OSM ile Polonya resmi topografik veri setinde yer alan bina detayları arasında semantik benzerlik analizi yaparak acil durum gibi uygulamalarda OSM verisinin resmi verilere tematik anlamda tamamlayıcı olup olamayacağını araştırmıştır. Bu analiz sonucunda entegrasyonun çeşitli acil durum senaryoları için umut verici olduğunu belirtmiştir. Son olarak, Basaraner (2020) OSM bina detaylarını 1:1 000 ölçekli topografik haritada bire-bir karşılık gelen bina detaylarıyla karşılaştırarak geometrik ve semantik doğruluk değerlendirmesini sunmuştur.

İfade edilen bu zenginleştirme çalışmalarının büyük bir kısmı resmi verilerin OSM verileriyle birlikte kullanılması üzerinedir. Resmi kurum mekansal verilerinin OSM verileriyle zenginleştirilmesi kurumlara zaman ve maliyet açısından da avantaj sağlayabilir. Resmi kurum verilerinin güncelleme sıklığı düşünüldüğünde, gönüllüler kendi yaşadığı veya iyi bildiği yerlerde OSM verisini daha hızlı güncelleyebilmektedir. Bu bağlamda, resmi kurumlar OSM'nin hızlı güncellenmesi durumunu avantaja çevirip güncellemeleri kendi veri setlerine aktarabilirler. Bununla birlikte, OSM'nin yaklaşık

ölçeğinden ($\sim 1:10\ 000$) daha küçük ölçeklerde üretilen mekansal veriler için OSM, öznitelik bağlamında bütünleyici bir rol oynayarak bazı detayların eksik özniteliklerini tamamlayabilir. Böyle bir zenginleştirme çalışması için de entegrasyon kilit rol oynamaktadır. Bu uygulamada çok kaynaklı ve çok ölçekli bina verileri OSM verileriyle geometrik olarak eşleştirilecek ve semantik web aracılığıyla zenginleştirilerek ontoloji tabanlı bir sınıflandırma gerçekleştirilecektir. Bu amaçla, OSM'deki bina kullanım türleriyle ilgili bilgiler ayıklanacak ve ontolojiler yardımıyla oluşturulan ilgili sınıflara tanımlanacaktır. Bu şekilde birbiriyle ilişkilendirilmiş farklı veri setlerindeki bina detayları OSM verileri yardımıyla ayrıntılı bir şekilde sınıflandırılabilir.

4.3.1 Kullanılan Veriler

Bu uygulamada, önceki uygulamada olduğu gibi, üç farklı kaynaktan gelen farklı ayrıntı düzeylerindeki bina veri setleri (TOPO1B, OSM ve TOPO25B) ve buna ek olarak OSM ilgi noktaları (POI) kullanılmıştır. Fakat önceki uygulamadan farklı olarak verilerin tamamı değil, sadece OSM ile bire-bir ilişkili TOPO1B binaları ve bunlar ile ilişkili TOPO25B münferit binaları ve OSM POI detayları seçilmiştir. Bunun nedeni, yapılacak sınıflandırmanın veri setleri arasında da kullanılabilir olmasını sağlamaktır. Örneğin, bire-çok ve çok-çok ilişkisinde yer alan bina detaylarının sınıflandırılması ve bunun veri setleri arasındaki ilişkisi daha karmaşık olacağından bu durumdaki binalar elenmiştir. Veri seçimi ile ilgili ayrıntılar sonraki bölümde ele alınmıştır.

4.3.2 Yöntem

Bire-bir eşleşen binaların tespit edilebilmesi için önceki uygulamadan farklı olarak Denklem 4.5'te belirtilen O_{KA-max} değeri kullanılmıştır. Bu denklem ile alanları kesişen iki binanın kesişim alanı binalardan büyük olanın alanına oranlanarak bir değer elde edilmektedir. Elde edilen bu O_{KA-max} değeri %60'dan fazla ise bu iki bina bire-bir eşleşen binalar olarak kabul edilmiştir. Literatürde aynı binaların tespiti için kullanılan bu yöntemde eşik değer %50-%70 aralığında değişebilmektedir (Basaraner, 2020; Erden, 2019). Bu çalışma için yapılan ön değerlendirme ve ampirik analiz sonuçlarına göre %60 eşik değeri bu bölge için yeterli görülmüştür.

$$O_{KA-max} = \frac{Alan_{kesisen}}{\max(Alan_{ref_i}, Alan_{hdf_i})} \quad (4.5)$$

Burada $Alan_{kesisen}$ iki binanın kesişen alanının, ref_i referans veri setinde yer alan bina alanının ve hdf_i ise hedef veri setinde yer alan bina alanının büyüklüğünü göstermektedir ($i : 1...n$).

Bununla birlikte OSM’de yer alan etiket değerlerine göre bir sınıflandırma çalışması yapılacağından OSM’deki bina detaylarından da en az bir etiketi olan binalar seçilmiştir (sınıflandırmada kullanılmak üzere belirlenen etiketler için). Burada building=’yes’ etiketi hariç tutulmuştur çünkü bu etiket binaların %80’den fazlasında yer almaktadır ve çalışmada zaten bina verileri kullanıldığından bu etiketin kullanılmasına gerek yoktur. Kısacası bu çalışmada kullanılacak veriler için aşağıdaki kriterleri sağlayan bina detayları seçilmiştir:

- O_{KA-max} değeri %60’tan büyük olanlar,
- En az bir etiketi olan OSM bina detayları (building=’yes’ hariç)

Ayrıca OSM bina alanları içine giren OSM POI’leri de tespit edilerek kullanılacak verilere eklenmiştir. Yukarıda yer alan kurallara uygun binalar Python kodları yardımıyla seçilerek GeoJSON formatında kullanıma hazır hale getirilmiştir. Seçilen bina detaylarına ilişkin bilgiler Tablo 4.8 ve Şekil 4.14’te yer almaktadır.

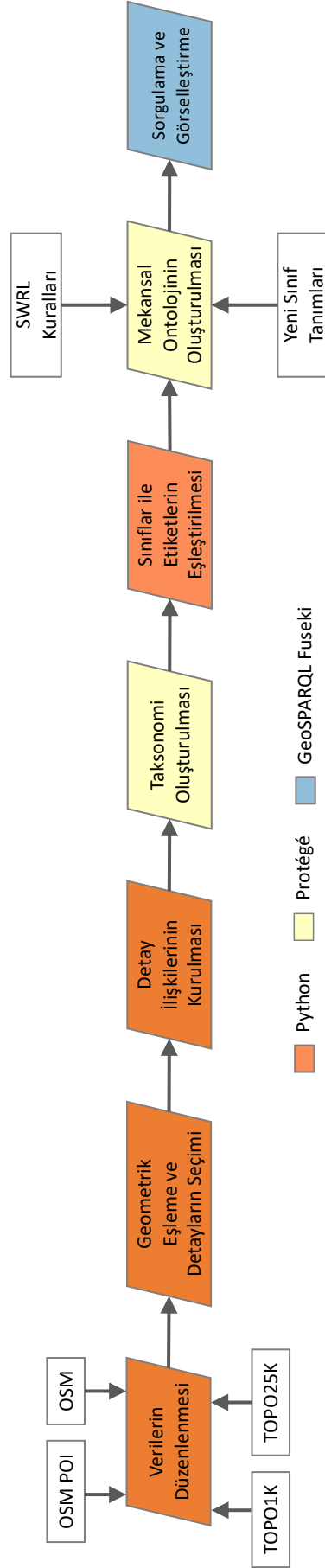
Tablo 4.8 Seçilen bina detay sayıları

Veri Seti	Detay Sayısı
OSM	1075
TOPO1B	1075
OSM POI	1512
TOPO25B	81



Şekil 4.14 Seçilen bina detaylarının gösterimi

Veri seçiminin ardından uygulamada kullanılacak bina detayları ve ilgili diğer verilerin RDF'e dönüştürülerek bir ontoloji haline getirilmesi gerekmektedir. Detay seçimi yapılırken %60 kesişim oranlı O_{KA-max} değeri kullanıldığından birbiriyle ilişkili detayların da tespiti yapılmıştır. Ardından, ontolojide bu detayların ilişkilerinin tanımlanması gerekmektedir. Bunun için önceki uygulamada kullanılan mekansal ontolojinin temelleri (bireyler olmadan sadece tanımların olduğu kısım) kullanılmış ve orada yer alan ilişki tanımları ile bu uygulamada oluşturulan detaylar birbirleriyle ilişkilendirilmiştir. Seçilen TOPO1B ve OSM bina detayları arasında mekansal ontolojide tanımlanmış olan sameAs (aynısıdır) ilişkisi kurulmuştur. OSM bina detayları ile OSM POI'leri arasında hasRelation (ilişkisiVar) ilişkisi ve son olarak da OSM ile TOPO25B münferit bina detayları arasında yine sameAs ilişkisi kurulmuştur. Görüldüğü üzere OSM tüm veri setlerinin ortak noktasıdır. Semantik web kuralları uygulandığında doğrudan ilişki kurulmayan ilgili tüm veri setleri arasında uygun ilişkiler çıkarım yoluyla elde edilecektir. Bu bağlamda, uygulamanın işlem adımları Şekil 4.15'te verilmiştir.



Şekil 4.15 Uygulamanın işlem adımları

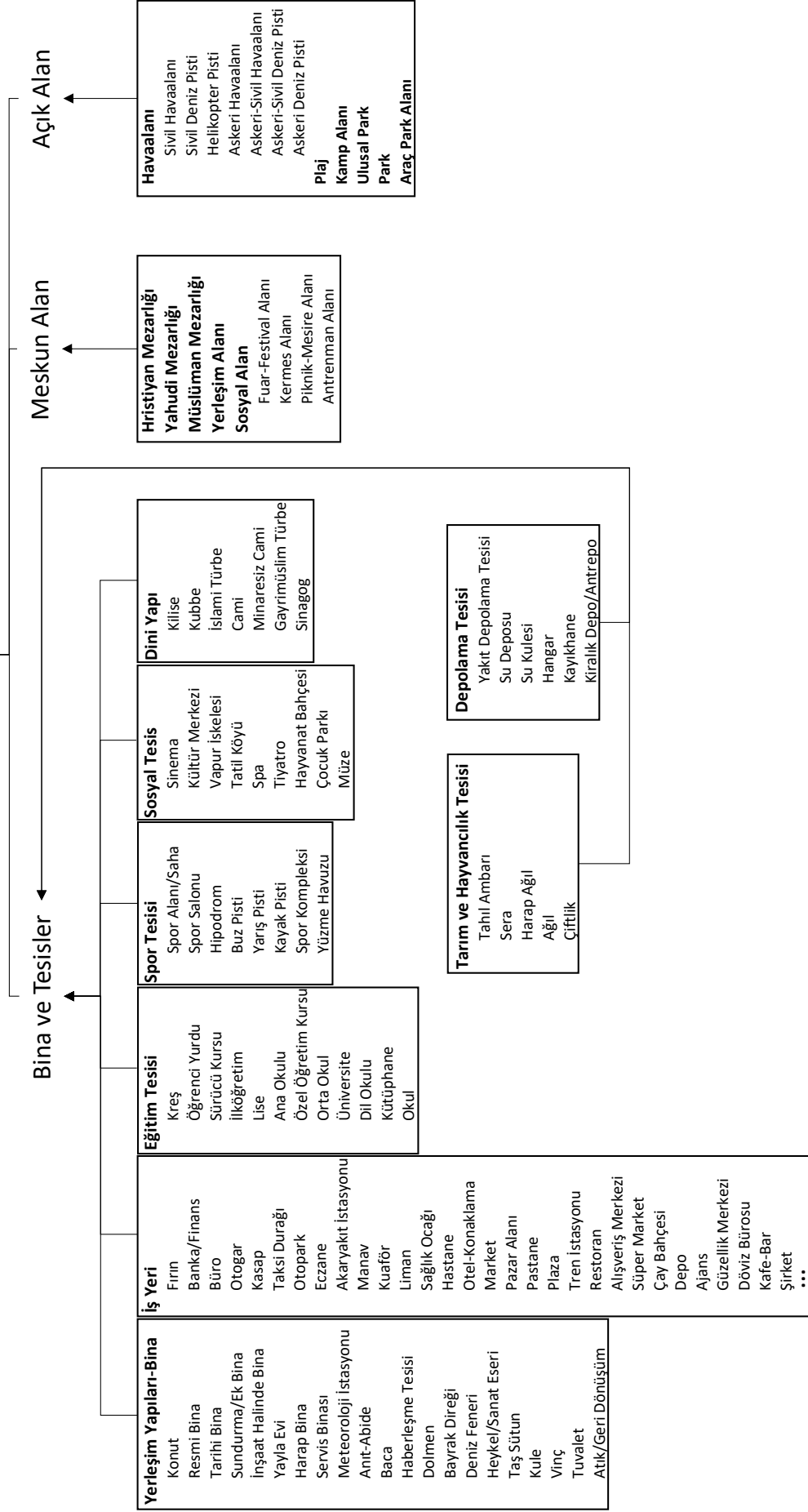
Daha önce de ifade edildiği gibi OSM’de gönüllüler özgürce veri tanımlamaları yapabilmektedir. Bunu yaparken anahtar=değer (key=value) şeklindeki etiket çiftlerini kullanırlar. OSM Wiki sayfasında yaygın etiketler ile ilgili gönüllülere yönelik bir rehber yer alsa da gönüllüler çoğu zaman bu rehberin dışına çıkarak özgün etiketler kullanırlar. Bu uygulamadaki amaçlardan biri, gönüllülerden gelen verilerle zenginleştirme yoluyla bina kullanım türlerini belirlemek olduğundan belirli etiketlerin seçilerek incelenmesi gerekmektedir. Ayrıca OSM’de yer alan bu etiket değerleri bir kent bilgi sistemindeki tematik coğrafi detayların özneteliklerini elde etmeye katkı sağlayabilir. Bu amaç için aşağıda sıralanan sekiz adet anahtar değeri seçilerek bu anahtarlar için gönüllülerin tanımladığı değerler kullanılmıştır. Örneğin, binalarda en çok kullanılan ‘building’ anahtarı için gönüllüler ‘house’, ‘apartment’, vb. değerler tanımlamışlardır. Bu bağlamda OSM ve OSM POI veri setleri, anahtarlar sütun başlığı değerler de satırlar olacak şekilde düzenlenmiştir. Ardından Python kodları yardımıyla mekansal ontolojiye hasTagVal ilişkisiyle eklenmiştir.

- *building*
- *leisure*
- *historic*
- *tourism*
- *amenity*
- *shop*
- *office*
- *healthcare*

Önceden oluşturulmuş olan mekansal ontolojide Şekil 4.9’da yer alan genel sınıf tanımlamaları olmasına rağmen bu uygulama için daha ayrıntılı sınıf tanımlarının yapılması gerekmektedir. Bunun için, yayınlanan Türkiye Ulusal Coğrafi Bilgi Sistemi (TUCBS) belgelerinden, Büyük Ölçekli Harita ve Harita Bilgileri Üretim Yönetmeliği’nden (BÖHHBÜY), çeşitli belediyelerin oluşturduğu Kent Bilgi Sistemi (KBS) belgelerinden ve Basaraner (2013)’ün çalışmasından faydalanarak mekansal ontolojinin ilgili yerleşim sınıfı altında yeni bir sınıflandırma (taksonomi) oluşturulmuştur (Şekil 4.16).

OSM etiket değerlerinden faydalanarak verilerin Şekil 4.16’da yer alan ontoloji sınıfları altına yerleştirilebilmesi için etiket değerleriyle ontoloji sınıflarının

Yerleşim



Şekil 4.16 Mekansal ontolojinin yerleşim sınıfı altında oluşturulan ayrıntılı sınıflar

eşleştirilmesi gerekmektedir. Örneğin, ‘apartments’ değerine sahip bir detayın ontolojide yer alan ‘Konut’ sınıfı altında yer alması gerekmektedir. Bunun için eşleştirme tabloları oluşturularak etiket değerleri uygun ontoloji sınıflarıyla manuel olarak eşleştirilmiştir. Ayrıca ontolojinin sadece burada yer alan verilere özgü olmaması ve daha genel bir şekilde diğer OSM bölgelerine de hizmet edebilmesi açısından, TagInfo adlı web sitesinden yararlanarak etiket değerlerinin kapsamı arttırılmıştır. TagInfo, OSM’deki tüm etiket istatistiklerinin yer aldığı ve belirli aralıklarla güncellenen bir web sitesidir (TagInfo, 2021). Bu siteden yararlanarak belirli kurallara uyan etiket değerleri de çalışmaya eklenmiş ve daha kapsayıcı bir hale getirilen sınıflandırmanın OSM’nin genelinde kullanılabilmesinin önü açılmıştır. Bu amaçla aşağıdaki iki kurala uyan etiket değerleri seçilmiştir.

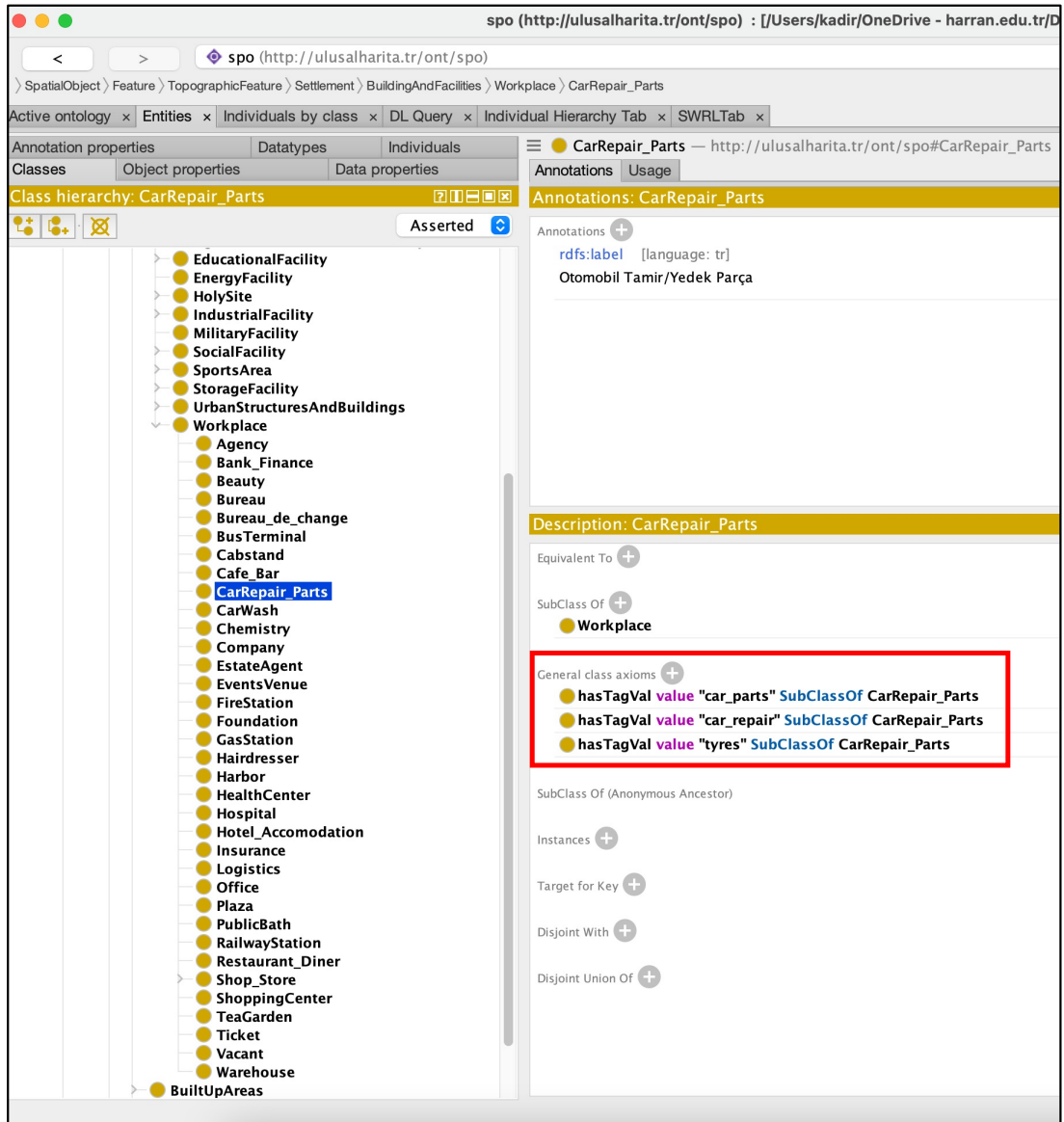
- *Her bir anahtar için, etiket değerleri tüm değerlerin en az %90’ını kapsamalıdır.*
- *Çok kullanılan ‘building’ ve ‘amenity’ anahtarları için en az 50, diğer anahtarlar için en yüksek kapsama oranına sahip en az 10 adet etiket değeri seçilmelidir.*

Bu bağlamda, ilk kural ile tüm OSM kullanıcılarının bir anahtar için girdiği etiket değerlerinden en az %90’ı çalışmaya dahil edilmiştir. İkinci kural ile de yüksek oranlı bir kaç adet etiket değeriyle %90 oranının tamamlanması ihtimaline karşın bazı anahtarlar için en az 50 adet etiket değerin çalışmaya dahil edilmesi sağlanmıştır. Yukarıdaki kurallara uyan etiket değerlerinin ve bunların ontoloji sınıfları ile eşleştirmelerinin yer aldığı örnek Tablo 4.9’da, tüm tablolar ise Ek C’de verilmiştir.

Seçilen sekiz adet anahtarın tamamı için elde edilen bu tablolarda, oluşturulan OSM/OSM POI sınıfları ve bunların mekansal ontolojideki eşlenikleri yer almaktadır. Bir Python kodu yardımıyla tablolardaki etiket değerleri ve eşlenmiş olan sınıfları mekansal ontolojiye aktarılmıştır. Şekil 4.17’de tanımlanan etiketlerden bir sınıf için olanı görülebilir.

Tablo 4.9 Leisure anahtarı için oluşturulan kapsam ve sınıf eşleniği tablosu

Sayı	Anahtar = leisure Değer	Kapsam (%)	Sınıf Eşleniği
1	pitch	27,47	Pitch
2	swimming_pool	20,36	SwimmingPool
3	park	15,39	Park
4	garden	10,96	Park
5	playground	9,65	Park
6	sports_centre	3,03	SportsCentre
7	picnic_table	2,50	PicnicArea
8	nature_reserve	1,56	NationalPark
9	track	1,37	Park
10	fitness_centre	0,77	SportsCentre
11	sauna	0,19	PublicBath
12	ice_rink	0,10	IceSkatingArea
13	turkish_bath	0,10	PublicBath
Toplam Kapsam =		%93,45	



Sınıflandırma işleminin yapılabilmesi için son olarak OSM POI'lerinden elde edilen etiket değerlerinin ilişkili olduğu OSM detaylarına aktarılması gerekmektedir. Bunun nedeni OSM POI bilgilerinin OSM detayları için yardımcı olarak kullanılacak olmasıdır. Kısacası OSM POI'leri doğrudan sınıflandırmaya katılmayacak, buradan elde edilen etiket değerleri ilişkili olduğu OSM detaylarının sınıflandırılmasına yardımcı olacaktır. Bu sayede OSM detaylarında var olmayan bazı bilgiler OSM POI'lerinden elde edilecektir. Bu işlem için aşağıda yer alan SWRL kuralı kullanılarak ilgili OSM POI değerleri ilişkili OSM detaylarına aktarılmıştır.

$$OSM_POI(?x) \wedge hasRelation(?x,?y) \wedge hasTagVal(?x,?z) \Rightarrow hasTagVal(?y,?z)$$

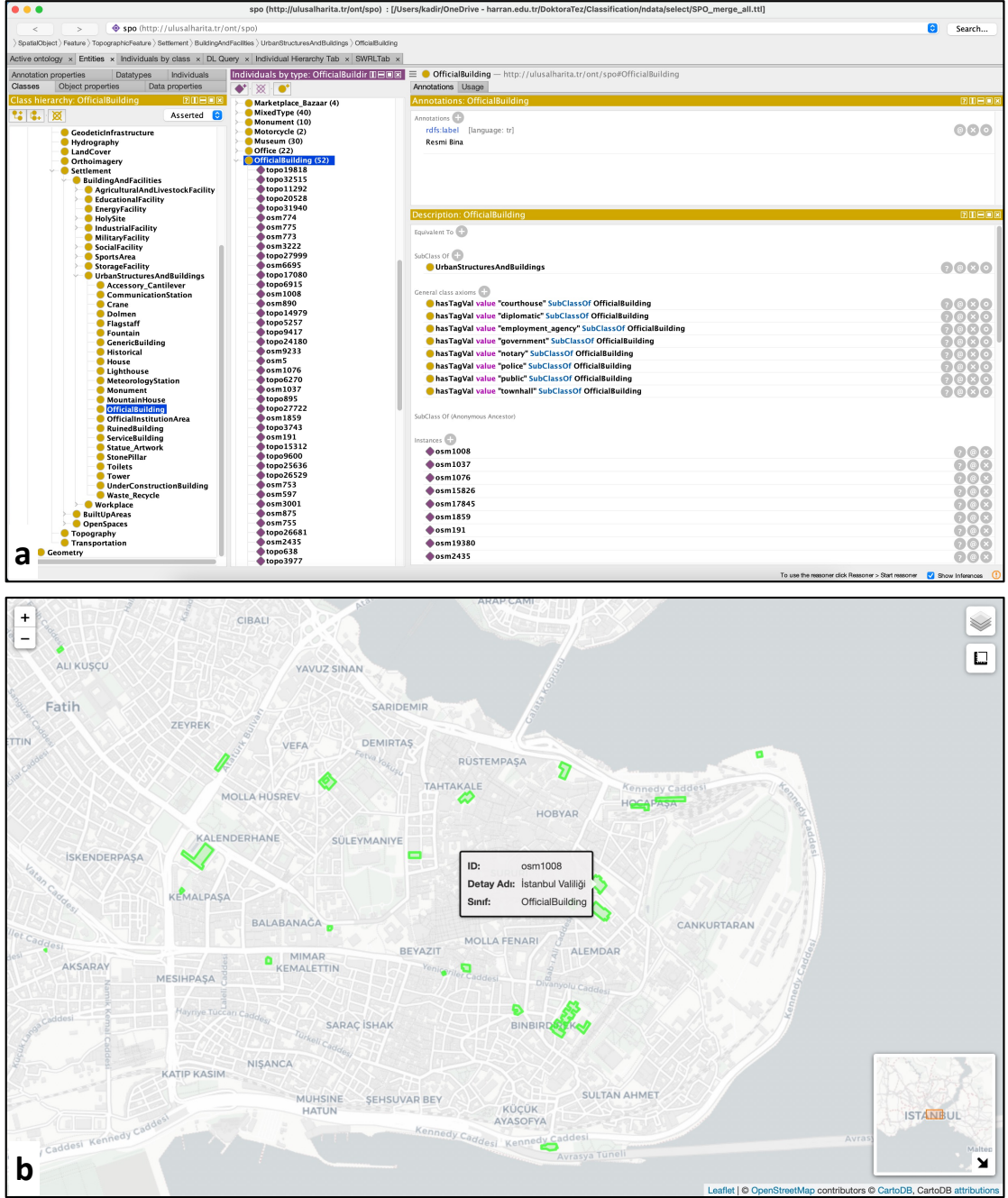
Yukarıda yer alan SWRL kuralının uygulanmasının ardından mekansal ontoloji artık semantik web yardımıyla sınıflandırılmaya ve kural tanımlamaları yapılmaya uygun hale gelmiştir. OWL 2 RL temelli çıkarım motoru çalıştırılarak OSM detayları ontolojide tanımlanan sınıflar ile sınıflandırılmıştır. Örneğin, car_parts veya car_repair etiket değerlerini içeren OSM detayları ontolojideki CarRepair_Parts sınıfı altına yerleşmiştir (Şekil 4.18).

Böylece ontolojide yer alan OSM detayları belirlenmiş olan ontoloji sınıfları altına yerleştirilmiştir. Fakat bu işlem ile üç farklı bina veri setinden sadece OSM sınıflandırılmıştır ve diğer iki veri setinin de sınıflandırılması gerekmektedir. TOPO1B ve TOPO25B veri setleri ise iki farklı şekilde: (1) sameAs ilişkisi ile bağlı olduğu OSM verileri yardımıyla SPARQL sorguları kullanılarak veya (2) OSM POI etiket değerlerinin OSM'ye aktarıldığı gibi OSM etiket değerlerinin TOPO1B ve TOPO25B veri setlerine de aktarılması ve çıkarım motorunun bunlar için de çalıştırılması yoluyla sınıflandırılabilir. Burada ikinci yöntem kullanılmış ve TOPO1B ve TOPO25B veri setleri de OSM gibi sınıflandırılarak mekansal ontolojiye kaydedilmiştir.

4.3.3 Bulgular

Bu uygulama sonucunda OSM POI değerleriyle zenginleştirilmiş OSM bina detayları sınıflandırma amacıyla kullanılmıştır. Bu amaç için, OSM ve OSM POI'lerden elde edilen 235 farklı etiket değeri, ontolojide yer alan 122 farklı taksonomi sınıfıyla eşleştirilmiştir. Bununla birlikte OSM'de yer alan 1075 adet detaydan 994'ü (yaklaşık %93'ü) en az bir ontoloji sınıfı altında tanımlanabilmiştir. Ardından, semantik web yöntemleri kullanılarak OSM için gerçekleştirilen bu sınıflandırma, TOPO1B ve TOPO25B veri setleri için de gerçekleştirilmiştir. Örnek bir sorgu olarak, mekansal ontolojide yer alan "Resmi Bina" sınıfı ve altında yer alan bina detaylarıyla ilgili ontoloji ve web haritasının bir görüntüsü Şekil 4.19'da yer almaktadır.

Ontolojinin sorgulanması ve sonuçların görselleştirilmesi aşamasında GeoSPARQL Fuseki, SPARQLWrapper ve Folium yazılımları kullanılmıştır. Mekansal ontoloji GeoSPARQL Fuseki ile HTTP üzerinden sorgulanmaya hazır olarak web üzerinden sunulmuştur. Ardından, Python ortamında sorguların HTTP istekleriyle istendiği ve alındığı SPARQLWrapper yazılımıyla sorgulanmıştır. Son olarak sorgu sonuçları Python ortamında düzenlenerek Folium kütüphanesi yardımıyla web haritası şeklinde gösterilmiştir.



Şekil 4.19 Resmi Bina sınıfı ayrıntılarını gösteren ontoloji (a) ve web haritası görüntüsü (b)

Elde edilen, POI verileriyle zenginleştirilmiş ve birbiriyle ilişkilendirilmiş olan mekansal ontolojiyle çeşitli mantıksal kurallar ile yeni sınıflandırmalar ortaya konabilir. Örneğin, bina detayları için konut + işyeri kullanım türünde olan karma binalar tespit edilebilir. Bunun için aşağıdaki mantıksal kural tanımlaması yapılarak yine çıkarım motoru çalıştırılır ve yeni sınıf altındaki bireyler tespit edilir.

- $KarmaBina \equiv Konut \sqcap İşyeri$

Yukarıda yer alan tanımlama ile oluşturulan sınıf Şekil 4.20’de yer almaktadır. Kurala uyan binalar bu yeni sınıf altında ‘KarmaBina’ olarak sınıflandırılmıştır.

Benzer şekilde, birden fazla sınıfın birleşiminden meydana gelen yeni sınıf tanımlamaları da yapılabilir. Aşağıda, bunlarla ilgili iki örnek görülebilir.

- $YıkılmışDiniTesis \equiv DiniTesis \sqcap Yıkılmış (1 \text{ adet yapı})$
- $TarihiDiniTesis \equiv Tarihi \sqcap DiniTesis (2 \text{ adet yapı, Örn. Sinan Paşa Türbesi})$

Sınıflandırılmış mekansal ontolojiden elde edilebilecek çok fazla yeni bilgi vardır. Burada sadece bir kısmı gösterilebilmiş olsa da amaca yönelik olarak yeni farklı sınıflar ve bilgiler elde edilebilir. Örneğin, bina detayları daha genel bir sınıf altında gösterilebilir. Mekansal ontoloji alt sınıf ve üst sınıf şeklinde bir taksonomi ortaya koyduğundan, bu detaylar ile ilgili bir üst sınıf gösterimi de gerçekleştirilebilir. TOPO1B detaylarının üst sınıfları ve bu sınıflara göre oluşturulmuş tematik harita Şekil 4.21’de yer almaktadır.

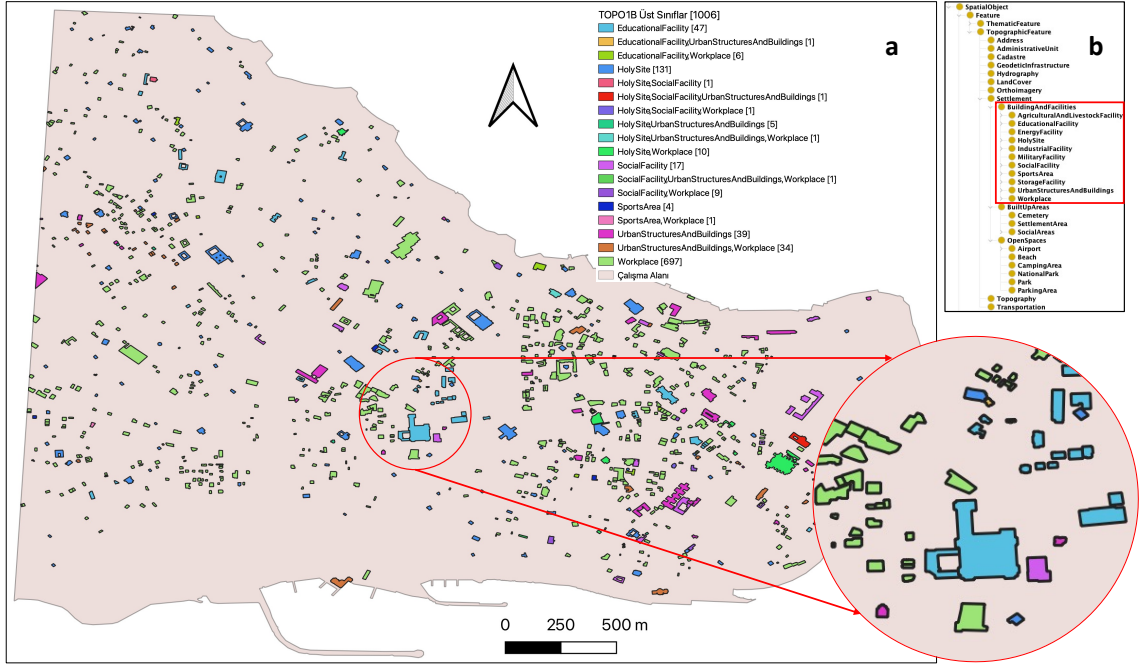
4.3.4 Tartışma

Bu bölümde farklı veri setlerinin entegre edilerek mekansal semantik web yöntemleriyle zenginleştirilmesi ve ardından ontoloji tabanlı sınıflandırılması aşamalarında ortaya çıkan zorluklar, uygulamanın avantajları ve sınırlılıkları sonuçlarıyla birlikte değerlendirilmektedir.

Bu çalışma ile Kent Bilgi Sistemi gibi bir tasarımda eksik olan semantik bilgilerin bir kısmı OSM tarafından elde edilebilir. Örneğin, hava fotoğrafları ile sadece bina sınırlarının elde edildiği bir projenin ilk aşamasında saha çalışmasıyla elde edilmesi gereken binaların semantik bilgilerinin bir kısmı buradan tamamlanabilir. Bununla birlikte, TOPO25B veri setinde tarihi camiler, pazarlar vb. yerler diğer binalardan karakteristik olarak farklı olduğundan meskun alanların dışında



Şekil 4.20 KarmaBina tanımlaması ile elde edilen binaları gösteren ontoloji (a) ve web haritası görünümü (b)



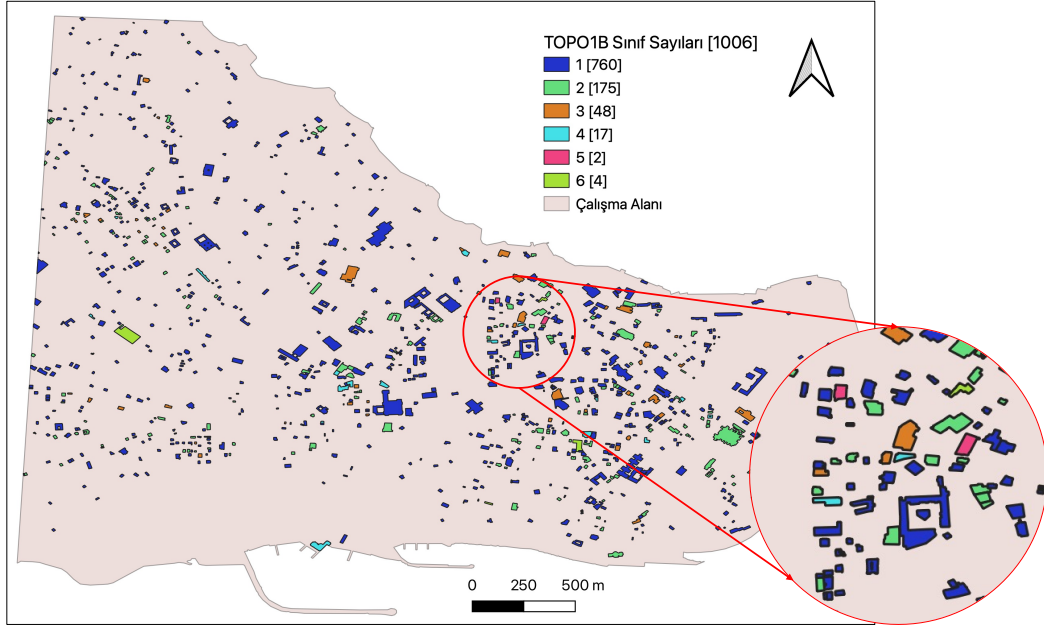
Şekil 4.21 Ontolojideki üst sınıflar (b) ve buna göre oluşturulmuş tematik harita (a)

münferit bina şeklinde gösterilmiştir. Daha küçük ölçekte yer alan ve semantik ayrıntısı az olan bu binaların bazı öznetelikleri ilişkili olan OSM veya TOPO1B binalarından tamamlanabilir. Bu çalışmada yer alan 81 adet TOPO25B münferit binası OSM detaylarıyla entegre edilerek zenginleştirilmiş ve ayrıntılı olarak oluşturulan taksonomiye göre sınıflandırılmıştır.

Diğer taraftan yapılan sınıflandırma, gönüllülerin oluşturduğu verilere dayandığından doğruluğu denetlenmemiş bilgiler içermektedir. Veri setlerinde, oluşturulan sınıflandırmaya dair herhangi bir bilgi bulunmadığından sınıflandırmanın doğruluğu bu çalışmanın konusu dışında bırakılmıştır.

Yapılan sınıflandırmada çok sayıda detay birden fazla sınıfa dahil olduğundan sınıflandırmaları kategorize ederek göstermek zorlaşmaktadır. Şekil 4.21’de görüldüğü üzere 11 üst sınıf altında gruplanmak istenen bina detayları birden çok sınıfa giren detayların varlığından dolayı 18 sınıf altında kategorize edilmiştir. Binalar doğal olarak birden fazla kullanım türüne sahip bölümlerden oluşabilmektedir. Bu durum, birden fazla sınıfa giren binaların ortaya çıkmasına sebep olmaktadır. Farklı kullanım türüne sahip binaları tek bir sınıf altında toplamak için doğrudan bir çözüm olmamasına rağmen, örtüşen sınıflar için ağırlık veya öncelik belirlenerek bu sorunun üstesinden bir ölçüde gelinebilir. Bu bağlamda, TOPO1B detaylarının sınıf sayılarını ve dağılımını gösteren harita Şekil 4.22’de görülmektedir. Birden çok sayıda sınıfa giren bazı detaylar incelendiğinde bunların ya konutla birlikte birden fazla iş yerinin yer aldığı binalar ya da birçok farklı türde işyerinin birlikte yer aldığı iş

merkezleri/işhanları gibi yerler olduğu görülmüştür.



Şekil 4.22 TOPO1B detaylarının sınıf sayılarını ve dağılımlarını gösteren harita

Yukarıda yer alan şekillerden de anlaşıldığı üzere ontoloji İngilizce dilinde oluşturulmuştur. Bu nedenle uygulamanın bazı kısımlarında bazı tanımlamalar İngilizce olarak sunulmuştur. Fakat ontolojilerde farklı dillerin kullanılması diğer veri yapılarına kıyasla çok daha kolaydır. Ontolojilerde her kaynak için yapılabilen çeşitli tanımlama ve açıklamalar farklı dillerde ve birden çok olabilir. Ontolojilerin birden çok dilde aynı dosyada tanımlanabilmesi olanağı, entegrasyonun önündeki önemli bir engel olan dil probleminin aşılmasında önemli bir avantaj sağlamaktadır. Bu uygulamada kullanılan ontolojideki tüm kaynaklar görüntüde İngilizce olmasına rağmen etiket (label) tanımları Türkçe olarak yapılmıştır. Protégé yazılımında yer alan bir seçenek ile tüm kaynakların Türkçe dilinde gösterilmesi sağlanabilmektedir.

Seçili etiketler için dünya genelinde kullanılan OSM değerlerinin %90'dan fazlasını kapsayan bu ontoloji ile dünyanın herhangi bir yerine ait OSM verileri mekansal ontolojideki tanımlarla sınıflandırılabilir. Semantik tanımlamalarla elde edilen sınıflardan farklı diğer sınıflar da elde edilebilir. Bununla birlikte mantıksal tanımlamaların burada kullanılan yazılımlar ve/veya donanımlar özelinde bir sınırı vardır. Yapılan mantıksal tanımlamalar (ne kadar basit olursa olsun) veya detay sayıları arttığında, ontolojide yer alan aksiyom (belit) sayısı da fazlaştığından (Çalışmadaki max. aksiyom sayısı: 278042) çıkarım motoru ile yapılması gereken işlemler gerçekleştirilememektedir (Kullanılan sistem özellikleri: 2,7 GHz Intel Core i5-5257U CPU, 8 GB bellek (RAM), 256 GB'lık bir katı hal sürücüsü (SSD) ve 64 bit macOS). Bu nedenle çalışmada sınırlı sayıda veri kümesi kullanılmıştır.

Bu tez kapsamında coğrafi bilgi sistemi ve mekansal semantik web teknikleriyle çok temsilli ve çok kaynaklı mekansal verilerin entegrasyonu ve zenginleştirilmesi konusu incelenmiş ve bu tekniklere dayalı bir yöntem geliştirilerek çeşitli bina veri setleriyle gerçekleştirilen iki uygulama sunulmuştur.

Birinci uygulamada, mekansal veri entegrasyonu ve çok kaynaktan gelen farklı ölçeklerdeki bina veri setlerinin zenginleştirilmesi için geometrik eşleştirme ve semantik web tabanlı bir yaklaşım sunulmuştur. Bu amaçla, kesişen alanların oranına (O_{KA}) dayalı olarak çok kaynaklı ve çok ölçekli üç farklı mekansal veri setine bir geometrik eşleştirme yöntemi uygulanmıştır. O_{KA} 'nın eşik değerinin belirlenmesi için ayrıca bir analiz yapılarak bu değer %35 olarak belirlenmiş ve eşleşme sonuçları, %90'ın üzerinde kesinlik ve duyarlılık değerleriyle manuel eşleştirme ile doğrulanmıştır. Ayrıca, kartografik ölçek desteği ve ölçekler arasında karmaşık kardinal ilişkilerini tanımlayan bir sözlük ile Ulusal Mekansal Veri Altyapısından ilgili detay sınıflarını içeren bir mekansal ontoloji tanımlanmıştır. Üç veri setinden mekansal detaylar ve ölçekler arasındaki detay ilişkileri, RDF veri modeline dönüştürülmüş ve daha sonra özel Python betikleri ile mekansal ontolojiye eklenmiştir. Farklı ölçeklerdeki mekansal detaylar arasında daha fazla bağlantı bulmak için Protégé yazılımında SWRL kuralları oluşturularak uygulanmıştır. Ek olarak, mekansal detaylar, popüler bilgi tabanları olan DBpedia ve YAGO ile ilişkilendirilmiştir. Sonuç olarak, oluşturulan mekansal ontoloji GeoSPARQL uyumlu bir HTTP sunucusu aracılığıyla sunulmuştur. Bu entegre edilmiş mekansal bağlantılı verilerin faydalarının gösterilmesi için üç örnek sorgu oluşturulmuş ve sonuçlar web harita uygulaması üzerinde görselleştirilmiştir. Ek olarak, geleneksel mekansal veri modellerinden web üzerinde mekansal bağlantılı veri sunmanın zorlukları tartışılmıştır. Sonuç olarak, bu çalışma ontolojileri kullanarak farklı mekansal veri setlerini bir bilgi tabanı olarak entegre etmenin farklı kullanıcılardan gelen soruları yanıtlamayı kolaylaştırabileceğini göstermiştir. Bu bağlamda, bir turist, bir engelli veya bir haritacı, bu çalışmada oluşturulan mekansal bilgi tabanında belirli sorularına cevap bulabilmektedir. Adına

Büyük Veri diyebileceğimiz yaşadığımız bu çağda, farklı kaynaklı ve ölçekli verileri entegre ederek değerli bilgiler elde etmenin, farklı gereksinimleri olan kişilerin coğrafi bilgiye erişimini kolaylaştırmak için umut verici bir yaklaşım olduğu söylenebilir.

İkinci uygulamada, üç farklı kaynaktan elde edilen ve OSM POI'lerini de içeren üç farklı ölçekteki veri setleri geometrik eşleştirme kullanılarak entegre edilmiş ve zenginleştirilmiştir. Ardından mekansal semantik web yardımıyla binaların kullanım türüne göre bir sınıflandırma gerçekleştirilmiştir. Bu amaç için mekansal ontolojide yer alan yerleşim sınıfına TUCBS, BÖHHBÜY ve çeşitli KBS dokümanlarından yararlanarak alt sınıflar oluşturulmuştur. Böylece yerleşim sınıfı altında kapsamlı bir taksonomi ortaya konmuştur. OSM'de en çok kullanılan etiket değerleri TagInfo web sitesi yardımıyla tespit edilmiş ve bu değerler uygun ontoloji sınıfları ile manuel olarak eşleştirilerek bir Python betiği aracılığıyla mekansal ontolojiye eklenmiştir. Veri setlerinin, sınıf tanımlarının ve bireyler arası ilişkilerin RDF formatında tanımlı olduğu bu mekansal ontolojide, OWL 2 RL temelli bir çıkarım motoru çalıştırılarak istenen sınıflandırma elde edilmiştir. Bu kapsamda kullanılan verilerin %93'ü sınıflandırılmıştır. Ayrıca, semantik web kuralları aracılığıyla sınıflandırmada kullanılan bilgiler diğer veri setlerine aktarılarak onların da sınıflandırılması gerçekleştirilmiştir. Mantıksal kurallarla var olan sınıflar üzerinden yeni sınıf tanımlamaları yapılarak yeni bilgiler ortaya çıkarılmıştır. Yapılan tanımlar veya sorgular sonucunda ortaya konan mekansal veriler web haritası uygulamasıyla görselleştirilerek sunulmuştur. Oluşturulan bu mekansal ontoloji, sınıflandırma için kullanılan etiketlerden dünya genelinde kullanılan değerlerin %90'ından fazlasını içerdiğinden farklı bölgelerin de aynı şekilde sınıflandırılabilmesine olanak sağlanmıştır. Ayrıca bu uygulama, sadece geometrilerin mevcut olduğu bazı projelerde çeşitli semantik bilgilerin elde edilebilmesi için de entegrasyon yoluyla zenginleştirme adına bir yöntem sunmaktadır.

Bununla birlikte, her iki uygulamada da bir gönüllü coğrafi bilgi uygulaması olan OSM verileri kullanıldığından doğruluğu denetlenmemiş bilgiler içerebilmektedir. Haritaların temelde tam ve doğru olması gerektiğinden buradan gelecek verileri harita oluşturmak için kullanırken dikkat edilmesi gerektiği göz önünde bulundurulmalıdır. OSM verilerinin doğruluğu ve güvenilirliği bu geliştirilen uygulamalar ve gönüllülerin her geçen artan katkısıyla zamanla daha ileri seviyeye taşınacaktır.

Gelecek çalışmalarda, ilk uygulama için ilişkisel modeldeki mekansal verileri (örn, shapefile) arka plan bilgisi gerektirmeden RDF'ye dönüştüren, veriden bağımsız bir paket veya araç geliştirilmesi, ikinci uygulama içinse manuel eşleştirme ile sağlanan etiket değerleriyle sınıf eşleştirmelerinin genel bir sözlük yapısı kullanılarak anlamların da yer aldığı otomatik bir yöntemle gerçekleştirilmesi üzerine çalışılacaktır.

Semantik web teknikleri mantık ve kural tanımlamalarına olanak sağlasa da ayrıntılı mekansal analizler söz konusu olduğunda yetersiz kalmaktadır. Çoğu durumda, hesaplanması gereken veriler, istatistikler veya mekansal analizler semantik web'in dışında elde edilerek buraya aktarılmaktadır. Bu aşamada semantik web, CBS gibi olgun sistemlere bir alternatif değildir. Gelecekte kullanılan yazılım ve donanımların geliştirilmesiyle CBS'de yapılan işlemler burada da kolay bir şekilde gerçekleştirilebilir. O zamana dek, CBS ve mekansal semantik web birbirlerini tamamlayan iki yöntem olarak ilişkilerini sürdürecektir.

İlk olarak 1900'lü yılların ortalarında insan hayatına giren bilgisayarlar, ilk zamanlardan bu yana insanlar tarafından makine dilleri geliştirilerek kullanılmıştır. Bu durum bilgisayarları kullanmak isteyen diğer insanları da bu makine dillerini öğrenmeye itmiştir. Fakat son zamanlarda, yapay zeka, doğal dil işleme ve akıllı asistanlar gibi yöntem ve bileşenlerin gelişmesiyle bu durum tersine dönmüştür. İnsanlar bu süreçte bilgisayarlara insanların dilini öğretmeye başlamışlar ve bilgisayarların insanlarla bu şekilde iletişim kurabilmesiyle eskiden belirli bir kitlenin yapabildiği insan-makine etkileşimi çok geniş kitlelere hatta neredeyse internetin olduğu her yere yayılmıştır. İnsan ve makine etkileşiminin bu denli arttığı bir dönemde makinelere insanlar için üretilmiş verileri öğretmek yerine onların da anlayabileceği verileri üretmek akıllıca olacaktır. Bu bağlamda, semantik web bu boşluğu doldurarak, insan ve makine ilişkileri arasında bir köprü görevi üstlenerek bu ilişkiyi bir üst seviyeye taşıyabilecek potansiyele sahiptir.

- Ai, T., Cheng, X., Liu, P., & Yang, M. (2013). A shape analysis and template matching of building features by the Fourier transform method. *Computers, Environment and Urban Systems*, 41, 219–233. <https://doi.org/10.1016/j.compenvurbsys.2013.07.002>
- Aranda, C. B., Corby, O., Das, S., Feigenbaum, L., Gearon, P., Glimm, B., Harris, S., Hawke, S., Herman, I., Humfrey, N., Michaelis, N., Ogbuji, C., Perry, M., Passant, A., Polleres, A., Prud'hommeaux, E., Seaborne, A., & Williams, G. T. (2013). *SPARQL 1.1 Overview* (tech. rep.). W3C.
- Athanasidou, S., Giannopoulos, G., Graux, D., Karagiannakis, N., Lehmann, J., Ngomo, A. C. N., Patroumpas, K., Sherif, M. A., & Skoutas, D. (2019). Big POI data integration with linked data technologies. *Advances in Database Technology - EDBT, 2019-March*, 477–488. <https://doi.org/10.5441/002/edbt.2019.44>
- Baglioni, M., Masserotti, M. V., Renso, C., & Spinsanti, L. (2007). Building geospatial ontologies from geographical databases. In F. Fonseca, M. A. Rodríguez, & S. Levashkin (Eds.), *Geospatial semantics* (pp. 195–209). Springer Berlin Heidelberg. https://doi.org/10.1007/978-3-540-76876-0_13
- Basaraner, M. (2009). Çok Çözünürlüklü Mekansal Veri Tabanları. XI. *Akademik Bilişim Konferansı, Şanlıurfa*, 727–736.
- Basaraner, M. (2012). An Investigation of Semantic, Geometric and Graphic Heterogenities of Building and Facility Objects in A Multi-Resolution Spatial Database. *Proceedings of 4th International Conference on Cartography and GIS, Albena*.
- Basaraner, M. (2013). Taxonomies of Building Objects towards Topographic and Thematic Geo-Ontologies. *26th International Cartographic Conference, Dresden*.
- Basaraner, M. (2016). Revisiting cartography: towards identifying and developing a modern and comprehensive framework. *Geocarto International*, 31(1), 71–91. <https://doi.org/10.1080/10106049.2015.1041560>
- Basaraner, M. (2020). Geometric and semantic quality assessments of building features in OpenStreetMap for some areas of Istanbul. *Polish Cartographical Review*, 52(3), 94–107. <https://doi.org/10.2478/pcr-2020-0010>
- Battle, R., & Kolas, D. (2012). Enabling the Geospatial SemanticWeb with Parliament and GeoSPARQL. *Semantic Web Journal*, 3(4), 355–370. <https://doi.org/10.3233/sw-2012-0065>
- Bechhofer, S., van Harmelen, F., Hendler, J., Horrocks, I., McGuinness, D. L., Patel-Schneider, P. F., & Stein, L. A. (2004). *OWL Web Ontology Language Reference* (tech. rep.). W3C.
- Bedard, Y., & Bernier, E. (2002). Supporting multiple representations with spatial databases views management and the concept of VUEL. *ISPRS/ICA Joint Workshop on Multi-Scale Representations of Spatial Data, Ottawa*, 1–14.

- Bellini, P., & Nesi, P. (2018). Performance assessment of RDF graph databases for smart city services. *Journal of Visual Languages & Computing*, 45, 24–38. <https://doi.org/https://doi.org/10.1016/j.jvlc.2018.03.002>
- Benslimane, D., Vangenot, C., Roussey, C., & Arara, A. (2003). Multirepresentation in Ontologies. In L. Kalinichenko, R. Manthey, B. Thalheim, & U. Wloka (Eds.), *Advances in Databases and Information Systems: 7th East European Conference, ADBiS 2003, Dresden, Germany, September 3-6, 2003. Proceedings*. Springer Berlin Heidelberg. https://doi.org/10.1007/978-3-540-39403-7_3
- Berners-Lee, T. (2006). Linked Data, <https://www.w3.org/designissues/linkedata.html>.
- Berners-Lee, T., Fielding, R., & Masinter, L. (1998). *RFC 2396: Uniform Resource Identifiers (URI)* (tech. rep.). IETF RFC.
- Berners-Lee, T., Hendler, J., & Lassila, O. (2001). The Semantic Web. *Scientific American*, 284(5), 28–37. <https://doi.org/10.1038/scientificamerican0501-34>
- Bobzien, M., Burghardt, D., Petzold, I., Neun, M., & Weibel, R. (2008). Multi-representation Databases with Explicitly Modeled Horizontal, Vertical, and Update Relations. *Cartography and Geographic Information Science*, 35(1), 3–16. <https://doi.org/10.1559/152304008783475698>
- Çabuk, S., Erdoğan, M., & Önal, E. (2015). Open Street Map Verilerinden Yararlanılarak 1/50 K Ölçekli Harita Üretilebilirliğinin Araştırılması. *Harita Dergisi*, 154, 26–34.
- Carral, D., Scheider, S., Janowicz, K., Vardeman, C., Krisnadhi, A. A., & Hitzler, P. (2013). An Ontology Design Pattern for Cartographic Map Scaling. In P. Cimiano, O. Corcho, V. Presutti, L. Hollink, & S. Rudolph (Eds.), *The semantic web: Semantics and big data: 10th international conference proceedings*. Springer. https://doi.org/10.1007/978-3-642-38288-8_6
- Chaabane, S., & Jaziri, W. (2018). A novel algorithm for fully automated mapping of geospatial ontologies. *Journal of Geographical Systems*, 20(1), 85–105. <https://doi.org/10.1007/s10109-017-0263-0>
- Claramunt, C. (2020). Ontologies for geospatial information: Progress and challenges ahead. *Journal of Spatial Information Science*, 20(20), 35–41. <https://doi.org/10.5311/josis.2020.20.666>
- Codescu, M., Vale, D. C., Kutz, O., & Mossakowski, T. (2012). Ontology-based route planning for OpenStreetMap. *CEUR Workshop Proceedings*, 901, 62–73.
- Cyganiak, R., Wood, D., & Lanthaler, M. (2014). *RDF 1.1 Concepts and Abstract Syntax* (tech. rep.). W3C.
- Davidovic, N., Mooney, P., Stoimenov, L., & Minghini, M. (2016). Tagging in volunteered geographic information: An analysis of tagging practices for cities and urban regions in OpenStreetMap. *ISPRS International Journal of Geo-Information*, 5(12), 1–22. <https://doi.org/10.3390/ijgi5120232>
- DBP. (2021). DBpedia, <https://www.dbpedia.org>.
- DC. (2021). Dublin Core Vocabulary, <http://purl.org/dc/terms/>.
- Debruyne, C., McGlinn, K., McNerney, L., & O’Sullivan, D. (2017). A lightweight approach to explore, enrich and use data with a geospatial dimension with semantic web technologies. *GeoRich 2017 - 4th International ACM Workshop*

- on Managing and Mining Enriched Geo-Spatial Data, in Conjunction with SIGMOD 2017, Illinois, Chicago, 1–6. <https://doi.org/10.1145/3080546.3080548>
- Debruyne, C., Meehan, A., Clinton, É., McNerney, L., Nautiyal, A., Lavin, P., & O’Sullivan, D. (2017). Ireland’s Authoritative Geospatial Linked Data. In C. D’Amato, M. Fernandez, V. Tamma, F. Lecue, P. Cudré-Mauroux, J. Sequeda, C. Lange, & J. Heflin (Eds.), *Lecture notes in computer science* (pp. 66–74). Springer. https://doi.org/10.1007/978-3-319-68204-4_6
- Decker, S., Melnik, S., van Harmelen, F., Fensel, D., Klein, M., Broekstra, J., Erdmann, M., & Horrocks, I. (2000). The Semantic Web: the roles of XML and RDF. *IEEE Internet Computing*, 4(5), 63–73. <https://doi.org/10.1109/4236.877487>
- Devogele, T., Trevisan, J., & Raynal, L. (1996). Building a multi-scale database with scale-transition relationships. *Proceedings of 7th International Symposium on Spatial Data Handling, Delft*, 337–351.
- Du, H. (2015). *Matching disparate geospatial datasets and validating matches using spatial logic* (Doctoral dissertation). University of Nottingham. UK.
- Du, H., Alechina, N., Jackson, M., & Hart, G. (2017). A Method for Matching Crowd-sourced and Authoritative Geospatial Data. *Transactions in GIS*, 21(2), 406–427. <https://doi.org/10.1111/tgis.12210>
- Du, H., Anand, S., Alechina, N., Morley, J., Hart, G., Leibovici, D., Jackson, M., & Ware, M. (2012). Geospatial Information Integration for Authoritative and Crowd Sourced Road Vector Data. *Transactions in GIS*, 16(4), 455–476. <https://doi.org/10.1111/j.1467-9671.2012.01303.x>
- DUL. (2021). DOLCE+DnS Ultralite ontology, <http://www.ontologydesignpatterns.org/ont/dul/DUL.owl#>.
- Egenhofer, M. J., & Mark, D. M. (1995). Naïve Geography. In A. U. Frank & W. Kuhn (Eds.), *Spatial Information Theory A Theoretical Basis for GIS* (pp. 1–15). Springer.
- Eiðsson, E. K. (2018). *Transformation and linking of authoritative multi-scale geo-data for the Semantic Web: a case study of Swedish national building data sets* (Master’s thesis). Lund University. Sweden.
- Erden, Ö. E. (2019). *OpenStreetMap Verilerindeki Bina Detaylarının Büyük Ölçekli Topografik Harita Verileri ile Karşılaştırmalı Geometrik Kalite Analizi* (Yüksek lisans tezi). Yıldız Teknik Üniversitesi. Türkiye.
- Fan, H., Yang, B., Zipf, A., & Rousell, A. (2016). A polygon-based approach for matching OpenStreetMap road networks with regional transit authority data. *International Journal of Geographical Information Science*, 30(4), 748–764. <https://doi.org/10.1080/13658816.2015.1100732>
- Fan, H., Zipf, A., & Fu, Q. (2014). Estimation of building types on openstreetmap based on urban morphology analysis. *Lecture notes in geoinformation and cartography* (pp. 19–35). Kluwer Academic Publishers. https://doi.org/10.1007/978-3-319-03611-3_2
- Fan, H., Zipf, A., Fu, Q., & Neis, P. (2014). Quality assessment for building footprints data on OpenStreetMap. *International Journal of Geographical Information Science*, 28(4), 700–719. <https://doi.org/10.1080/13658816.2013.867495>
- Feigenbaum, L. (2008). *SPARQL By Example: The Cheat Sheet* (tech. rep.). W3C SPARQL Working Group.
- FOAF. (2021). Friend of A Friend Vocabulary, <http://xmlns.com/foaf/spec/>.

- Fonte, C. C., Minghini, M., Antoniou, V., Patriarca, J., & See, L. (2018). Classification of building function using available sources of VGI. *International Archives of the Photogrammetry, Remote Sensing and Spatial Information Sciences - ISPRS Archives*, 42(4), 283–290. <https://doi.org/10.5194/isprs-archives-XLII-4-209-2018>
- Friis-Christensen, A., Jensen, C. S., Nytnun, J. P., & Skogan, D. (2005). A conceptual schema language for the management of multiple representations of geographic entities. *Transactions in GIS*, 9(3), 345–380. <https://doi.org/10.1111/j.1467-9671.2005.00222.x>
- Fu, Z., & Wu, J. (2008). Entity matching in vector spatial data. *The International Archives of the Photogrammetry, Remote Sensing and Spatial Information Sciences. Vol. XXXVII. Part B4. Beijing*, 1467–1472.
- Gandon, F. (2014). URL-URI-IRI, <http://www-sop.inria.fr/members/fabien.gandon/docs/url-uri-iri.png>.
- Gangemi, A., & Presutti, V. (2009). Ontology Design Patterns. In S. Staab & R. Studer (Eds.), *Handbook on ontologies* (2nd ed., pp. 221–243). Springer Berlin Heidelberg. <https://doi.org/10.1007/978-3-540-92673-3>
- Garbis, G., Kyzirakos, K., & Koubarakis, M. (2013). Geographica: A Benchmark for Geospatial RDF Stores (Long Version). In H. Alani, L. Kagal, A. Fokoue, P. Groth, C. Biemann, J. X. Parreira, L. Aroyo, N. Noy, C. Welty, & K. Janowicz (Eds.), *The semantic web – iswc 2013* (pp. 343–359). Springer Berlin Heidelberg.
- Glaser, H., Millard, I. C., Sung, W.-k., Lee, S., Kim, P., & You, B.-j. (2009). Research on Linked Data and Co-reference Resolution. *Proceeding of International Conference on Dublin Core and Metadata Applications, Seoul*, 113–117.
- GN. (2021). GeoNames, <https://www.geonames.org>.
- Goodwin, J., Dolbear, C., & Hart, G. (2008). Geographical linked data: The administrative geography of Great Britain on the semantic web. *Transactions in GIS*, 12(Suppl. 1), 19–30. <https://doi.org/10.1111/j.1467-9671.2008.01133.x>
- Gruber, T. R. (1993). A translation approach to portable ontology specifications. *Knowledge Acquisition*, 5(2), 199–220. <https://doi.org/10.1006/knac.1993.1008>
- Guarino, N. (1997). Semantic matching: Formal ontological distinctions for information organization, extraction, and integration. In M. T. Pazienza (Ed.), *Information extraction a multidisciplinary approach to an emerging information technology* (pp. 139–170). Springer. https://doi.org/10.1007/3-540-63438-x_8
- Hacar, M. (2019). *Yol Ağlarının Geometrik Entegrasyonu için Nesne Eşleme Yöntemlerinin Geliştirilmesi* (Doktora tezi). Yıldız Teknik Üniversitesi. Türkiye.
- Hacar, M., & Gökgöz, T. (2019). A New, Score-Based Multi-Stage Matching Approach for Road Network Conflation in Different Road Patterns. *ISPRS International Journal of Geo-Information*, 8(2). <https://doi.org/10.3390/ijgi8020081>
- Hahmann, S., & Burghardt, D. (2010). Linked Data - A Multiple Representation Database at Web Scale? *13th Workshop of the ICA commission on Generalisation and Multiple Representation, Zurich*.

- Hakimpour, F. (2003). *Using Ontologies to Resolve Semantic Heterogeneity for Integrating Spatial Database Schemata* (Doctoral dissertation). University of Zurich. Switzerland.
- Homburg, T., Steppan, S., & Würriehausen, F. (2019). Semantic Data Integration and Quality Assurance of Thematic Maps in the German Federal Agency for Cartography and Geodesy. In W. Abramowicz & R. Corchuelo (Eds.), *Business information systems workshops* (pp. 543–555). Springer International Publishing.
- Horridge, M., Knublauch, H., Rector, A., Stevens, R., Wroe, C., Jupp, S., Moulton, G., Drummond, N., & Brandt, S. (2011). *A Practical Guide To Building OWL Ontologies Using Protege 4 and CO-ODE Tools Edition 1.3*. The University Of Manchester Copyright.
- Horrocks, I., Patel-Schneider, P. F., Boley, H., Tabet, S., Grosof, B., Dean, M., vd. (2004). *SWRL: A semantic web rule language combining OWL and RuleML* (tech. rep.). W3C.
- Hu, Y., & Janowicz, K. (2016). Enriching Top-down Geo-ontologies Using Bottom-up Knowledge Mined from Linked Data. In H. Onsrud & W. Kuhn (Eds.), *Advancing geographic information science: The past and next twenty years* (pp. 183–198). GSDI Association Press.
- Hu, Y., Janowicz, K., Carral, D., Scheider, S., Kuhn, W., Berg-Cross, G., Hitzler, P., Dean, M., & Kolas, D. (2013). A Geo - ontology Design Pattern for Semantic Trajectories. *Spatial information theory, cosit 2013, lecture notes in computer science, vol 8116*. (pp. 438–456). Springer, Cham. https://doi.org/10.1007/978-3-319-01790-7_24
- Huang, W. (2020). *Geospatial data and knowledge on the Web: Knowledge-based geospatial data integration and visualisation with Semantic Web technologies* (Doctoral dissertation). Lund University. Sweden.
- Huang, W., Mansourian, A., Abdolmajidi, E., Xu, H., & Harrie, L. (2018). Synchronising geometric representations for map mashups using relative positioning and Linked Data. *International Journal of Geographical Information Science*, 32(6), 1117–1137. <https://doi.org/10.1080/13658816.2018.1441416>
- Huang, W., Raza, S. A., Mirzov, O., & Harrie, L. (2019). Assessment and Benchmarking of Spatially Enabled RDF Stores for the Next Generation of Spatial Data Infrastructure. *ISPRS International Journal of Geo-Information*, 8(7), 310. <https://doi.org/10.3390/ijgi8070310>
- INSPIRE. (2014). *Drafting Team "Data Specifications", Generic Conceptual Model, Version 3.4* (tech. rep.). INSPIRE.
- Jena. (2021). Apache Jena, <https://jena.apache.org>.
- Kashyap, V., & Sheth, A. (1997). Semantic heterogeneity in global information systems: The role of metadata, context and ontologies. *Cooperative Information Systems: Current Trends and Directions*, 139–178.
- Kavouras, M., & Kokla, M. (2002). Developing Multi-scale, Multi-context Databases Through the Semantic Integration of Heterogeneous Datasets. *The Proceedings of the 8th EC-GI&GIS Workshop, Dublin*.
- Kettouch, M. S. (2017). *A New Approach for Interlinking and Integrating Semi-Structured and Linked Data* (Doctoral dissertation). Anglia Ruskin University. UK.

- Kilpeläinen, T. (2000). Maintenance of multiple representation databases for topographic data. *The Cartographic Journal*, 37(2), 101–107. <https://doi.org/10.1179/caj.2000.37.2.101>
- Kim, J. O., Yu, K., Heo, J., & Lee, W. H. (2010). A new method for matching objects in two different geospatial datasets based on the geographic context. *Computers and Geosciences*, 36(9), 1115–1122. <https://doi.org/10.1016/j.cageo.2010.04.003>
- Kuhn, W. (2001). Ontologies in support of activities in geographic space. *International Journal of Geographic Information Science*, 15(7), 613–631.
- Kunze, C., & Hecht, R. (2015). Semantic enrichment of building data with volunteered geographic information to improve mappings of dwelling units and population. *Computers, Environment and Urban Systems*, 53, 4–18. <https://doi.org/10.1016/j.compenvurbsys.2015.04.002>
- Laurini, R. (2017). Geographic Ontologies. In R. Laurini (Ed.), *Geographic knowledge infrastructure* (pp. 111–137). ISTE Press - Elsevier. <https://doi.org/10.1016/b978-1-78548-243-4.50006-2>
- Li, W., Raskin, R., & Goodchild, M. F. (2012). Semantic similarity measurement based on knowledge mining: an artificial neural net approach. *International Journal of Geographical Information Science*, 26(8), 1415–1435. <https://doi.org/10.1080/13658816.2011.635595>
- Liu, L., Zhu, X., Zhu, D., & Ding, X. (2018). M:N Object matching on multiscale datasets based on MBR combinatorial optimization algorithm and spatial district. *Transactions in GIS*, 22(6), 1573–1595. <https://doi.org/10.1111/tgis.12488>
- LOD. (2021). Linked Open Data Cloud, <https://lod-cloud.net/>.
- Memduhoglu, A., & Basaraner, M. (2017). Potential of Spatial Semantics for Developing Multi-Representation Spatial Databases. *Proceeding of UCTEA International Geographic Information Systems Congress, Adana*, 71–75.
- Memduhoglu, A., & Basaraner, M. (2018). Possible Contributions of Spatial Semantic Methods and Technologies to Multi-Representation Spatial Database Paradigm. *International Journal of Engineering and Geosciences*, 3(3), 108–118. <https://doi.org/10.26833/ijeg.413473>
- Mitchell, P. (2010). Measuring similarity between polygons, <https://student.cs.uwaterloo.ca/~cs763/projects/phil.pdf>.
- Musen, M. A. (2015). The protégé project: a look back and a look forward. *AI Matters*, 1(4), 4–12. <https://doi.org/10.1145/2757001.2757003>
- Mustière, S., & van Smaalen, J. (2007). Database Requirements for Generalisation and Multiple Representations. In W. Mackaness, A. Ruas, & L. Sarjakoski (Eds.), *Generalisation of geographic information: Cartographic modelling and applications* (pp. 113–136). Elsevier Ltd. <https://doi.org/10.1016/b978-008045374-3/50008-9>
- Nikolaou, C., Dogani, K., Bereta, K., Garbis, G., Karpathiotakis, M., Kyzirakos, K., & Koubarakis, M. (2015). Sextant: Visualizing time-evolving linked geospatial data. *Journal of Web Semantics*, 35(Part 1), 35–52. <https://doi.org/10.1016/j.websem.2015.09.004>
- Nowak Da Costa, J. (2016). Towards Building Data Semantic Similarity Analysis: OpenStreetMap and the Polish Database of Topographic Objects. *Proceedings*

- 2016 Baltic Geodetic Congress (Geomatics), Gdansk, 269–275. <https://doi.org/10.1109/BGC.Geomatics.2016.55>
- Olteanu-Raimond, A. M., Hart, G., Foody, G. M., Touya, G., Kellenberger, T., & Demetriou, D. (2017). The Scale of VGI in Map Production: A Perspective on European National Mapping Agencies. *Transactions in GIS*, 21(1), 74–90. <https://doi.org/10.1111/tgis.12189>
- Parliament. (2021). Parliament, <https://github.com/semwebcentral/parliament>.
- Perry, M., & Herring, J. (2012). *GeoSPARQL - A Geographic Query Language for RDF Data* (tech. rep.). OGC.
- Regalia, B., Janowicz, K., Mai, G., Varanka, D., & Usery, E. L. (2018). GNIS-LD: Serving and Visualizing the Geographic Names Information System Gazetteer as Linked Data. In A. Gangemi, R. Navigli, M.-E. Vidal, P. Hitzler, R. Troncy, L. Hollink, A. Tordai, & M. Alam (Eds.), *The semantic web* (pp. 528–540). Springer International Publishing. <https://doi.org/10.1007/978-3-642-29923-0>
- Renteria-Agualimpia, W., & Levashkin, S. (2011). Multi-criteria geographic information retrieval model based on geospatial semantic integration. In C. Claramunt, S. Levashkin, & M. Bertolotto (Eds.), *Lecture notes in computer science (geos 2011)* (pp. 166–181). Springer-Verlag Berlin Heidelberg. https://doi.org/10.1007/978-3-642-20630-6_11
- Revell, P., & Antoine, B. (2009). Automated matching of building features of differing levels of detail: A case study. *Proceedings of the 24th International Cartographic Conference, Santiago de Chile*, 15–21.
- Ronzhin, S., Folmer, E., Lemmens, R., Mellum, R., von Brasch, T. E., Martin, E., Romero, E. L., Kytö, S., Hietanen, E., & Latvala, P. (2019). Next generation of spatial data infrastructure: lessons from linked data implementations across Europe. *International journal of spatial data infrastructures research*, 14, 83–107.
- Ruiz-Lendínez, J. J., Ariza-López, F. J., Ureña-Cámara, M. A., & Blázquez, E. B. (2011). Digital map conflation: A review of the process and a proposal for classification. *International Journal of Geographical Information Science*, 25(9), 1439–1466. <https://doi.org/10.1080/13658816.2010.519707>
- Ruiz-Lendínez, J. J., Ureña-Cámara, M. A., & Ariza-López, F. J. (2017). A polygon and point-based approach to matching geospatial features. *ISPRS International Journal of Geo-Information*, 6(12). <https://doi.org/10.3390/ijgi6120399>
- Rutzinger, M., Rottensteiner, F., & Pfeifer, N. (2009). A Comparison of Evaluation Techniques for Building Extraction From Airborne Laser Scanning. *IEEE Journal of Selected Topics in Applied Earth Observations and Remote Sensing*, 2(1), 11–20. <https://doi.org/10.1109/jstars.2009.2012488>
- Savoy, J., & Gaussier, E. (2010). Information retrieval. In N. Indurkha & F. Damerau (Eds.), *Handbook of natural language processing* (2nd, pp. 455–484). Chapman; Hall/CRC. <https://doi.org/10.4324/9781351044677-24>
- Schwering, A. (2006). *Semantic Similarity Measurement including Spatial Relations for Semantic Information Retrieval of Geo-Spatial Data* (Doctoral dissertation). Universität Münster. Germany. <https://doi.org/10.1002/ejoc.201200111>
- Sinha, G., Mark, D. M., Kolas, D., Varanka, D. E., Romero, B. E., Feng, C.-C., Usery, E. L., Liebermann, J., & Sorokine, A. (2014). An Ontology Design Pattern for Surface Water Features. In M. Duckham, E. Pebesma, K. Stewart, & A. U. Frank (Eds.), *Geographic Information Science: 8th International Conference, GIScience*

- 2014, *Proceedings* (pp. 187–203). Springer International Publishing. https://doi.org/10.1007/978-3-319-11593-1_13
- SKOS. (2021). Simple Knowledge Organization System Vocabulary, <http://www.w3.org/2004/02/skos/core#>.
- Smith, B., & Mark, D. (1998). Ontology and geographic kinds. *Proceedings of the 8th International Symposium on Spatial Data Handling, Vancouver*, 308–320.
- Stadler, A., Nagel, C., König, G., & Kolbe, T. H. (2009). Making Interoperability Persistent: A 3D Geo Database Based on CityGML. *3D Geo-Information Sciences*, 175–192. https://doi.org/10.1007/978-3-540-87395-2_11
- Stadler, C., Lehmann, J., Höffner, K., & Auer, S. (2012). LinkedGeoData: A Core for a Web of Spatial Open Data. *Semantic Web*, 3(4), 333–354. <https://doi.org/10.3233/sw-2011-0052>
- Stoter, J., Lemmens, R., Kobben, B., & Bakker, N. J. (2006). Semantic data integration in a multiple representation environment. *ISPRS - Workshop on Multiple Representation and Interoperability of Spatial Data, Hannover*, 22–29.
- Stoter, J., Visser, T., van Oosterom, P., Quak, W., & Bakker, N. (2011). A semantic-rich multi-scale information model for topography. *International Journal of Geographical Information Science*, 25(5), 739–763. <https://doi.org/10.1080/13658816.2010.490218>
- Strabon. (2021). Strabon - A Spatiotemporal RDF Store, <http://www.strabon.di.uoa.gr/publications.html>.
- Sun, K., Zhu, Y., & Song, J. (2019). Progress and Challenges on Entity Alignment of Geographic Knowledge Bases. *ISPRS International Journal of Geo-Information*, 8(2), 77. <https://doi.org/10.3390/ijgi8020077>
- TagInfo. (2021). TagInfo OpenStreetMap, <https://taginfo.openstreetmap.org>.
- Thomson, M.-K. (2009). *Dwelling on ontology - semantic reasoning over topographic maps* (Doctoral dissertation). University College London. UK.
- Uitermark, H. T., van Oosterom, P., Mars, N. J. I., & Molenaar, M. (1999). Ontology-Based Geographic Data Set Integration. In M. Böhlen, C. Jensen, & M. Scholl (Eds.), *Spatio-temporal database management* (Lncs:1678, pp. 60–78). Springer. https://doi.org/10.1007/3-540-48344-6_4
- Uitermark, H. T., van Oosterom, P., Mars, N. J. I., & Molenaar, M. (2005). Ontology-based integration of topographic data sets. *International Journal of Applied Earth Observation and Geoinformation*, 7(2), 97–106. <https://doi.org/10.1016/j.jag.2005.03.002>
- Ulutaş, D., Kara, G., & Cömert, Ç. (2016). Semantic definition and matching for implementing national spatial data infrastructures. *Journal of Spatial Science*, 61(2), 441–459. <https://doi.org/10.1080/14498596.2016.1142397>
- Usery, E. L., & Varanka, D. E. (2012). Design and development of linked data from The National Map. *Semantic Web Journal*, 1(3), 371–384. <https://doi.org/10.3233/sw-2011-0054>
- van den Brink, L., Janssen, P., Quak, W., & Stoter, J. (2017). Towards a high level of semantic harmonisation in the geospatial domain. *Computers, Environment and Urban Systems*, 62, 233–242. <https://doi.org/10.1016/j.compenvurbsys.2016.12.002>
- Vangenot, C., Parent, C., & Spaccapietra, S. (2002). Modelling and Manipulating Multiple Representations of Spatial Data. In D. E. Richardson & P. van Oosterom

- (Eds.), *Advances in spatial data handling: 10th international symposium on spatial data handling*. Springer Berlin Heidelberg. https://doi.org/10.1007/978-3-642-56094-1_7
- van Wijngaarden, F., van Putten, J., van Oosterom, P., & Uitermark, H. (1997). Map integration - update propagation in a multi-source environment. *Proceedings of the ACM Workshop on Advances in Geographic Information Systems*, 71–76. <https://doi.org/10.1145/267825.267844>
- Varanka, D. E., & Cheatham, M. (2016). Spatial Concepts for Hydrography Ontology Alignment. *The 19th International Research Symposium on Computer-based Cartography-AutoCarto2016, Albuquerque*, 224–237.
- Varanka, D. E., & Usery, E. L. (2015). An Applied Ontology for Semantics Associated with Surface Water Features. In O. Ahlqvist, D. Varanka, S. Fritz, & K. Janowicz (Eds.), *Land use and land cover semantics: Principles, best practices, and prospects* (pp. 145–170). CRC Press.
- Varanka, D. E., Usery, E. L., & Mattli, D. M. (2015). Adapting the U.S. National Hydrography Dataset to Linked Open Data. In R. S. Claudia, C. B. M. Cruz, & P. M. L. d. Menezes (Eds.), *Cartography - maps connecting the world, icc2015* (pp. 129–146). Springer International Publishing. https://doi.org/10.1007/978-3-319-17738-0_10
- Vargas-Muñoz, J. E., Lobry, S., Falcão, A. X., & Tuia, D. (2019). Correcting rural building annotations in OpenStreetMap using convolutional neural networks. *ISPRS Journal of Photogrammetry and Remote Sensing*, 147(May 2018), 283–293. <https://doi.org/10.1016/j.isprsjprs.2018.11.010>
- Vilches-Blázquez, L. M., & Saavedra, J. (2019). A framework for connecting two interoperability universes: OGC Web Feature Services and Linked Data. *Transactions in GIS*, 23(1), 22–47. <https://doi.org/10.1111/tgis.12496>
- Vilches-Blázquez, L. M., Saquicela, V., & Corcho, O. (2012). Interlinking geospatial information in the web of data. In J. Gensel, D. Josselin, & D. Vandenbroucke (Eds.), *Bridging the Geographic Information Science, AGILE'2012* (pp. 119–139). Springer Berlin Heidelberg. https://doi.org/10.1007/978-3-642-29063-3_7
- Vilches-Blázquez, L. M., Villazón-Terrazas, B., Corcho, O., & Gómez-Pérez, A. (2014). Integrating geographical information in the Linked Digital Earth. *International Journal of Digital Earth*, 7(7), 554–575. <https://doi.org/10.1080/17538947.2013.783127>
- Volz, S. (2005). Data-driven matching of geospatial schemas. *International Conference on Spatial Information Theory, L'Aquila*, 115–132.
- Wiemann, S., & Bernard, L. (2016). Spatial data fusion in Spatial Data Infrastructures using Linked Data. *International Journal of Geographical Information Science*, 30(4), 613–636. <https://doi.org/10.1080/13658816.2015.1084420>
- Xavier, E. M., Ariza-López, F. J., & Ureña-Cámara, M. A. (2016). A survey of measures and methods for matching geospatial vector datasets. *ACM Computing Surveys*, 49(2), 1–34. <https://doi.org/10.1145/2963147>
- Xu, Y., Xie, Z., Chen, Z., & Wu, L. (2017). Shape similarity measurement model for holed polygons based on position graphs and Fourier descriptors. *International Journal of Geographical Information Science*, 31(2), 253–279. <https://doi.org/10.1080/13658816.2016.1192637>
- Yaghouti, N., Kahani, M., & Behkamal, B. (2015). A metric-driven approach for interlinking assessment of RDF graphs. *International Symposium on Computer*

- Science and Software Engineering (CSSE)*, Tabriz, 1–8. <https://doi.org/10.1109/csicsse.2015.7369244>
- YAGO. (2021). YAGO: A High-Quality Knowledge Base, <https://yago-knowledge.org>.
- Yan, H., & Li, J. (2015). *Spatial Similarity Relations in Multi-scale Map Spaces* (H. Yan & J. Li, Eds.). Springer International Publishing. <https://doi.org/10.1007/978-3-319-09743-5>
- Yi, S. (2013). Learning ontologies for geographic entity matching and multi-sources data fusion. *International Conference on Geoinformatics, Kaifeng, Henan*, 1–5.
- Yu, L., Qiu, P., Liu, X., Lu, F., & Wan, B. (2018). A holistic approach to aligning geospatial data with multidimensional similarity measuring. *International Journal of Digital Earth*, 11(8), 845–862. <https://doi.org/10.1080/17538947.2017.1359688>
- Zhang, X., Ai, T., Stoter, J., & Zhao, X. (2014). Data matching of building polygons at multiple map scales improved by contextual information and relaxation. *ISPRS Journal of Photogrammetry and Remote Sensing*, 92, 147–163. <https://doi.org/10.1016/j.isprsjprs.2014.03.010>
- Zhang, X., Yin, W., Yang, M., Ai, T., & Stoter, J. (2018). Updating authoritative spatial data from timely sources: A multiple representation approach. *International Journal of Applied Earth Observation and Geoinformation*, 72, 42–56. <https://doi.org/10.1016/j.jag.2018.05.022>
- Zhang, Y., Huang, J., Deng, M., Chen, C., Zhou, F., Xie, S., & Fang, X. (2019). Automated matching of multi-scale building data based on relaxation labelling and pattern combinations. *ISPRS International Journal of Geo-Information*, 8(1), 38. <https://doi.org/10.3390/ijgi8010038>
- Zhou, X., Chen, Z., Zhang, X., & Ai, T. (2018). Change detection for building footprints with different levels of detail using combined shape and pattern analysis. *ISPRS International Journal of Geo-Information*, 7(10). <https://doi.org/10.3390/ijgi7100406>

A

KAYNAK KODLAR

Yapılan uygulamanın birçok aşamasında Python ortamında yazılan kodlardan faydalanılmıştır. Geliştirme ortamı olarak Jupyter Lab'ın tercih edildiği ve aşağıda ilk aşama olan veri düzenlemesi kısmı verilen bu Python kodlarının tamamı GitHub sitesinde (<https://github.com/carto-web/building-integration>) GPL-3.0 lisansı ile ilgililerin kullanımına sunulmuştur.

```
#!/usr/bin/env python
# coding: utf-8

import geopandas as gpd
import pandas as pd

# Verileri geojson formatından Pandas DataFrame olarak içe aktar

topo1k = gpd.GeoDataFrame.from_file('data/topo1k.geojson')

empty_cols_topo1k =
[col for col in topo1k.columns if topo1k[col].isnull().all()]
topo1k.drop(empty_cols_topo1k, axis = 1, inplace = True)

# Detaylara ID ekle

Feature_topo1k = []

for i in range(len(topo1k)):
    Feature_topo1k.append('topo1k' + str(i+1))

topo1k['fid'] = Feature_topo1k

topo1k.drop(['OBJECTID', 'GLOBALID', 'SHAPE_Leng',
'SHAPE_Area'],axis=1,inplace=True)

topo1k.rename(columns={"BINA_ADI":"building_name",
"IL_KODU":"province_code","ILCE_KODU":"district_code",
"MAHALLE_KO":"neighborhood_code",
"BINA_NO":"building_num"}, inplace=True)
```

```

# building_num column has NaN values that
# cause datatype being float instead of an integer.
# So it needs to be fixed.

topo1k['building_num'].fillna(0,inplace=True)
topo1k['building_num'] = topo1k['building_num'].astype('int64')


# Veri setinde yer alan Türkçe karakterleri
# latin karakterlerle değiştir.

tr2en = {'ğ':'g', 'ç':'c', 'ş':'s', 'ü':'u',
'ö':'o', 'ı':'i', 'Ğ':'G', 'Ç':'C', 'Ş':'S',
'Ü':'U', 'Ö':'O', 'İ':'I'}

for i in topo1k.columns:
    try:
        topo1k.replace({i:tr2en}, regex=True, inplace=True)
    except Exception:
        pass

osm = gpd.GeoDataFrame.from_file('data/osm.geojson')

# building sütunu null olan satırları temizleme
# (sadece binaları al)

osm = osm[osm["building"].notna()]

# sadece bina olan detayları al

osm = osm[osm['building'] != 'transportation']

# Boş kolonları temizleme

empty_cols_osm =
[col for col in osm.columns if osm[col].isnull().all()]
osm.drop(empty_cols_osm, axis = 1, inplace = True)

# Veri setinde yer alan Türkçe karakterleri
# latin karakterlerle değiştir.

tr2en = {'ğ':'g', 'ç':'c', 'ş':'s', 'ü':'u',
'ö':'o', 'ı':'i', 'Ğ':'G', 'Ç':'C', 'Ş':'S',
'Ü':'U', 'Ö':'O', 'İ':'I'}

for i in osm.columns:
    try:

```

```

        osm.replace({i:tr2en}, regex=True, inplace=True)
    except Exception:
        pass

# Detaylara URI, ID ekleme

Feature_osm = []

for i in range(len(osm)):
    Feature_osm.append('osm' + str(i+1))

osm['fid'] = Feature_osm

# osmId sütununu var olan id sutununun
# ikinci kısmını kullanarak oluştur.

osm['osmid'] = osm['id'].str.split('/', expand=True)[1]
osm.drop(['@id', 'id'], axis=1, inplace=True)

# Sütun başlıklarında yer alan bazı karakterleri düzenle.

osm.columns = osm.columns.str.replace(':', '_')
osm.columns = osm.columns.str.replace('@', '')

osm['timestamp'] =
pd.to_datetime(osm['timestamp']).dt.strftime('%Y-%m-%d')

topo25k = gpd.GeoDataFrame.from_file('data/topo25k.geojson')

# Boş kolonları temizleme

empty_cols_topo25k =
[col for col in topo25k.columns if topo25k[col].isnull().all()]

topo25k.drop(empty_cols_topo25k, axis = 1, inplace = True)

# Detaylara URI, ID ekleme

Feature_topo25k = []

for i in range(len(topo25k)):
    Feature_topo25k.append('topo25k' + str(i+1))

topo25k['fid'] = Feature_topo25k

# topo25k sütun başlıklarını ingilizceye çevirerek düzenle.

```

```

topo25k.rename(columns={'DETAY_KODU':'feature_code',
'DETAY_ADI':'feature_name', 'SEMBOL':'symbol',
'OZEL_ISMI':'special_name', 'Shape_Leng':'shape_leng',
'Shape_Area':'shape_area'}, inplace=True)

# Veri setinde yer alan Türkçe karakterleri
# latin karakterlerle değiştir.

tr2en = {'ğ':'g', 'ç':'c', 'ş':'s', 'ü':'u',
'ö':'o', 'ı':'i', 'Ğ':'G', 'Ç':'C', 'Ş':'S',
'Ü':'U', 'Ö':'O', 'İ':'I'}

for i in topo25k.columns:
    try:
        topo25k.replace({i:tr2en}, regex=True, inplace=True)
    except Exception:
        pass

topo1k.to_file('data/clean_topo1k.geojson', driver='GeoJSON')
osm.to_file('data/clean_osm.geojson', driver='GeoJSON')
topo25k.to_file('data/clean_topo25k.geojson', driver='GeoJSON')

```

KESİŞİM ALANI ORANI PERFORMANS DEĞERLERİ

Şekil 4.5'te gösterilen, O_{KA} için kesinlik ve duyarlılık değerlerinin hesaplandığı farklı aralıkların tüm ilgili değerleri Tablo B.1'de görülebilir.

Tablo B.1 O_{KA} aralıklarının Doğru Pozitif (TP), Yanlış Pozitif (FP) ve Yanlış Negatif (FN) değerleri

O_{KA}	TP	FP	FN
%25	646	99	29
%30	633	71	42
%35	610	50	65
%40	574	30	101
%45	536	18	139
%50	491	14	184

C

SINIF TABLOLARI

OSM'den elde edilen etiket değerleriyle ontoloji sınıfı eşleştirmelerinin ve tüm verideki kapsama oranının yer aldığı tablolar her etiket için ayrı toplam 8 adet olacak şekilde aşağıda verilmiştir.

Tablo C.1 building anahtarı için oluşturulan kapsam ve sınıf eşleniği tablosu

Sayı	Anahtar = building			
	Değer	Kapsam (%)	Sınıf Eşleniği	
1	yes >>building	81,09	GenericBuilding	*”building’ olarak değiştirilmiştir
2	house	9,19	House	
3	residential	2,79	House	
4	garage	1,03	Accessory_Cantilever	
5	apartments	0,90	House	
6	detached	0,83	House	
7	shed	0,46	Accessory_Cantilever	
8	hut	0,43	Accessory_Cantilever	
9	industrial	0,39	IndustrialBuilding	
10	roof	0,30	Accessory_Cantilever	
11	farm_auxiliary	0,29	Farm	
12	terrace	0,23	Accessory_Cantilever	
13	commercial	0,20	Shop	
14	school	0,18	School	
15	garages	0,15	Accessory_Cantilever	
16	retail	0,13	Shop	
17	construction	0,13	UnderConstructionBuilding	
18	barn	0,12	Accessory_Cantilever	
19	semidetached_house	0,11	House	
20	greenhouse	0,11	Greenhouse	
21	cabin	0,09	Accessory_Cantilever	
22	church	0,07	Church	
23	static_caravan	0,06	Accessory_Cantilever	
24	warehouse	0,06	Warehouse	
25	service	0,05	ServiceBuilding	
26	farm	0,05	Farm	

Table C.1 önceki sayfadan devam etmektedir

Sayı	Anahtar = building		
	Değer	Kapsam (%)	Sınıf Eşleniği
27	bungalow	0,04	Accessory_Cantilever
28	civic	0,03	GenericBuilding
29	office	0,03	Office
30	ruins	0,03	RuinedBuilding
31	public	0,03	OfficialBuilding
32	university	0,03	University
33	hospital	0,02	Hospital
34	manufacture	0,02	Factory
35	hotel	0,02	Hotel_Accommodation
36	carport	0,02	Accessory_Cantilever
37	collapsed	0,02	RuinedBuilding
38	chapel	0,02	Church
39	kindergarten	0,01	Kindergarten
40	hangar	0,01	Hangar
41	semi	0,01	House
42	ger	0,01	Accessory_Cantilever
43	boathouse	0,01	Boathouse
44	storage_tank	0,01	WaterTank
45	mosque	0,01	Mosque
46	dormitory	0,01	Dormitory
47	allotment_house	0,01	House
48	train_station	0,01	RailwayStation
49	college	0,01	HighSchool
50	stable	0,01	Stable
Toplam Kapsam =		%99,87	

Tablo C.2 shop anahtarı için oluşturulan kapsam ve sınıf eşleniği tablosu

Sayı	Anahtar = shop		
	Değer	Kapsam (%)	Sınıf Eşleniği
1	convenience	12,46	Shop
2	supermarket	8,58	SuperMarket
3	clothes	6,37	Clothes
4	hairdresser	5,18	Hairdresser
5	car_repair	4,13	CarRepair_Parts
6	bakery	4,08	Bakery
7	Shop	3,64	Shop
8	car	2,56	CarSell
9	beauty	1,99	Beauty
10	kiosk	1,75	Kiosk
11	butcher	1,61	Butcher
12	hardware	1,60	Telecommunication_IT
13	furniture	1,51	Furniture
14	mobile_phone	1,46	Telecommunication_IT
15	electronics	1,36	Electronics
16	car_parts	1,35	CarRepair_Parts
17	florist	1,35	Florist
18	alcohol	1,32	Liquor
19	shoes	1,29	Shoes
20	mall	1,26	ShoppingCenter
21	doityourself	1,15	Shop
22	variety_store	1,12	Shop
23	optician	1,10	Optician
24	jewelry	1,08	Jewelry
25	gift	1,03	Gift
26	greengrocer	0,97	Greengrocer
27	department_store	0,96	Clothes
28	books	0,96	Books
29	bicycle	0,94	Bicycle
30	travel_agency	0,83	Agency
31	chemist	0,76	Chemistry
32	sports	0,74	Sports
33	laundry	0,73	Cleaning
34	confectionery	0,68	Food
35	stationery	0,67	Shop
36	pet	0,66	Shop
37	computer	0,61	Electronics
38	vacant	0,60	Vacant
39	beverages	0,55	Liquor
40	tyres	0,55	CarRepair_Parts
41	newsagent	0,53	Agency
42	dry_cleaning	0,52	Cleaning
43	cosmetics	0,49	Cosmetics

Table C.2 önceki sayfadan devam etmektedir

Sayı	Anahtar = shop		
	Değer	Kapsam (%)	Sınıf Eşleniği
44	motorcycle	0,48	Motorcycle
45	garden_centre	0,46	Shop
46	funeral_directors	0,45	Office
47	copyshop	0,45	Shop
48	tailor	0,43	Clothes
49	toys	0,40	Toys
50	tobacco	0,40	Shop
51	farm	0,39	Farm
52	deli	0,36	Food
53	interior_decoration	0,36	Office
54	seafood	0,35	Food
55	ticket	0,34	Ticket
56	massage	0,34	HealthCenter
57	storage_rental	0,33	StorageRental
58	trade	0,32	Shop
59	houseware	0,31	Shop
60	photo	0,30	Shop
61	pastry	0,29	Patisserie
62	wine	0,28	Liquor
Toplam Kapsam =		%90,12	

Tablo C.3 amenity anahtarı için oluşturulan kapsam ve sınıf eşleniği tablosu

Sayı	Anahtar = amenity		
	Değer	Kapsam (%)	Sınıf Eşleniği
1	parking	20,80	ParkingArea
2	bench	8,15	Park
3	place_of_worship	6,70	Religious
4	restaurant	6,19	Restaurant_Diner
5	school	5,99	School
6	parking_space	3,90	ParkingArea
7	waste_basket	2,55	Waste_Recycle
8	fuel	2,48	GasStation
9	cafe	2,39	Cafe_Bar
10	fast_food	2,31	Restaurant_Diner
11	bicycle_parking	2,03	ParkingArea
12	bank	1,92	Bank_Finance
13	shelter	1,88	Accessory_Cantilever
14	toilets	1,76	Toilets
15	pharmacy	1,71	Chemistry
16	post_box	1,70	Logistics
17	recycling	1,64	Waste_Recycle
18	kindergarten	1,39	Kindergarten
19	drinking_water	1,25	Fountain
20	hospital	1,03	Hospital
21	bar	1,03	Cafe_Bar
22	atm	0,97	Bank_Finance
23	vending_machine	0,95	Accessory_Cantilever
24	post_office	0,95	Logistics
25	grave_yard	0,93	Cemetery
26	pub	0,91	Cafe_Bar
27	hunting_stand	0,90	TrainingFacility
28	clinic	0,76	HealthCenter
29	doctors	0,72	HealthCenter
30	fountain	0,71	Fountain
31	police	0,71	OfficialBuilding
32	townhall	0,70	OfficialBuilding
33	waste_disposal	0,69	Waste_Recycle
34	community_centre	0,69	CulturalCenter
35	fire_station	0,63	FireStation
36	social_facility	0,59	CulturalCenter
37	library	0,47	Library
38	car_wash	0,45	CarWash
39	marketplace	0,36	Marketplace_Bazaar
40	public_building	0,29	CulturalCenter
41	veterinary	0,23	HealthCenter
42	theatre	0,21	Theatre
43	ice_cream	0,18	Restaurant_Diner

Table C.3 önceki sayfadan devam etmektedir

Sayı	Anahtar = amenity		
	Değer	Kapsam (%)	Sınıf Eşleniği
44	courthouse	0,13	OfficialBuilding
45	nightclub	0,11	Cafe_Bar
46	bureau_de_change	0,07	Bureau_de_change
47	internet_cafe	0,06	Cafe_Bar
48	events_venue	0,05	EventsVenue
49	public_bath	0,05	PublicBath
50	social_centre	0,03	CulturalCenter
51	language_school	0,02	LanguageSchool
52	money_transfer	0,02	Bank_Finance
Toplam Kapsam =		%93,34	

Tablo C.4 leisure anahtarı için oluşturulan kapsam ve sınıf eşleniği tablosu

Sayı	Anahtar = leisure		
	Değer	Kapsam (%)	Sınıf Eşleniği
1	pitch	27,47	Pitch
2	swimming_pool	20,36	SwimmingPool
3	park	15,39	Park
4	garden	10,96	Park
5	playground	9,65	Park
6	sports_centre	3,03	SportsCentre
7	picnic_table	2,50	PicnicArea
8	nature_reserve	1,56	NationalPark
9	track	1,37	Park
10	fitness_centre	0,77	SportsCentre
11	sauna	0,19	PublicBath
12	ice_rink	0,10	IceSkatingArea
13	turkish_bath	0,10	PublicBath
Toplam Kapsam =		%93,45	

Tablo C.5 historic anahtarı için oluşturulan kapsam ve sınıf eşleniği tablosu

Sayı	Anahtar = historic			
	Değer	Kapsam (%)	Sınıf Eşleniği	
1	memorial	19,76		*yapı olmadığından alınmamıştır.
2	archaeological_site	10,32	Historical	
3	wayside_cross	10,22	Historical	
4	ruins	9,20	RuinedBuilding	
5	yes >>historic	8,01	Historical	*'historic' olarak değiştirilmiştir.
6	wayside_shrine	6,60	Historical	
7	monument	5,04	Monument	
8	building >>historic_building	4,06	Historical	*'historic_building' olarak değiştirilmiştir.
9	boundary_stone	3,61	Historical	
10	castle	3,50	Historical	
11	charcoal_pile	3,24	Historical	
12	tomb	2,29	Historical	
13	citywalls	1,11	Historical	
14	heritage	0,88	Historical	
15	railway	0,82	Historical	
16	mine_shaft	0,73	Historical	
17	city_gate	0,35	Historical	
18	aqueduct	0,20	Historical	
19	cistern	0,10	Historical	
Toplam Kapsam =		%90,04		

Tablo C.6 tourism anahtarı için oluşturulan kapsam ve sınıf eşleniği tablosu

Sayı	Anahtar = tourism			
	Değer	Kapsam (%)	Sınıf Eşleniği	
1	information	30,78	Office	
2	hotel	14,30	Hotel_Accommodation	
3	attraction	8,06		*yapı olmadığından alınmamıştır.
4	viewpoint	7,24		*yapı olmadığından alınmamıştır.
5	artwork	6,64	Statue_Artwork	
6	guest_house	5,97	Hotel_Accommodation	
7	picnic_site	5,43	PicnicArea	
8	camp_site	4,68	CampingArea	
9	museum	3,47	Museum	
10	chalet	2,30	House	
11	apartment	2,28	House	
12	hostel	2,03	Hotel_Accommodation	
13	motel	1,69	Hotel_Accommodation	
14	caravan_site	1,15	CampingArea	
Toplam Kapsam =		%96,02		

Tablo C.7 office anahtarı için oluşturulan kapsam ve sınıf eşleniği tablosu

Sayı	Anahtar = office			
	Değer	Kapsam (%)	Sınıf Eşleniği	
1	government	22,31	OfficialBuilding	*’office’ olarak değiştirilmiştir.
2	company	22,19	Company	
3	yes >>office	8,53	Office	
4	estate_agent	6,92	EstateAgent	
5	insurance	5,49	Insurance	
6	lawyer	3,62	Bureau	
7	educational_institution	2,99	School	
8	telecommunication	2,76	Telecommunication_IT	
9	association	2,24	Bureau	
10	ngo	2,10	Bureau	
11	diplomatic	1,58	OfficialBuilding	
12	it	1,52	Telecommunication_IT	
13	administrative	1,22	AdministrativeUnit	
14	employment_agency	1,18	OfficialBuilding	
15	accountant	1,17	Office	
16	research	1,11	Office	
17	religion	1,09	Religious	
18	architect	0,92	Office	
19	financial	0,84	Bank_Finance	
20	tax_advisor	0,79	Bank_Finance	
21	advertising_agency	0,61	Agency	
22	notary	0,60	OfficialBuilding	
23	political_party	0,58	Bureau	
24	logistics	0,48	Logistics	
25	travel_agent	0,47	Agency	
26	foundation	0,31	Foundation	
27	chamber	0,00	Bureau	
Toplam Kapsam =		%93,62		

Tablo C.8 healthcare anahtarı için oluşturulan kapsam ve sınıf eşleniği tablosu

Sayı	Anahtar = healthcare		
	Değer	Kapsam (%)	Sınıf Eşleniği
1	pharmacy	28,92	Chemistry
2	hospital	13,76	Hospital
3	doctor	13,57	HealthCenter
4	clinic	11,71	HealthCenter
5	centre	9,74	HealthCenter
6	dentist	9,56	HealthCenter
7	physiotherapist	2,30	HealthCenter
8	yes >>healthcare	2,00	HealthCenter
9	alternative	1,76	HealthCenter
10	laboratory	1,69	HealthCenter
Toplam Kapsam =		%95,01	

*'healthcare' olarak değiştirilmiştir.

TEZDEN ÜRETİLMİŞ YAYINLAR

Makale

1. Memduhoglu, A., & Basaraner, M. (2021). An Approach for Multi-Scale Urban Building Data Integration and Enrichment through Geometric Matching and Semantic Web. *Cartography and Geographic Information Science*. <https://doi.org/10.1080/15230406.2021.1952108>
2. Memduhoglu, A., & Basaraner, M. (2018). Possible Contributions of Spatial Semantic Methods and Technologies To Multi-Representation Spatial Database Paradigm. *International Journal of Engineering and Geosciences*, 3(3), 108–118. <https://doi.org/10.26833/ijeg.413473>

Konferans Bildirisi

1. Memduhoglu, A., & Basaraner, M. (2019). Use of Semantic Web Technologies for Integrating Multiple Representations of Spatial Features: A Preliminary Study on Buildings. In B. Bayram (Ed.), *International Symposium on Applied Geoinformatics (ISAG-2019)*, Istanbul, 1(1), 166–170.
2. Memduhoglu, A., & Basaraner, M. (2017). Potential of Spatial Semantics for Developing Multi-Representation Spatial Databases. *UCTEA International Geographic Information Systems Congress*, Adana, 71–75.