

KARADENİZ TEKNİK ÜNİVERSİTESİ
FEN BİLİMLERİ ENSTİTÜSÜ

HARİTA MÜHENDİSLİĞİ ANABİLİM DALI

KİTLE KAYNAK VERİLERİN YETKİLİ VERİLERLE
KARŞILAŞTIRILMASI VE VERİ KALİTESİNİN DEĞERLENDİRİLMESİ

DOKTORA TEZİ

Deniz YILDIRIM

NİSAN 2024

TRABZON



KARADENİZ TEKNİK ÜNİVERSİTESİ
FEN BİLİMLERİ ENSTİTÜSÜ

HARİTA MÜHENDİSLİĞİ ANABİLİM DALI

**KİTLE KAYNAK VERİLERİN YETKİLİ VERİLERLE
KARŞILAŞTIRILMASI VE VERİ KALİTESİNİN DEĞERLENDİRİLMESİ**

Deniz YILDIRIM

**Karadeniz Teknik Üniversitesi Fen Bilimleri Enstitüsünde
"DOKTOR (HARİTA MÜHENDİSLİĞİ)"
Unvanı Verilmesi İçin Kabul Edilen Tezdir.**

Tezin Enstitüye Verildiği Tarih : 16/02/2024

Tezin Savunma Tarihi : 04/04/2024

Tez Danışmanı : Dr. Öğr. Üyesi Gülten KARA

Trabzon 2024

ÖNSÖZ

“Kitle Kaynak Verilerin Yetkili Verilerle Karşılaştırılması ve Veri Kalitesinin Değerlendirilmesi” isimli bu tez Karadeniz Teknik Üniversitesi Fen Bilimleri Enstitüsü Harita Mühendisliği Anabilim Dalı, Doktora Programı’nda hazırlanmıştır.

Tez çalışmam boyunca hiçbir desteğini esirgemeyen ve bilim insanı olma yolundaki adımlarımı atarken kıymetli bilgileriyle ve tecrübesiyle yolumu aydınlatan, önüme çıkan engelleri aşmamda yardımcı olan danışman hocam Sayın Dr. Öğr. Üyesi Gülten KARA’ya teşekkürü bir borç bilirim. Ayrıca, yapıcı eleştirileri ve önerileri ile tezime büyük katkıları bulunan saygıdeğer hocalarım Sayın Prof. Dr. Fevzi KARSLI’ya ve Sayın Prof. Dr. Bekir DİZDAROĞLU’na, eğitim öğretim hayatımda katkısı olan tüm hocalarıma ve ortak olarak çalışmalar yürüttüğümüz ve bu uzun soluklu tez çalışması süresinde ve hayatın geri kalanında hiçbir yardımdan kaçınmayan canım arkadaşım Arş. Gör. Dr. Cemre YILMAZ’a, genel doktora bursu (2211-A) kapsamında destek aldığım TÜBİTAK ve BİDEB birimine, süreçte sürekli olarak danışabildiğim enstitü eski sekreterimiz Sayın Ayşe BAYRAKTAR’a, ihtiyaç duyduğumda yardımını esirgememiş arkadaşım Dr. Öğr. Üyesi Kemal Ilgar EROĞLU’na teşekkür ederim.

Son olarak, doğduğum günden beri elimi hiç bırakmayan, hayatım boyunca yaptığım işlerde bilgi ve tecrübelerinden faydalandığım ve küçük yaşlarımdan beri her kararımda arkamda olan, maddi ve manevi desteklerini hiçbir zaman esirgemeyen ve bugünlere gelmemde en büyük paya sahip olan anneme, babama, ablama ve erkek kardeşime teşekkür ederim. Bu tez çalışmasının bundan sonraki çalışmalara katkı sağlamasını temenni ederim.

Deniz YILDIRIM

Trabzon 2024

TEZ ETİK BEYANNAMESİ

Doktora Tezi olarak sunduğum “Kitle Kaynak Verilerin Yetkili Verilerle Karşılaştırılması ve Veri Kalitesinin Değerlendirilmesi” başlıklı bu çalışmayı baştan sona kadar danışmanım Dr. Öğr. Üyesi Gülten KARA’nın sorumluluğunda tamamladığımı, verileri/örnekleri kendim topladığımı, deneyleri/analizleri ilgili laboratuvarlarda yaptığımı/yaptırdığımı, başka kaynaklardan aldığım bilgileri metinde ve kaynakçada eksiksiz olarak gösterdiğimi, çalışma sürecinde bilimsel araştırma ve etik kurallara uygun olarak davrandığımı ve aksinin ortaya çıkması durumunda her türlü yasal sonucu kabul ettiğimi beyan ederim. 16/02/2024

Deniz YILDIRIM

İÇİNDEKİLER

	<u>Sayfa No</u>
ÖNSÖZ.....	III
TEZ ETİK BEYANNAMESİ.....	IV
İÇİNDEKİLER.....	V
ÖZET	VII
SUMMARY	VIII
ŞEKİLLER DİZİNİ.....	IX
TABLOLAR DİZİNİ.....	XI
SEMBOLLER VE KISALTMALAR DİZİNİ	XII
1. GENEL BİLGİLER	1
1.1. Giriş.....	1
1.1.1. Problem Tanımı ve Motivasyon.....	3
1.1.2. Amaçlar ve Hedefler	7
1.2. Literatür Taraması	7
1.2.1. Yol Ağlarına Yönelik Araştırmalar.....	8
1.2.2. Binalara Yönelik Çalışmalar	11
1.2.3. Ontoloji Tabanlı Kalite Değerlendirme	19
1.2.4. OSM Veri Kalitesi	23
1.3. Temel Tanımlar ve Kavramlar	24
1.3.1. Gönüllü Coğrafi Bilgi	24
1.3.2. Kitle Kaynak Kullanımı	26
1.3.3. OpenStreetMap	28
1.3.4. Yetkili Konumsal Veri Kaynakları ve Özellikleri	32
1.3.5. Konumsal Veri Kalitesi.....	33
1.3.6. ISO 19157:2023 Konumsal Veri Kalite Elemanları	38
1.3.7. Gönüllü Coğrafi Bilgi Kalitesi.....	40
1.3.8. Gönüllü Coğrafi Bilgide Konumsal Doğruluk.....	43
1.3.9. Semantik Web ve İlgili Kavramlar	45
2. YAPILAN ÇALIŞMALAR.....	49
2.1. Genel Çerçeve	49
2.2. SDEO Üst Hiyerarşik Yapısı.....	51
2.2.1. Üst Sınıflar	51
2.2.2. Veri Tipi Üst Özellikleri	52

2.2.3. Nesne Özellikleri.....	53
2.2.4. Notasyon Özellikleri	53
2.3. HGM Verilerinin Hazırlanması.....	53
2.3.1. ESRI Shapefile HGM Verisine sdeo:shortID Değer Atanması ..	56
2.3.2. SDQO ile Kalite Denetimi	59
2.4. OpenStreetMap Verilerinin Hazırlanması.....	61
2.4.1. Veri indirme	61
2.5. HGM Verisinin Ontolojiye ve Dosyalara Aktarılması.....	63
2.5.1. Öznitelik Verileri	64
2.5.2. Geometrik Veriler ve Yapı Geometri Tipine Göre Uygulanan Yöntemler.....	66
2.5.3. OSM ve HGM Verilerinin Karşılaştırılması için Ön İşlemler	68
2.5.4. SAP Yöntemi	73
3. BULGULAR ve TARTIŞMA	81
3.1. Çalışmanın Alanlara Uygulanabilirliği ve Hata Tespiti	81
3.2. Veri Üst Bilgisi.....	85
3.3. OSM İstatistikleri ve İşlemleri	85
3.3.1. Kaynaklar ve Uydu Görüntülerinin Kullanımı	86
3.3.2. Kapalı Yol İncelemesi.....	88
3.3.3. Tarlabahçe	88
3.3.4. Veri Seçme ve İndirme	89
3.4. Dini Yapı Karşılaştırması.....	91
3.5. Sonuç Metrikleri.....	94
3.6. Olası Vandalizm ve İlgili Hatalar.....	98
4. SONUÇLAR ve ÖNERİLER	99
4.1. Metodolojik İnovasyonlar ve Uygulamalar.....	100
4.2. Karşılaşılan Zorluklar ve Çözüm Önerileri	100
4.3. Gelecek Araştırmalar İçin Öneriler	101
4.4. SAP Yöntemi.....	101
5. KAYNAKLAR	103
ÖZGEÇMİŞ.....	119

Doktora Tezi

ÖZET

KİTLE KAYNAK VERİLERİN YETKİLİ VERİLERLE KARŞILAŞTIRILMASI VE VERİ KALİTESİNİN DEĞERLENDİRİLMESİ

Deniz YILDIRIM

Karadeniz Teknik Üniversitesi
Fen Bilimleri Enstitüsü
Harita Mühendisliği Anabilim Dalı
Danışman: Dr. Öğr. Üyesi Gülten KARA
2024, 119 sayfa

Gönüllü Coğrafi Bilgi ve kitle kaynak yaklaşımlarla üretilen verilerin birçok alanda popülaritesinin artmasıyla birlikte, ulusal harita kurumları, verilerini güncellemek için bu yaklaşımlarla üretilen verilerin potansiyelini değerlendirmeye başlamıştır. Bu iki yaklaşımın doğaları gereği, veri kalitesinin değerlendirilmesi gerekmektedir. Bu noktadan hareketle tez çalışmasının amacı, Gönüllü Coğrafi Bilgi ve kitle kaynak yaklaşımlarla elde edilen konumsal verilerin yetkili verilerle karşılaştırılması, kalitesinin değerlendirilmesi ve kullanılabilirliğinin artırılmasıdır. Bu süreçte, Harita Genel Müdürlüğü ve OpenStreetMap veri setlerinin entegrasyonu ve analizi önemli rol oynamaktadır. Bu noktada, Konumsal Veri İyileştirme Ontolojisi geliştirilmiştir; bu ontolojinin güdümünde Trabzon ilinde belirlenen bölgelerde yapılan örnek uygulama ile konumsal veri kalitesini değerlendirme ve iyileştirme hedefinin gerçekleştiği görülmüştür. Ontoloji tabanlı bir yaklaşım kullanılarak ortaya konan metodolojiye ek olarak, iki boyutlu yapılar için konumsal veri kalitesinde, özellikle konumsal doğruluk açısından hedefe ulaştıran yöntem tasarlanmıştır. Bu yöntem için gerekli algoritmalar geliştirilmiş ve MATLAB ortamında uygulanmıştır. Elde edilen sonuçlar, yaklaşımın etkinliğini ve başarısını göstermiştir.

Anahtar Kelimeler: Gönüllü Coğrafi Bilgi, Kitle Kaynak, Konumsal Doğruluk, Ontoloji, OpenStreetMap, Açık Veri.

PhD. Thesis

SUMMARY

COMPARISON OF CROWD SOURCE DATA WITH AUTHORIZED DATA AND ASSESSMENT OF DATA QUALITY

Deniz YILDIRIM

Karadeniz Technical University
The Graduate School of Natural and Applied Sciences
Geomatics Engineering Graduate Program
Supervisor: Asst. Prof. Gülten KARA
2024, 119 Pages

With the increasing popularity of data produced by Volunteer Geographic Information and crowdsourcing paradigms, in many fields, national mapping agencies have started to evaluate the potential of data produced by these paradigms to update their data. Due to the nature of these two paradigms, data quality needs to be assessed. This thesis is dedicated to assessing the quality and applicability of spatial data derived from Volunteer Geographic Information and crowdsourcing approaches by juxtaposing it with data from authoritative sources. Through a case study conducted in a specifically chosen area within the Trabzon province, this research seeks to elevate the quality of spatial data. The principal goal is to refine the existing spatial data's accuracy and accessibility, employing the frameworks of the Spatial Data Quality Ontology and the newly created Spatial Data Enhancement Ontology. A significant portion of this process entails the integration and meticulous analysis of datasets from the Turkish General Directorate of Mapping and OpenStreetMap. Leveraging an ontology-based approach, an algorithm that surpasses existing methods in spatial accuracy has been devised using MATLAB software. The efficacy and success of the proposed method have been validated through its application to real-world datasets, demonstrating its potential to significantly improve spatial data quality.

Keywords: Volunteered Geographic Information, Crowdsourced, Positional Accuracy, Ontology, OpenStreetMap, Open Data.

ŞEKİLLER DİZİNİ

	<u>Sayfa No</u>
Şekil 1. Web 1.0 ve Web 2.0’da veri akışı.	25
Şekil 2. Ulusal KVA’nın kapsamı ve yapısı (Akıncı ve Cömert & Akıncı, 2004).26	
Şekil 3. Basitleştirilmiş OSM kavramsal veri modeli (OpenGeoEdu, 2024).	29
Şekil 4. OSM XML örneği, düğüm (nokta, node).....	30
Şekil 5. OSM XML örneği, çizgi (way).	30
Şekil 6. OSM XML örneği, ilişki (relation)	31
Şekil 7. İçsel ve dışsal veri kalitesi (R. Devillers vd., 2007).	33
Şekil 8. ISO19157 2013 şeması (ISO, 2013).	38
Şekil 9. Veri Kalitesi Elemanları (ISO, 2023).	39
Şekil 10. Semantik Web katmanlı yapısı.....	46
Şekil 11. SWT katmanlı yapısı (Idehen, 2017).	46
Şekil 12. RDF üçlü örneği ex:Deniz ex:livesIn ex:Istanbul.	47
Şekil 13. RDF çizgesi.	47
Şekil 14. Önerilen metodolojinin mimarisi.	50
Şekil 15. Detay Sınıfları.	52
Şekil 16. Çıkarsama sonrası sdeo:ResmıBuyukBına.....	56
Şekil 17. sdeo:shortID için notasyon özelliği ayarlama.	59
Şekil 18. SDQO hiyerarşisi.	60
Şekil 19. SDEO’daki SDQO sınıfları.	60
Şekil 20. OSM verisi indirmek için hazırlanan kullanıcı arayüzü.....	62
Şekil 21. Ontolojide HGM sınıfları hiyerarşisi.	63
Şekil 22. HGM ve OSM etiket verileri ve sdeo:shortID.	66
Şekil 23. Çapa noktası A1H ve bilgileri.	70
Şekil 24. Potsdam patatesleri ve dünya.	71

Şekil 25. sdeo:MBRRotOrigin.	72
Şekil 26. Ara poligon örneği.....	74
Şekil 27. İçbükey köşede iç aç ve alan ilişkisi.	74
Şekil 28. Poligon (mavi) ve ara köşeler (yeşil yıldız).	75
Şekil 29. SAP (siyah), dışbükey kabuk (kırmızı) ve ara köşeler (yeşil).....	76
Şekil 30. Poligon yakınlığı.	77
Şekil 31. sdeo:ConvexHullCoeff.....	78
Şekil 32. sdeo:HGM2 sınıf ve ilişkileri OntoGraf görüntüsü.....	81
Şekil 33. 223447 nolu OSM ilişkisi, Trabzon..	82
Şekil 34. Bölgeler ve HGM küçük bina dağılımları.....	83
Şekil 35. OSM ilişkisi 1236156 (Akçaabat) ve HGM bölgeleri.....	84
Şekil 36. 1km çapındaki yeryüzünde dairesel alanlar.	84
Şekil 37. Akçaabat camilerinin eşleşmesi.	92
Şekil 38. Yeşilyurt yakınında eşleşen cami eşdeğer çifti ve diğer OSM camisi. ..	93
Şekil 39. HGM’de olmayan binalar (sol taraf). HGM binaları (sağ taraf)	95
Şekil 40. Akçaabat merkezi bölgesinden bir parça.....	96
Şekil 41. Akçaabat şehir merkezinden bir parça (SAP).	97
Şekil 42. KTÜ mantıksal SAP raster, HGM(yeşil) ve OSM (kırmızı).....	98

TABLÖLAR DİZİNİ

	<u>Sayfa No</u>
Tablo 1. OSM yol verilerini karşılaştıran çalışmalar.	10
Tablo 2. OSM bina verilerini karşılaştıran çalışmalar.....	18
Tablo 3. Konumsal veri kalitesi sınıflandırması	36
Tablo 4. Bina Kullanım Amacı ve Yapı Fonksiyon Kategorisi.	54
Tablo 5. ESRI Shapefile teknik tanım PoligonZ kayıt verisi bilgisi.	57

SEMBOLLER VE KISALTMALAR DİZİNİ

BAN	: Fransız Ulusal Adres Veri Tabanı
BKG	: Almanya Federal Kartografya ve Jeodezi Kurumu
CBS	: Coğrafi Bilgi Sistemleri
CEN	: Avrupa Standardizasyon Komitesi
CGDI	: Kanada Konumsal Veri Altyapısı
ESRI	: Environmental Systems Research Institute
GCB	: Gönüllü Coğrafi Bilgi
GSI	: Japonya Konumsal Bilgi Kurumu
HGM	: Harita Genel Müdürlüğü
ICA	: Uluslararası Haritacılık Birliği
IGN	: Fransa Ulusal Coğrafi ve Orman Bilgi Enstitüsü
INSPIRE	: Infrastructure for Spatial Information in the European Community
ISO	: Uluslararası Standartlar Teşkilatı
KVA	: Konumsal Veri Altyapısı
NCDCDS	: Ulusal Dijital Kartografik Veri Standartları
NLS	: Finlandiya Ulusal Arazi Araştırması Kurumu
NRCan	: National Resources Canada
NTDB	: Kanada Ulusal Topografik Veri Tabanı
OGC	: Açık Coğrafi Bilgi Konsorsiyumu
OS	: Ordnance Survey
OSM	: OpenStreetMap
OWL	: Web Ontology Language
RCMRD	: Regional Centre for Mapping of Resources for Development
RDF	: Kaynak Açıklama Çerçevesi
SAP	: Sanal Ara Poligon
SDEO	: Konumsal Veri İyileştirme Ontolojisi
SDQM	: Semantik Veri Kalitesi Yönetim Çerçevesi
SDQO	: Konumsal Veri Kalitesi Ontolojisi
SDTS	: Konumsal Veri Transfer Standardı
SOI	: Survey of India
SWT	: Semantik Web Teknolojileri

TIGER	: Topologically Integrated Geographic Encoding and Referencing
TOPOVT	: Topografik Veri Tabanı
URI	: Tekdüzen Kaynak Tanımlayıcısı
USGS	: ABD Jeoloji Araştırmaları Kurumu
WKT	: Well Known Text



1. GENEL BİLGİLER

1.1. Giriş

Coğrafi Bilgi Sistemleri (CBS) modern dünyada, karar verme süreçlerinin ve günlük yaşamın ayrılmaz bir parçası haline gelmiştir. Navigasyon, coğrafi araştırmalar, doğal afet yönetimi, çevre izleme ve kentsel planlama gibi çeşitli alanlarda kullanılan CBS, coğrafi verilerin toplanması, analizi, depolanması ve görselleştirilmesinde güçlü bir araç seti sunar (Uluğtekin ve Bildirici, 1997). Harita Genel Müdürlüğü (HGM) ve diğer harita kurumları, onlarca hatta yüzlerce yıldır güvenilir konumsal veri sağlayıcıları olmuş ve günümüzde yetkili veri diye adlandırılan veriyi üretmiş ve kullanıma açmışlardır. İnternetin ve mobil cihazların yaygınlaşması ile kullanıcılar tarafından hızlı ve erişilebilir konumsal verilere olan ihtiyaç artmış; bu da konumsal veri kaynaklarının doğruluğu ve güvenilirliği üzerindeki baskıyı artırmıştır. Bu bağlamda, bilginin sürekli evrildiği ve her zaman tamamlanmamış olabileceği "açık dünya varsayımı" önem kazanmıştır.

Açık dünya varsayımı, özellikle Gönüllü Coğrafi Bilgi (GCB) toplama süreçlerinde, mevcut bilginin her zaman eksik olabileceği ve yeni verilerin sürekli olarak eklenmesi gerektiği anlayışına dayanır. Bu da konumsal veri analizi ve karar verme süreçlerinde yeni zorluklar ve fırsatlar sunmaktadır.

Öte yandan, "kapalı dünya varsayımı", mevcut veri setinin tüm bilgiyi içerdiği ve hiçbir şeyin dışarıdan eklenmeyeceği beklentisiyle hareket eder. Bu da dinamik dünyamızda sürekli değişen coğrafi olayları ve yapıları yakalama konusunda sınırlılıklar sunar.

Semantik Web Teknolojileri (SWT), konumsal veri yönetimi ve paylaşımında yeni olanaklar sunarak, verilerin sadece insanlar tarafından değil, aynı zamanda makineler tarafından da anlaşılabilir ve işlenebilir olmasını sağlar, konumsal veri kaynakları arasında daha derin entegrasyon ve otomatik akıl yürütme kapasitesi sunar.

GCB ve kitle kaynak yaklaşımları, konumsal veri toplama ve paylaşımında devrim yaratmış, konumsal bilgi topluluğuna, yetkili veri kaynaklarının ulaşamadığı veya göz ardı edilen alanlarda değerli veri sağlamıştır. Bu yaklaşımlarla toplanan verilerin heterojen doğası, çeşitli bilgi türlerini ve kaynaklarını kapsayarak, katkıda bulunanların çeşitli perspektiflerini yansıttığı için başka bir avantajdır. Ancak, kitle kaynağına dayalı verilere olan güven, veri kalitesi, tutarlılığı ve doğrulama ile ilgili zorlukları da beraberinde getirir.

OpenStreetMap (OSM), Ushahidi gibi platformlar, dünya çapında gönüllülerin katkılarıyla zengin konumsal veri tabanları oluşturmuştur. OSM, farklı beceri ve bilgi seviyelerine sahip çok sayıda katılımcıdan veri toplar. Bu geniş katılımcı tabanı, veri katkısının geniş bir yelpazeden geldiğini ve kitle kaynaklı bir yaklaşımı koruduğunu gösterir. Literatürde OSM, sıkça GCB olarak sınıflandırılmaktadır. Bu çalışmada, literatür taraması dışında bu tanımlamadan kaçınılacaktır. OSM'nin veri kaynakları incelendiğinde önemli bir kısmının görevlendirilmiş kişiler, profesyoneller ya da OSM ile anlaşma yapmış ticari kurumlar tarafından üretildiği ya da aktarıldığı görülebilir. Gönüllü katılımcı sayısı düşük değildir ama OSM katkılarının % 90'ı, kullanıcıların yüzde bir civarı tarafından gerçekleştirilmektedir ve bu yüzde birdeki birçok kullanıcı gönüllü değildir (Andersson vd., 2019).

OSM verilerinin açık veri olması, bazı kısıtlamalara engel olmamaktadır. Örneğin Maxar uydu görüntüleri kaynak alınarak oluşturulmuş verilerin OSM dışında kullanımı kısıtlıdır. Çeşitli ülkeler bazı bölgelerin mozaiklenmesini istemiştir.

OSM de Wikipedia gibi, verilerinin geçmişini tutar; eski verilere “tavan-arası” (attic) veri denir. Kısıtlı bölgelerdeki tavan-arası veriler, güncel tarihli verilerden daha güncel olabilir. Kullanıcı düzenlemesine açık veri olmanın yan etkilerinden biri vandalizmdir. Vandalizm sonucunda 11 milyondan fazla OSM kayıtlı düzenleyicisinden 1600 civarı engellenmiştir. Diğer açık veri örneklerinde olduğu gibi vandalizmin üçte ikisi bir gün içinde; kalanın yarısı da bir hafta içinde düzeltilmektedir (Juhász vd., 2020). Vandalizme denk gelme riski düşüktür; tavan-arası veriler kullanılarak bu risk ortadan kaldırılabilir.

Kitle kaynaklı verilerin ve yetkili verilerinin avantajları ve dezavantajları incelendiğinde, bunların birbirini tamamladığı ve doğru bir yöntemle bu verilerin birleştirilerek çok daha iyi bir veri setinin oluşturulabileceği görülebilir. Böylece konumsal veri kaynaklarının doğruluğu, güvenilirliği ve güncelliği üzerinde iyileştirmeler yapılabilir.

SWT ve ontolojiler, konumsal veri yönetimi ve paylaşımı konusunda önemli yenilikler sunarak, verilerin sadece insanlar tarafından değil, aynı zamanda makineler tarafından da anlaşılır ve işlenebilir olmasını sağlar. Bu teknolojiler, CBS sistemlerinde veri entegrasyonu ve analitik işlemleri otomatize ederek, veri setleri arasında daha zengin ve anlamlı bağlantılar kurulmasına olanak tanır.

Ontolojiler, belirli bir alanla ilgili kavramları ve bu kavramlar arasındaki ilişkileri formalize eder, konumsal detaylar, olaylar ve süreçler hakkında bilgi sağlar. Bu, farklı kaynaklardan gelen konumsal verilerin daha iyi anlaşılmasını ve entegre edilmesini

sağlayarak, makinelerin veri üzerinde mantıksal çıkarımlar yapmasına ve konumsal verileri daha zengin ve bağlamsal bir şekilde ifade etmeye yardımcı olur. Bu durum, CBS'nin analitik yeteneklerinin genişlemesine, daha karmaşık sorgulamaların yapılabilmesine ve kullanıcıların konumsal verilere daha kolay erişimine olanak tanır.

SWT uygulama alanları arasında çevresel izleme, afet yönetimi ve şehir planlaması gibi çeşitli alanlar bulunur. Bu teknolojilerin gelişimi, CBS'nin potansiyelini artırarak farklı alanlarda yenilikçi uygulamalara olanak tanımakta ve çalışmalarını daha etkili, erişilebilir ve anlamlı hale getirmektedir.

Bu çalışmada, Yılmaz (2018) tarafından geliştirilmiş olan Konumsal Veri Kalitesi Ontolojisi (Spatial Data Quality Ontology, SDQO) kullanılmış; ayrıca Konumsal Veri İyileştirme Ontolojisi (Spatial Data Enhancement Ontology, SDEO) adında bir ontoloji tasarlanıp geliştirilmiştir. Uygulamalar bu ontolojiler kullanılarak gerçekleştirilmektedir.

1.1.1. Problem Tanımı ve Motivasyon

CBS uygulamalarının etkinliği, büyük ölçüde sağlanan konumsal verilerin kalitesine ve doğruluğuna bağlıdır. Geleneksel olarak bu veriler yerel yönetimler, ulusal harita kurumları ve diğer yetkili kurumlar tarafından sağlanmaktadır.

Ulusal ya da bölgesel düzeyde harita üretiminden sorumlu kurum/kuruluş tarafından belirlenen veri kalitesi kuralları kullanılarak üretilen/kullanılan konumsal veriler günümüzde “yetkili veri” olarak isimlendirilmektedir. Ancak yetkili veri üretiminin yanı sıra, Web 2.0 ve internet teknolojilerindeki gelişmelerle birlikte, kullanıcılar da kitle kaynak ve GCB yaklaşımlarını kullanarak veri üretme sürecine dâhil olmuştur (Turner, 2006; O'Reilly ve Battelle, 2009; Zielstra ve Zipf, 2010).

Birçok alanda gücünü hissettiren ve yaygın kullanıma sahip kitle kaynak yaklaşımla üretilen verilerin kullanımının büyük bir hızla yaygınlaşması; harita kurumlarını da bu yönde teşvik etmiştir. Kitle kaynak veriler, genellikle daha zengin kullanıcı tabanlı bilgiler içerdiğinden, gerçek dünyadaki değişiklikleri (örneğin yeni bina inşaatları) daha hızlı yansıtabildiğinden ve çok daha düşük veri toplama maliyetine sahip olduğundan büyük bir potansiyel sunmaktadır. Kitle kaynak verilerinin doğası gereği konumsal veri kalitesi açısından heterojen olması, bu verilerin doğrudan kullanımının önündeki en büyük engeldir. Kitle kaynak ve yetkili veriler bir arada kullanılırsa, eksik yönler bertaraf edilip daha da kaliteli veriler oluşturulabilir.

Yetkili veri faaliyetleri uzmanlık gerektiren araçların ve yazılımların kullanılmasını gerektirir. Buna bağlı olarak; konumsal veri üretimi, konusunda donanımlı personeller tarafından gerçekleştirilir. Birimlerinde zamanla uzmanlaşmış bu kurum yetkililerinin başka konularda da yetkin olmasını beklemek yerinde değildir, ama alışageldiklerinin dışındaki ilgili diğer konularda temel işleri yapabilmeleri beklenebilir. Buna örnek olarak var olan bir ontolojinin editör kullanılarak hiyerarşik düzenlenmesinin yapılabilmesini verebiliriz. Bu araştırmada üretilen yöntemler, yetkili kurum personelleri tarafından, ontoloji güdümünde kullanılmaya yönelik tasarlanmıştır.

OSM ve Kitle Kaynak Platformları

Kitle kaynak ve yetkili veri arasındaki eşleştirmeler veri entegrasyonu, veri güncelleme ve zenginleştirme gibi uygulamalarda kullanılır. OSM haricinde önde gelen kitle kaynak veya GCB platformları aşağıda listelenmiştir.

OSM: 2004 yılında geliştirilmiştir. Kullanıcılar tarafından etiketlenmiş milyarlarca objeyi içeren en yaygın konumsal kitle kaynak platformudur.

Ushahidi: Kriz durumlarında, olaylarda veya genel olarak herhangi bir konuda kitle kaynaklı bilgi toplama ve haritalama amacıyla kullanılan açık kaynaklı bir platformdur. 2008 yılında Kenya'daki şiddet olaylarını haritalamak amacıyla geliştirilen bu platform, dünya genelinde çeşitli sosyal, politik ve çevresel olayların raporlanması ve analiz edilmesi için kullanılmaktadır.

Mapbox: Ticari bir hizmet olan Mapbox, özelleştirilebilir haritalar sunar ve geliştiricilere güçlü araçlar sağlar (Kastanakis, 2016). Kullanıcılar tarafından sağlanan verilerle desteklenen geniş bir veri tabanına sahiptir.

Google Map Maker: Google, kullanıcıların haritalarına katkıda bulunmalarını sağlayan Map Maker aracını sunmuştu (Pfund, 2008). Ancak, Google bu hizmeti 2017 yılında durdurdu ve kullanıcı katkılarını doğrudan Google Maps üzerinden almaya başladı (Google, 2017).

HERE WeGo: Kullanıcı katkıları ile desteklenen bir başka haritalama servsidir. Araç sürücülerini için yol tarifleri, trafik bilgileri ve çeşitli taşımacılık modlarına yönelik navigasyon hizmetleri sunar (Alomari vd., 2020).

Wikimapia: Kullanıcılar tarafından etiketlenmiş milyonlarca objeyi içeren, kitle kaynaklı bir web haritası ve coğrafi ansiklopedidir. Wikimapia, kullanıcıların harita üzerine yerler eklemesine ve düzenlemesine izin verir (Ballatore ve Arsanjani, 2019).

Mapillary: Sokak görünümü görüntüleri toplamak için kitle kaynak yaklaşımını kullanan bir hizmettir; 2020 yılında Facebook (Meta) tarafından satın alınmıştır (Juhász ve Hochmair, 2016; Mapillary, 2020). Kullanıcılar, dünya çapında sokak seviyesinde fotoğraflar çekip yükleyerek Mapillary veri tabanını genişletirler. OSM üzerinden source: etiketi yerine mapillary: etiketi ile Mapillary görüntülerine direkt ilişkilendirme yapılabilir (OSM Wiki, 2022).

OpenSeaMap: Denizcilik ve su yolları odaklı bir kitle kaynaklı harita projesidir. Denizler, göller ve nehirler üzerinde navigasyon bilgileri, liman bilgileri ve diğer denizcilikle ilgili verileri içerir (Bärlocher, 2013). OSM'nin bir parçasıdır.

OpenTopoMap: 2011 yılında geliştirilmiş, topografik haritalar sunan kitle kaynaklı bir projedir, OSM'nin bir parçasıdır (Erhardt, 2016).

Dünyada Yetkili Veri Kaynakları

Yetkili veri olarak nitelendirilen veriler ulusal harita kurumları veya konumsal veri üreticileri tarafından oluşturulur.

Harita Genel Müdürlüğü (HGM), 1895 yılında kurulmuştur.

Ordnance Survey (OS), Birleşik Krallık'ın ulusal harita kurumudur, coğrafi veri toplama ve dağıtımında öncü bir role sahiptir. 1783 yılında kurulmuştur ve iki asrı aşan bir geçmişe sahiptir (Ordnance Survey, 2023).

Almanya Federal Kartografya ve Jeodezi Kurumu (Bundesamt für Kartographie und Geodäsie, BKG), Almanya'nın ulusal kartografya ve jeodezi kurumudur (Nalci, 2011). 70 yılı aşkın bir geçmişe sahiptir (BKG, 2023).

Fransa Ulusal Coğrafi ve Orman Bilgi Enstitüsü (Institut National de l'Information Géographique et Forestière, IGN), Fransa'nın CBS ve orman bilgileri ile ilgili ulusal kurumudur ve 1940 yılında kurulmuştur (IGN, 2023).

Kartverket, Norveç'in CBS, haritalama, arazi ölçümü, kadastro ve konumsal veri yönetimi alanlarında lider kurumudur. Ülkenin coğrafi verilerinin toplanması, işlenmesi, dağıtılması ve yönetilmesinden sorumludur. CC 4.0 ile sundukları ücretsiz veriler, OSM verilerinde kaynak olarak kullanılmaktadır (Kartverket, 2021; OSM Wiki, 2023a).

Maanmittauslaitos, Finlandiya Ulusal Arazi Araştırması Kurumu, İngilizce yaygın adıyla National Land Survey of Finland (NLS) Finlandiya'nın ulusal kartografya ajansıdır. Kurum on yedinci yüzyılda bölge İsveç egemenliğinde iken kurulmuş, ülke bağımsızlığı kazandıktan sonra günümüzdeki ismini almıştır (NLS, 2023).

Survey of India (SoI), Hindistan'ın ulusal harita yapım kurumu, konumsal veri ve haritaların üretiminde liderdir. On sekizinci yüzyılda kurulmuştur ve sadece askeri birimlere açık verinin yanında Open Series Map adı altında satışı halkına, kısmen ücretsiz olarak açık olan veri de sunar (SoI, 2023).

Geospatial Information Authority of Japan (GSI), Japonya'nın on dokuzuncu yüzyılda kurulmuş 2010 yılında şimdiki ismini almış coğrafi bilgi ve haritalama ajansıdır; verileri ücretli olarak satışıdır (GSI, 2023).

Afrika Kalkınma için Kaynakların Haritalandırılması Bölgesel Merkezi (Regional Centre for Mapping of Resources for Development, RCMRD) 1975 yılında kurulmuştur; Afrika kıtasında doğal kaynakların yönetimi ve sürdürülebilir kalkınma için CBS, uzaktan algılama ve diğer konumsal teknolojilerin kullanımını destekleyen bir hükûmetler arası kuruluştur (RCMRD, 2023).

ABD Jeoloji Araştırmaları Kurumu (United States Geological Survey, USGS), ABD'nin konumsal bilgiler, doğal kaynaklar konusunda bilgi sağlayan bir hükûmet kurumudur. Landsat uydu görüntüleri, düşük çözünürlüklü olsalar da OSM'de kaynak olarak kullanılmaktadır. Hava kuvvetlerinin idaresindeki GPS verileri ve Okyanus ve Atmosfer İdaresi'nin PGS verileri de OSM kaynakları arasında yer alır (OSM Wiki, 2019; 2023b; USGS, 2023).

Natural Resources Canada (NRCan) Kanada'nın doğal kaynakları ve konumsal bilgi altyapısını yöneten kurumudur. Canadian Geospatial Data Infrastructure (CGDI) Kanada için üretilmiş Konumsal Veri Altyapısı (KVA)'dır.

Ayrıca Avrupa'nın KVA'sı, Infrastructure for Spatial Information in the European Community (INSPIRE), Avrupa Birliği düzeyinde kabul görmüş bir çerçeve olduğu için, direktif kapsamında toplanan, işlenen ve paylaşılan veriler yetkili ve güvenilir veri kaynakları olarak kabul edilir. Direktif, veri kalitesi ve erişilebilirlik standartlarını belirleyerek, kullanıcıların ihtiyaç duyduğu konumsal bilgilere tutarlı bir şekilde ulaşmalarını sağlar (INSPIRE, 2023). Bu nedenle, INSPIRE tarafından yönlendirilen veya desteklenen konumsal bilgi kaynakları, araştırma, planlama ve politika yapımı gibi çeşitli amaçlar için yetkili ve güvenilir veri kaynakları olarak değerlendirilebilir.

1.1.2. Amaçlar ve Hedefler

Bu çalışmanın temel amacı, Türkiye'deki yetkili verilerin ve kitle kaynaklı veri kaynağı olan OSM verilerinin karşılaştırılması ve kalitesinin değerlendirilmesidir. Ayrıca, bu iki tür veriden daha kaliteli ve güvenilir veri üretme potansiyelini incelemek de hedeflenmektedir. Bu amaçlar doğrultusunda, aşağıdaki hedefler belirlenmiştir:

- Kitle kaynaklı ve yetkili veri sistemlerinin veri kalitesi ve güvenilirlik sorunlarına dair mevcut literatürü araştırmak.
- Kitle kaynaklı ve yetkili veri sistemlerinin genel çalışma prensiplerini, benzer sistemleri ve bu sistemlerin veri kalitesi üzerindeki etkilerini incelemek.
- OSM ve yetkili verilerin temel özelliklerini karşılaştırmak, analiz etmek ve bu iki veri kaynağından elde edilen verilerin kalitesini değerlendirmek.
- OSM verilerinin konumsal doğruluğunu değerlendirmek ve yetkili verilerle karşılaştırmalı bir doğruluk analizi gerçekleştirmek.
- OSM ve yetkili verilerin birleştirilmesiyle oluşturulabilecek daha kapsamlı ve doğru veri setlerinin potansiyelini araştırmak.
- Veri kalitesi ve doğruluğunu değerlendirmek için çeşitli yöntemler kullanmak.

1.2. Literatür Taraması

Konumsal verilerin daha eksiksiz, güncel, insan odaklı ve daha zengin bir veri içeriğinin sağlanması için yetkili kurum ve kitle kaynaklı verilerin birbirini tamamlayacak şekilde kullanılması arzu edilir. Ulusal harita kurumlarının yetkili verilerin zenginleştirilmesine ve güncellenmesine yardımcı olmak için kitle kaynaklı konumsal verileri kullanmaktadır. Bu bağlamda kitle kaynak verilerinin yetkili verilerle karşılaştırılmasına yönelik çalışmalar incelenmiştir. OSM verilerinin kalitesini farklı yönleriyle değerlendiren birçok çalışma literatürde mevcuttur. Bunların çoğu özellikle yol ağı ve bina verilerinin kalitesinin değerlendirilmesine odaklanmıştır. Aşağıda, özellikle OSM-yetkili yol verileri ve OSM-yetkili bina verileri arasındaki karşılaştırmaya odaklanan en önemli çalışmaların bir incelemesi sunulmuştur.

1.2.1. Yol Ağlarına Yönelik Araştırmalar

Brovelli vd. (2015), FOSS4G (Özgür ve Açık Kaynaklı Coğrafi Bilgi Sistemleri) tabanlı bir prosedür kullanarak OSM ile yetkili yol ağı veri setlerini karşılaştırarak OSM'nin konumsal doğruluğunu ve tamlığını değerlendirmeyi amaçlamaktadır. OSM ve herhangi bir yetkili yol ağı veri setinin otomatik kalite karşılaştırması için FOSS4G tabanlı bir prosedür önerilmiştir. Önerilen prosedür bir GRASS komutuna uygulanmıştır ve bunun bir QGISplugin ve FOSS4G tabanlı bir WPS'ye genişletilmesi planlanmıştır. Karşılaştırma prosedürü, geometrik ve topolojik karşılaştırmalar gibi çeşitli ölçümleri ve analizleri içerir. Bu analizler, OSM'nin konumsal doğruluğunu ve tamlığını değerlendirmek için önemli bir çerçeve sunar. Özellikle, geometrik karşılaştırmalar, yol geometrileri arasındaki farklılıkları inceleyerek, OSM'nin gerçek dünyadaki yollarla uyumluluğunu değerlendirir. Topolojik karşılaştırmalar ise, OSM'nin yol ağının yetkili veri setleriyle kesişimleri, parçalı yollar gibi topolojik özellikler açısından uyumlu olup olmadığını belirlemeye yardımcı olur. Ayrıca, OSM'nin güncellik sürecini değerlendirmek için zaman analizi ve eksikliklerin belirlenmesi için fark analizleri gibi ek yöntemler de ele alınmaktadır.

Dorn vd. (2015), OSM arazi kullanımı ve arazi örtüsü (Land Use and Land Cover-LULC) verilerinin kalitesi güney Almanya'daki bir alan için incelenmiştir. İki konumsal veri kalitesi unsuru, tematik doğruluk ve bütünlük, OSM verileri yetkili bir Alman referans veri seti ile karşılaştırılarak ele alınmıştır. Sonuçlar, kappa değerinin OSM ve yetkili veri seti arasında önemli bir uyum olduğunu göstermektedir. Bununla birlikte, çalışma bölgesi için LULC sınıfları arasında açık farklılıklar vardır. OSM orman veri setinin tamlığı (%97,6) doğruluğu (%95,1)'dir. Buna karşılık, OSM tarım arazilerinin tamlık değeri (%45,9)'dur.

Brovelli, Minghini ve Molinari (2016), OSM ve IGN yol ağları veri setlerini karşılaştırmıştır. GRASS GIS yazılımı kullanılarak geliştirilen bir metodoloji ile OSM yol ağının konumsal doğruluğu ve tamlığı incelenmiştir. Önerilen metodoloji, OSM ve IGN yol ağı veri setiyle karşılaştırmak için geliştirilen otomatik bir prosedüre dayanmaktadır. Açık kaynak kodlu GRASS GIS yazılımı için üç tane Python modülüne dayanan bu yöntem, OSM yol ağının konumsal doğruluğu ve tamlığına ilişkin ölçümler sağlamaktadır. OSM ve yetkili kurum verisini karşılaştıracak kullanıcı, yetkili veri setine aşina olması koşuluyla, prosedürün esnekliği sayesinde ilgili parametrelerin değerlerini ayarlayabilir. Yöntem, Paris OSM yol ağı veri setinin kalitesini, IGN tarafından sağlanan yetkili kurum veri setiyle karşılaştırarak değerlendirmek için uygulanmıştır. Sonuçlar, Paris OSM yol ağının hem

yüksek bir tamlığa hem de konumsal doğruluğa sahip olduğunu göstermektedir. OSM yol veri setindeki yol ağının uzunluğu IGN yol ağından daha uzundur ve 5-6 m'ye kadar konumsal doğruluk gerektiren uygulamalar için uygun olduğu bulunmuştur. Ayrıca sonuçlar, kullanıcıların OSM ve referans yol veri setleri arasında kendi karşılaştırmalarını yapmalarını desteklemek için prosedürün esnekliğini doğrulamaktadır.

Mahabir vd. (2017), Kenya, Nairobi ve çevresindeki nüfus verileriyle birlikte yetkili (Regional Center for Mapping of Resources for Development - RCMRD) ve kitle kaynak yol verileri arasındaki benzerlikleri ve farklılıkları belirlemek için çeşitli analizler yapmıştır. Bu analizler arasında, geometrik karşılaştırma, topolojik uyumsuzlukların belirlenmesi, veri eksikliklerinin tespiti ve veri doğruluğunun değerlendirilmesi yer alır. RCMRD, tüm çalışma alanı dikkate alındığında daha az güncel olsa da tam kapsamı sağlarken, OSM ve Google Map Maker, Nairobi'nin merkezinden kırsal alanlara doğru gidildikçe kapsamın azaldığını göstermiştir. OSM ayrıca büyük gecekondü mahallelerinde daha yüksek içerik yoğunluğuna sahip olup bu konumlarda yetkili veri setlerini geride bırakırken, Google Map Maker kırsal konut alanlarında daha iyi kapsama alanı göstermiştir. Google Map Maker'dan alınan Kenya'nın Nairobi kentindeki yol verileri incelenmiş ve bölgedeki nüfus yoğunluğu ile yol kapsamı arasındaki ilişki analiz edilmiştir. Yol kapsamının tamlığı, 2014 yılında elde edilen OSM ve Google Map Maker verileri arasında karşılaştırılmıştır. Çalışma, OSM verilerinde "sınıflandırılmamış" olarak sınıflandırılan yolların Google Map Maker ve RCMRD deki yollarla örtüştüğünü, RCMRD ve Map Maker'da OSM'ye kıyasla daha sınırlı bir kapsama alanı olduğunu göstermiştir. Nairobi ve çevresindeki yol verilerinin özet istatistikleri, OSM 2011, OSM 2014 ve Google Map Maker'ın Nairobi'nin merkezinde daha yoğun kapsama alanına sahip olduğunu göstermiştir. Çalışma ayrıca yol verilerinin zaman içindeki gelişimini analiz etmek için OSM 2011 ile RCMRD ve OSM 2014 üç açıdan karşılaştırılmıştır: (i) aynı zaman dilimi için kapsama alanı açısından yetkili ve GCB yol verilerinin karşılaştırılması; (ii) sürekli güncellenen ve mevcut GCB verilerinin kapsamının yetkili veri setlerinin sabit güncelleme döngüleriyle nasıl karşılaştırıldığının değerlendirilmesi ve (iii) çalışma alanındaki nüfus yoğunluğu ile yol kapsamı arasındaki ilişkinin araştırılmasıdır.

Brovelli vd. (2017), OSM yol ağı verilerinin geometrisini diğer yol veri kümeleriyle karşılaştırmak için esnek bir metodoloji önermiştir. OSM yol verilerinin diğer harita veri tabanlarıyla uyumluluğu da dahil olmak üzere OSM'nin tamlığı ve konumsal doğruluğu için nicel ölçütler hesaplanmıştır. Çalışmada önerilen metodoloji, kullanıcılara parametreleri

kendi uygulamalarına göre deęiřtirme esneklięi saęlar. Bu yazılım açık kaynak kodlu olarak geliřtirilmiřtir ve OSM yol verilerinin dięer türlerine uyarlanabilir.

Silva ve Camboim (2021) Brezilya ulusal harita kurumu tarafından üretilen haritaların tamlięını ve doęruluęunu saęlayacak bir metodoloji önermiřtir. Metodoloji, yetkili ve kitle kaynak verilerinde benimsenen kavramsal modeller arasındaki semantik uyumluluęu test ederek bařlar, veri entegrasyonu için en önemli potansiyele sahip detay sınıflarının belirlenmesiyle devam eder. Son olarak, OSM veri entegrasyon sürecini 1:25.000 ölçekli bir kartografik veri tabanında deęerlendirmek ve doęrulamak için bazı yöntemler geliřtirilmiřtir. Yetkili verinin güncel olarak oluřturulduęu bölgelerde, ön deęerlendirmenin sonuçları yol detaylarında yaklaşık % 52 ve % 16'lık bir artış potansiyeline iřaret etmektedir; bu da eksik veya güncel olmayan veri anlamına gelir.

Sarretta ve Minghini (2021) OSM ve açık verilerin Avrupa ulusal harita kurumlarının yetkili konumsal veri setleriyle entegre edilmesinin artı ve eksilerini incelemiřtir. Veri entegrasyonunu yalnızca yerel veya bölgesel düzeyde test eden çalışmaların aksine, Finlandiya NLS tarafından yayınlanan ulusal adres veri setini OSM'nin ilgili veri setiyle entegre etmek için bir çalışma yapmıřtır. Çalışma, iki veri setinin analizini, veri modelleri arasındaki eřleřtirmeyi ve içeriklerini dönüřtürmek ve nihai olarak birleřtirmek için açık kaynak kodlu QGIS yazılımı kullanılarak gerçekleřtirilen bir dizi iřlem adımını içermektedir. Sonuçta ortaya çıkan veri seti, NLS'den alınan adreslerin genel olarak Finlandiya genelinde daha tam olmasına raęmen, bazı bölgelerde OSM adreslerinin yetkili veriyi tamamlamak için daha ayrıntılı ve daha güncel bilgiler saęladığını doęrulamaktadır.

OSM yol verilerini referans bir veri seti ile karřılařtıran çalışmalar Tablo 1 üzerinde gösterilmiřtir.

Tablo 1. OSM yol verilerini karřılařtıran çalışmalar.

	Bölge	Referans Veri	Kalite kriteri
Ather (2009)	Londra	Britanya NMA	Tamlık, konumsal doęruluk
M. Haklay (2010)			Tamlık, konumsal doęruluk
Girres ve Touya (2010)	Fransa	IGN	Tamlık, konumsal doęruluk, öznitelik doęruluęu, semantik doęruluk, tamlık, mantıksal tutarlılık, zamansal doęruluk, soy ve kullanım amacı
Zielstra ve Zipf (2010)	Almanya	TeleAtlas	Tamlık
Helbich vd. (2012)		TomTom	Konumsal doęruluk

Tablo 1'in devamı

Ludwig vd. (2011)	Almanya	Navteq	Tamlık
Neis vd. (2011)	Almanya	TomTom	Tamlık, öznitelik doğruluğu
Zielstra ve Hochmair (2011)	Florida	Navteq ve TeleAtlas	Tamlık
Kounadi (2009)	Atina	Yunan Askeri Coğrafi Servisi	Tamlık, konumsal doğruluk, isim doğruluğu, uzunluk.
Ciepluch vd. (2011)	İrlanda	OS İrlanda	Tamlık
Arsanjani vd. (2013)	Almanya	BKG	Tamlık, konumsal doğruluk, semantik doğruluk
Zheng ve Zheng (2014)	Çin	Çin	Tamlık, konumsal doğruluk
Brovelli vd. (2015)	İtalya	Lombardiya	Tamlık, Konumsal doğruluk
Brovelli, Minghini, Molinari, vd. (2016)	Paris, Fransa	IGN	Tamlık, konumsal doğruluk
Antunes vd. (2015a)	Portekiz	Coimbra Belediyesi	Konumsal Doğruluk
Mahabir vd. (2017)	Nairobi, Kenya	RCMRD	Verinin kapladığı alan (Coverage)

1.2.2. Binalara Yönelik Çalışmalar

OSM Bina verileri farklı kullanıcılar tarafından farklı teknikler ve farklı veri toplama kaynaklarından elde edilmektedir. Veriler, GPS cihazları kullanılarak, hava fotoğraflarından/uydu görüntülerinden sayısallaştırma (armchair-mapping), sokak düzeyinde ölçümler, ulusal harita kurumları tarafından eklenen veriler ve diğer mevcut konumsal verilerden elde edilmektedir. OSM bina verileri, veri toplama yöntemine, kullanıcı algısına göre bina dış çizgisinin genelleştirilme düzeyine ve OSM'yi düzenleyen kişinin beceri ve deneyim düzeyine bağlı olarak heterojen doğruluk ve ayrıntı seviyesine sahiptir. Ulusal harita kurumlarının veri zenginleştirme ve güncellemesini engelleyebilecek faktörler arasında sınırlı mali kaynaklar, idari kısıtlamalar ve bazı durumlarda gizlilik ve veri kullanımıyla ilgili sıkı düzenlemeler vardır. Bu nedenle özellikle dünya çapında bina verilerini içeren ve kullanıcılara açık olan en kapsamlı kaynak hiç şüphesiz OSM'dir. Bu bağlamda OSM bina verilerinin kalitesini referans veri seti ile karşılaştırarak değerlendiren çalışmalar aşağıda verilmiştir:

Fan vd. (2014) OSM'deki bina izi verilerinin kalite değerlendirmesine odaklanarak Almanya'nın Münih şehri için bir uygulama gerçekleştirmiştir. Veriler, tamlık, semantik doğruluk, konum doğruluğu ve şekil doğruluğu açısından ATKIS (Alman Yetkili Topografik-Kartografik Bilgi Sistemi) verilerindeki bina izleri kullanılarak değerlendirilmiştir. Uygulama için, OSM ve ATKIS verileri arasında eş bina izlerini bulma, dört kalite kriterinin parametrelerini hesaplama ve istatistiksel analiz işlemleri gerçekleştirilmiştir. İki poligon arasındaki benzerliği ölçmek için Turning fonksiyonu (veya tangent function) kullanılmıştır. Kardinalite türüne göre poligonlar için benzer noktalar belirlenir. Sonuçlar, Münih'teki OSM izi verilerinin yüksek bir tamlık ve semantik doğruluğa sahip olduğunu göstermektedir. Konum doğruluğu açısından ortalama olarak yaklaşık 4 metrelik bir kayma bulunmaktadır. Şekil açısından, OSM bina izleri, ATKIS verilerine oldukça benzerlik göstermektedir ancak bina özneteliklerinin eksik olduğu görülmüştür. Bununla birlikte, bazı mimari detayların eksik olduğu ve OSM bina izlerinin, ATKIS verilerinin basitleştirilmiş bir versiyon olarak kabul edilebildiği belirtilmiştir.

Bina çıkarımı için değerlendirme yöntemlerinin standartlaştırılması problemine odaklanarak Avbelj vd. (2015), iki poligonu karşılaştırmak için “PoLiS metriği” önermiştir. PoLiS metriği, bir üçgen eşitsizliğini karşılayan pozitif tanımlı ve simetrik bir fonksiyondur. Çokgenler arasındaki şekil ve doğruluk farklılıklarını hesaba katar, uygulaması kolaydır ve eşik gerektirmez. İki çokgen arasındaki PoLiS metriğinin küçük öteleme, döndürme ve ölçek değişikliklerine göre yaklaşık doğrusal olarak değiştiği belirlenmiştir. Ayrıca, sayısal yüzey modelinden çıkarılan bina poligonları, PoLiS, Hausdorff ve Chamfer mesafeleri hesaplanarak referans bina poligonlarıyla karşılaştırılmıştır. Sonuçlar, poligonlar arasındaki farklılığın PoLiS mesafesi ile ölçülmesinin görsel algı ile tutarlı olduğunu göstermiştir. Hausdorff ve Chamfer mesafeleri, bir çokgenin diğerinden daha fazla köşeye sahip olduğu durumlarda, farklılığı olduğundan fazla göstermektedir. Bu nedenle, şekil benzerliği alanında daha geniş uygulamalara da sahip olabilecek yapı çıkarma değerlendirmesini standartlaştırmaya yönelik bir yaklaşım önerilmiştir. Dikdörtgen ve uzun şekilli binalar iyi tahmin edilirken birkaç küçük yapıya sahip binalarda daha kötü tahminler yapılmıştır. Tez çalışması kapsamında önerilen çalışmada kitle kaynak ve yetkili veri setindeki binaların Convex Hull ları arasında karşılaştırma yapıldığı için şekil tahminleri daha doğru sonuç vermiştir.

Brovelli, Minghini, Molinari, vd. (2016) OSM binalarının konumsal veri kalitesini değerlendirmek için OSM bina verilerini referans veri setiyle karşılaştıran bir yöntem

önermiştir. Yöntemle iki veri seti arasındaki eş binalar yarı otomatik olarak belirlenir. OSM binalarının referans veri setindeki eşdeğer binalara konumsal olarak eşleştirilmesi için Warping algoritmaları (affine transformation, multi-resolution splines, vb.) kullanılmıştır. Afin dönüşüm uygulandıktan sonra eş binaların belirlenmesi amacıyla her bir bina detayının bina köşelerinde benzer noktaların belirlenmesi için istatistiksel mesafe kullanılır. Seçilen benzer binaların eşleştirme doğruluğunu artırmak için yönlü analiz gerçekleştirilir. Bu işlem, her homolog çiftindeki noktalardan çıkan tüm çizgi parçalarının açılarının karşılaştırılmasını içerir ve tolerans eşiğinin altında olması durumunda, doğrulanır.

OSM gibi kitle kaynak veri setleri ile yetkili veri setleri arasındaki tutarsızlıkların azaltılmasına yönelik bir çalışma (Du vd., 2017), tarafından gerçekleştirilmiştir. Bu tutarsızlıkların, her iki veri kaynağının farklı veri toplama yöntemlerinden kaynaklandığı belirtilmiştir. OSM ile İngiltere (Ordnance Survey) - Fransa (Institut national de l'information géographique et forestière (IGN) – France National Institute of Geographic and Forest Information) ulusal harita kurumlarının konumsal verilerini eşleştirmek ve tutarsızlıkları gidermek amacıyla bir yöntem önerilmiştir. Yöntem, ilk olarak iki bina veri seti arasında bir benzerlik skoru hesaplamak için belirli özelliklerin kullanılmasını içerir. Bu özellikler, geometrik ve topolojik öznitelikler, yol tipleri, yer isimleri gibi çeşitli kriterlerden oluşur. Ardından, bu benzerlik skorlarına dayanarak en iyi eşleşmeyi belirlemek için eşleştirme algoritmaları kullanılır. Önerilen yöntem farklı veri kaynaklarındaki konumsal detayları eşleştirmek için “geometrikleri eşleştirme” ve “konumsal nesneleri eşleştirme” olmak üzere iki işlem adımını içerir. Geometri eşleştirme 'possibly partOf' ve 'possibly sameAs' kavramlarına dayanmaktadır. Yöntem, yetkili ve kitle kaynak verilerin eşleştirilmesini sağlarken konumsal verinin güncellenmesi ve veri zenginleştirme açısından bir çözüm sunmamaktadır.

Y. Xu vd. (2017) OSM bina verilerinin tamlığını, konumsal doğruluğunu, şekil doğruluğunu, semantik doğruluğunu ve yönelme tutarlılığını değerlendirmek için OSM ve referans bina verileri arasında karşılaştırma yapmıştır. OSM bina izi verilerinin kalitesini derin otokodlayıcı ağı (deep autoencoder network) kullanarak değerlendirmek için önerilen bu yöntem, veri tamlığını, konumsal doğruluğu, şekil doğruluğunu, semantik doğruluğu ve yönelme tutarlılığını kalite değerlendirme ölçütleri olarak hesaplar. Hesaplanan ölçütler girdi olarak kabul edilir ve eğitilmiş bir ağ tarafından kodlanır ve yeniden oluşturulur. Sonuç olarak, anormal bir performans gösteren OSM verileri daha büyük yeniden oluşturulmuş

hataya sahip olacaktır ki bu da OSM verilerinin kalitesine karşılık gelir; ortalama değerden daha büyük olan anormal bir veri noktası, OSM'de daha iyi kaliteye sahiptir.

Brovelli ve Zamboni (2018), bina ayak izi verilerinin konumsal doğruluğunu kontrol etmek için bir referans veri seti ile karşılaştırmaya dayalı bir detay eşleştirme yöntemi sunmaktadır. İki veri seti arasında eş poligonların belirlenmesi için otomatik bir arama algoritması önerilmiştir. Ayrıca, detay eşleştirme ve benzerlik kontrolünden binalar tespit edilebilmekte ve böylece bir tamlik indeksi de hesaplanabilmektedir. Bu süreç, Kuzey İtalya'da 23.900 km²'lik bir alanı kapsayan ve sırasıyla OSM'de yaklaşık 1 milyon bina ve yetkili veri kümesinde 2,8 milyon bina içeren bir bölge olan Lombardiya'da uygulanmıştır. Karşılaştırmanın sonuçları, OSM binalarının konumsal doğruluğunun en azından 1:5000 ölçeğindeki referans veri setinin kalitesiyle uyumlu olduğunu göstermektedir, çünkü yetkili veriye göre ortalama sapma 3 m'lik beklenen toleransın altındadır. Yetkili veri setinde olan ve OSM'de bulunmayan bina sayısı açısından verilen tamlik analizi, tüm bölgede ortalama %57'ye eşit bir yüzde göstermektedir. Ancak, bunun tersinin, yani OSM'de olup referans veri setinde olmayan bina sayısının sıfır olmayıp %9'a karşılık geldiğini belirtmek gerekir. Bu nedenle OSM bina verisi, 1:5000'e karşılık gelen bir doğruluğun gerekli olduğu tüm durumlarda doğrudan kullanım (bölgesel çerçeveler, harita navigasyonu, kentsel analiz, vb.) ve türetilmiş kullanım (tematik haritaların üretimi için arka plan) için geçerli bir temel harita olarak kabul edilebilir. İki veri setindeki binaların oluşturulması ve güncellenmesi prosedürlerinde doğal (ve kaçınılmaz) farklılıklara sahip olduğundan, önerilen metodoloji yanlışlıkla yıkılmış binaların yerine yaklaşık olarak aynı konumda yeni binalarla eşleşebilir.

Basaraner (2020), İstanbul'un bazı bölgelerine ait büyük ölçekli topografik verilerle OSM bina izlerinin geometrik ve semantik kalite değerlendirmelerine odaklanmaktadır. Karşılaştırma, geometrik eşleşme oranına göre birbirine eşlenmiş bina detayları temel alınarak gerçekleştirilir. Geometrik olarak, konum (örneğin, X, Y), boyut (örneğin, alan, çevre ve granülarite), şekil (örneğin, konveksiyon, dairesellik, uzunluk, eşdeğer dikdörtgen indeksi, dikdörtgenlik ve pürüzlülük indeksi) ve yönelim (örneğin, yönelim açısı) unsurlarının çeşitli parametreleri hesaplanmıştır ve karşılaştırılmıştır. Geometrik kalite bulgularına göre, ortalama konumsal fark üç metreden az bulunmuştur. Ayrıca, çevre değerlerinin azalma eğiliminde olduğu, alan ve granülarite değerlerinin ise TOPO verilerine göre OSM verilerinde artma eğiliminde olduğu görülmüştür. Bu, OSM bina izlerinin genel olarak TOPO bina detaylarından daha düşük detay seviyesine sahip olduğunu ortaya koymuştur. OSM veri setinin ölçeğinin, homojen olmasa da, genellikle orta ölçekli

haritaların alt sınırına (yani 1:10.000) veya biraz daha küçük bir ölçeğe yaklaştığı tespit edilmiştir. Semantik kalite bulgularına göre, belirli bir tür tanımının varlığı durumunda, OSM ve TOPO bina detayları arasındaki tutarlılık oldukça yüksek çıkmıştır. Önerilen yaklaşımın, geometrik ve semantik yönlerden eşleşen binaların değerlendirilmesinde büyük ölçüde başarılı olduğu belirtilmiştir.

Küçük ve Anbaroğlu (2020), OSM bina verisinin konumsal doğruluğunu, HGM tarafından üretilen Türkiye Topografik Vektör Veri Tabanı (TOPOVT) ile 'bina' katmanı bağlamında değerlendirmiştir. Çalışmadaki amaç, TOPOVT'de güncelleme gerektiren konumların otomatik olarak tespit edilmesi ve OSM verileriyle güncellenebilirliğinin araştırılmasıdır. Bu amaçla, çalışmada OSM verisinin konumsal doğruluğunu değerlendirmek için OSM ve TOPOVT'deki eşleşen poligonlar arasındaki Hausdorff mesafesi hesaplanmıştır. Ayrıca önerilen metodolojide eşleşen poligonların tespiti için örtüşme yöntemi ve Centroid yöntemi kullanılmıştır. Hausdorff mesafesi sadece her iki katmanda da (OSM ve TOPOVT) kesişen binalar için hesaplanmıştır. Her iki katmanda da kesişen poligonların aynı konumsal detay olduğu kabulüyle, hesaplanan Hausdorff mesafesiyle söz konusu konumsal detayın yani bu çalışma için "bina"nın konumsal doğruluğu hesaplanmış olur. Çalışmada yapılan uygulama, örnek olarak Ankara'nın kırsal ve kentsel kesiminde iki farklı konumdaki veriye uygulanmıştır. Yazarlar, kırsal kesimdeki OSM verisinin TOPOVT ile eşleşen veri oranının düşük olduğunu belirterek, kırsal kesimdeki OSM verisinin doğru sonuçlar verebilmesi için geliştirilmesi gerektiğini belirtmiştir. Ortalama Hausdorff mesafesi her iki yöntem için de yaklaşık olarak aynı hesaplanmıştır. Buna rağmen, Centroid metodunun eşleşen detaylarda daha az hata verdiğini belirterek, yine de bu yöntemin örtüşme yönteminden daha iyi olduğunu söylemek için daha fazla çalışmaya ihtiyaç duyulduğunu belirtmişlerdir.

Salvucci ve Salvati (2021), OSM veri tabanından elde edilen yerleşim haritalarının ulusal ölçekte kapsamını ve tamlığını doğrulayan basitleştirilmiş bir çerçeve sunmaktadır, bu da resmi istatistiklerde potansiyel kullanım için uygundur. OSM veri tabanından elde edilen binaların tamlığını ölçmek; bina/nüfus sayımları, tarihsel değişimli diğer ölçümler ve bina izinleri ile arazi kullanımı, kadastro kayıtları gibi idari kaynaklarda potansiyel kullanımını destekler. Bu çalışmada, OSM bina verileri ile karşılaştırmak için yetkili veri olarak İtalya Bina Açık Veri Arşivi (Building Open Data Archive-BODA) kullanılmıştır. Karşılaştırma için; sayı tabanlı karşılaştırma, bina ağırlık merkezi ve bina bindirme, poligon kesişimine dayanan nesne tabanlı yaklaşım olmak üzere üç yöntem kullanılmıştır. OSM ve

BODA bina veri setlerinin karşılaştırmalı bir analizi sonucunda kapsama alanına göre bina sayıları belirlenmiştir. OSM binalarının hangilerinin BODA'da bulunduğunu belirlemek için BODA binalarının ağırlık merkezlerinin kaç tanesinin OSM binaları içinde kaldığı hesaplanmıştır. İtalya'nın farklı bölgelerindeki OSM kapsamının karşılaştırmalı analizi ile referans verilerinin belirsizliğine ve heterojenlik kaynaklarını belirlemenin zorluğuna vurgu yapılmıştır.

OSM, bina verilerini elde etmek için önemli bir kaynaktır ancak bu tür verilerde kalite sorunları ile ilgili çalışmaların çok azı küresel ölçekte gerçekleştirilmiştir. Zhou vd. (2022), OSM bina verisinin tamlık kalite kriterini, dünya genelinde 164 ülke, 12.975 şehir için değerlendirmeyi amaçlamaktadır. Referans bina verileri için yetkili veri olarak nüfus grid verileri kullanılmıştır. Sadece her bir şehrin tamlığı değil, aynı zamanda o şehirdeki gridler de değerlendirilmiştir. Sonuçlar, genel doğruluk ve R-Kare hesaplamasına dayalı olarak değerlendirilmiştir ve şehirlerin %75'i için tamlığın %20'den düşük olduğunu göstermiştir; şehirlerin %9'undan fazlası %80'den daha yüksek tamlığa sahip değildir. Çoğu ülkenin genel doğruluk oranı %80'den yüksektir. Tahmini tamlık, önerilen yaklaşımın etkinliğini doğrulayan referans tamlık ile yüksek oranda korelasyon göstermiştir. Bu sonuçlar OSM'de bina verilerinin elde edilmesi ve güncellenmesi için faydalı olabileceği ileri sürülmüştür. OSM bina tamlığı ile ilgili küresel ve açık bir veri seti kullanıcıların kullanımına sunulmuştur.

Borkowska vd. (2023) tarafından OSM bina verilerinin kalitesini değerlendirmek ve kullanıcının veri kümesini seçmesine yardımcı olmak için yöntem geliştirilmiştir ve bu yöntemin herhangi bir OSM poligon detayı için uygulanabilir olduğu belirtilmiştir. Sonuçlar, OSM binalarının tamlığının çoğunlukla yapılaşmış alanlarda neredeyse %100'e yakın olduğunu göstermektedir. Daha az yapılaşmış alanlarda tamlık oranlarının daha düşük olduğu gözlemlenmiştir. Çalışmanın amacı, OSM verilerindeki binaların sayısı ile referans veri tabanına göre ilişkisini sağlamanın yanında, verilerin tamlığı ve kullanılabilirliği açısından kalitesini görselleştirmektir. OSM binalarının tamlığının görselleştirilmesinde üç index kullanılmıştır. Alan tamlığı için C Index, konum tamlığı için TP Index ve sayısal tamlık için Count Index değerleri hesaplanmıştır. Sınıf aralıkları index değerlerinin istatistiksel dağılımına dayanmaktadır. C Index, TP Index ve COUNT Index gibi tamlık indekslerinin görselleştirilmesi için 1 km²'lik altıgen gridler kullanılmıştır. Değerlendirme sonucunda, OSM ve Polonya vektör veri tabanındaki binaların tamlığını görselleştirmek için iki değişkenli renk tonlu haritalar (Bivariate choropleth maps) hazırlanmıştır. Çalışma,

kullanıcıların iki konumsal veri kümesinden hangisinin ihtiyaçlarına daha uygun olduğunu seçerek analiz edilen verilerin kullanılabilirliğini değerlendirmesine olanak tanıyan bir veri kalitesi unsuru olarak OSM verilerinin tamlığının kartografik temsiline odaklanmaktadır. İki değişkenli renk tonlu haritaların sonuçları, kullanıcıya sadece OSM binalarının sayısı hakkında bilgi sağlamakla kalmaz, aynı zamanda OSM verilerinin kalitesini değerlendirmek ve belirli bir amaç için konumsal verilerin uygunluğunu değerlendirmek için bir temel oluşturur.

Detay eşleştirme, iki veri kümesindeki poligonların eşleştirilmesini içeren kalite değerlendirme yönteminde kritik bir adımdır. Araştırmacılar yaygın olarak iki ana poligon eşleştirme algoritması kullanmaktadır: 1) Tampon kesişim yöntemi (Buffer intersection method) ve 2) Ağırlık merkezi karşılaştırma yöntemi (Centroid comparison method). Bu yöntemler OSM bina ayak izlerinin büyük çoğunluğu için etkili olsa da karmaşık durumlarda yüksek doğruluğa ulaşamayabilir. Bunun olası bir nedeni, her iki yöntemin de OSM poligonunun referans poligonuna kıyasla yalnızca konumunu dikkate almasıdır. Bu eşleştirme algoritmalarını iyileştirmek ve daha etkili bir çözüm önermek için Moradi vd. (2023a), iki veri setinde karşılık gelen poligonları daha iyi tanımlamak için konum benzerliğine ek olarak şekil benzerliğini (ortalama mesafe yöntemini kullanarak) dikkate alan bir algoritma önermektedir. Algoritma, OSM binası için referans veri setinde eşleşen birçok bina olması durumunda (1:n) poligonların içbükey kabuklarının (Concave Hull) hesaplanması temeline dayanır. Algoritma OSM-Referans veri seti arasında uygulandıktan sonra Referans veri seti-OSM arasında uygulanır. Quebec eyaletindeki beş şehir için yapılan deney sonuçları, önerilen algoritmanın önceki detay eşleştirme algoritmalarının eşleştirme hatasını %8'den yaklaşık %3'e düşürebileceğini göstermektedir. Ayrıca, önerilen poligon eşleştirme algoritmasının, birden fazla poligondan oluşan binalarda önceki yöntemlere göre daha doğru performans gösterdiğini ortaya koymuştur. Önerilen yöntem, eş binaları bulurken iki veri setindeki öznitelikleri dikkate almamaktadır.

OSM, bina bazında ayrıntılı bir risk değerlendirmesi için büyük bir potansiyel sunsa da OSM bina ayak izlerinin tamlığı hala heterojendir. Ullah vd. (2023), bu boşluğu, gönüllülerin bir bölgenin uydu görüntülerini kaydırarak sınıflandırma görevleri hakkında kullanıcı geri bildirimi topladığı MapSwipe mobil uygulamasına dayanan kitle kaynak kullanımı yoluyla kapatmaya yönelik bir yaklaşım önermiştir. MapSwipe, bir karo (tile) için "bina yok", "tam" veya "eksik" olarak sınıflandırılmasına olanak tanıyacak şekilde "tamlık" özelliği ile genişletilmiştir. Üretilen verilerin kalitesini değerlendirmek için, tamlık özelliği

dört bölgeye uygulanmıştır. MapSwipe tabanlı değerlendirme, mevcut bir modelin tahminleriyle karşılaştırılmıştır. Sonuçlar, kitle kaynak tabanlı yaklaşımın OSM bina çizimlerinin tamlık derecesi için makul bir sınıflandırma performansı sağladığını göstermiştir. MapSwipe tabanlı değerlendirme, çalışma bölgeleri için tutarlı tahminler ürettiğini gösterirken, mevcut yaklaşımın daha yüksek bir değişkenlik gösterdiğini ortaya koymuştur. Ayrıca, gönüllülerin özellikle yüksek OSM bina yoğunluğuna sahip bölgelerde neredeyse tamamen haritalanmış karoları "tamamlanmış" olarak sınıflandırmaya eğilimli olduğunu ve sınıflandırma performansını etkileyen diğer bir faktörün OSM katmanının uydu görüntüleriyle uyum düzeyi olduğunu ortaya koymuştur.

Mevcut çalışmalar incelendiğinde, OSM verilerindeki bina verilerinin veri kalitesini değerlendirmek için referans veri setleri kullanılmıştır. OSM verilerini referans veri setleri ile karşılaştırmak, OSM verilerinin kalitesini değerlendirmek için yaygın kullanılan bir yöntemdir. Bu karşılaştırma, genel olarak değerlendirildiğinde, binaların ağırlık merkezlerinin veya buffer kesişimlerinin karşılaştırıldığı görülmektedir. OSM veri setindeki bina verilerinin referans veri setinde karşılık gelen binaların belirlenmesi için Tablo 2 üzerinde gösterildiği gibi birçok veri kalitesi kriteri değerlendirilmiştir.

Tablo 2. OSM bina verilerini karşılaştıran çalışmalar.

		Referans Veri	Kalite kriteri
Fan vd. (2014)	Münih, Almanya	ATKIS	Tamlık, semantik doğruluk, konum doğruluğu, şekil doğruluğu
Brovelli, Minghini, Molinari, vd. (2016)	Milan, Kuzey İtalya	Milan Belediyesi	Konumsal doğruluk, tamlık
Du vd. (2017)	Büyük Britanya, Fransa	OS, IGN	Konumsal doğruluk
Y. Xu vd. (2017)	Toronto, Kanada	Şehir Planlama ve Belediye Mülkiyet Değerlendirme Kurumu	Tamlık, konumsal doğruluk, şekil doğruluğu, semantik doğruluk yönelme doğruluğu
Brovelli ve Zamboni (2018)	Lombardiya, Kuzey İtalya	Topografik Veri Tabanı	Konumsal doğruluk, tamlık
Basaraner (2020)	İstanbul, Türkiye	1:1,000 ölçekli topografik harita	Konumsal veri kalitesi, semantik veri kalitesi
Küçük ve Anbaroğlu (2020)	Ankara, Türkiye	1:25000 TOPOVT (HGM)	Konumsal doğruluk
Salvucci ve Salvati (2021)	İtalya	Bina Açık Veri Arşivi	Tamlık
Q Zhou vd. (2022)	İngiltere, Fransa, Yeni Zelanda, Avustralya, Birleşik Devletler, Kanada, Uganda, Tanzanya,...(164 ülke)	Nüfus Grid Verisi	Tamlık

Tablo 2'nin devamı

Borkowska vd. (2023)	Piasieczno, Polonya	Polonya Vektör Veri Tabanı-BDOT10k 1:10000	Tamlık
Moradi vd. (2023a)	Quebec, Kanada	Données Québec Açık Veri Merkezi	Konumsal doğruluk
Ullah vd. (2023)	Medellin (Kolombiya) Taipei Tokyo, Japan	MapSwipe bina verisi	Tamlık

Çalışmalarda iki veri setindeki binaların karşılaştırılması sırasında (içinde, kısmen kesişir, tamamen dışında) ortaya çıkan ve belirtilen işlem karmaşıklığı ve yükünü, tez çalışması kapsamında karşılaştırmaların raster veri üzerinde ve bu işlemlerin ontoloji destekli olarak gerçekleştirilmesi önemli ölçüde azaltmaktadır. Literatürde önerilen yöntemlerin çoğu (Moradi vd., 2023) hariç şekil benzerliğini dikkate almamıştır. Tezde önerilen yöntem konum ve şekil benzerliğini dikkate alarak veri kalitesini değerlendirmektedir.

1.2.3. Ontoloji Tabanlı Kalite Değerlendirme

Ontolojiler kalite yönetiminde; kalite değerlendirme, kalitenin raporlanması gibi konularda farklı çalışma alanlarında kullanılmaktadır. Bu kısımda ontolojilerin veri kalitesi yönetimi alanında kullanıldığı çalışmalar ve kullanım detayları kısaca özetlenmiştir.

Mostafavi vd. (2004), kural tabanlı kalite değerlendirmesi için ontolojilerin kullanılmasının avantajlarını araştırmaya yönelik ilk çalışmalardandır. Çalışma, verilerin spesifikasyonuna uygunluğunun test edilmesi amacıyla Kanada Ulusal Topografik Veri Tabanı (NTDB) (Canada, 1997) için uygulanmıştır. Bu spesifikasyon, NTDB'deki detayları, desteklenen geometrileri ve detaylar arasında izin verilen konumsal ilişkileri tanımlar. Dört temel konumsal ilişki içerir; bağlantı, paylaşım, bitişiklik ve üst üste binme. Bağlantı ve paylaşım ilişkileri, ilgili varlıkların aynı temaya veya farklı temalara ait olmasına göre kategorilere ayrılmaktadır. Çalışmada, NTDB için üretilen verilerin NTDB spesifikasyonlarına uygunluğunu test etmek için ontolojiler ve kurallar birlikte kullanmıştır. Ontolojilerden yararlanılmasına rağmen çalışmada kapalı dünya kavramı hakimdir. Kurallar Prolog ile yazılmıştır. Her ne kadar amaçlanan görevi, yani verilerin mantıksal tutarlılık testini yapsa da yeniden kullanılabilir amaçlar için tasarlanmamıştır. Kullanıcı bu şemaya uymayan verileri bulabilir; kuralları özelleştiremez veya benzer değişiklikler yapamaz.

Şemadaki değişiklikler tamamen yeniden kural yazmayı gerektirir. Çalışmanın Prolog kısmında verilerdeki belirli hata kaynaklarına yönelik tutarsızlık kuralları uygulanmaktadır.

Wang vd. (2005), mobil platformlar için sahadan toplanan konumsal verilerin belirli kalite kontrollerinin yapılması ve yeni eklenen veriler nedeniyle herhangi bir tutarsızlık olması durumunda kullanıcıyı uyarması için kullanılan bir sistem geliştirmiştir. Sistem verileri GML olarak sunarken WFS-T servislerini kullanır. Veri ile veri modelinin uyumluluğu kontrol edilerek kalite kontrolü yapılır. GML uygulama şemasında özellikler arasındaki kısıtlamalar SWRL kurallarıyla belirlenmiştir. Bu araştırma, mobil cihazlar kullanarak sahadan veri toplarken ve veri girerken veri doğrulamaya yönelik tasarlanmıştır. Ontoloji tabanlı olan bu uygulamada değerlendirilen kurallar belirli bir uygulama şemasına yöneliktir. Tezde sunulan ontoloji tabanlı kalite değerlendirme uygulamasında alandan bağımsız kurallar ve OSM verisi kullanılmaktadır.

Nash, Wiebenson, vd. (2011), tarım alanında spesifikasyon kurallarının otomasyonunu araştırmaya yönelik bir çalışma yapmıştır. Yapılan çalışmada spesifikasyonlarda yer alan kurallar türlerine göre; “zorunluluk gerektiren” ve “yasaklı” olmak üzere ikiye ayrılmıştır. Nash, Nikkilä, vd. (2011), konumsal kuralların değiştirilebilir kural formatı kullanarak uygulanabilmesi için GeoRIF'i önererek, coğrafi ilişkiler ve geometri destekleyen kurallarla genişletmişlerdir.

Debattista vd. (2014), daQ veri seti kalite ontolojisini geliştirmiştir. Bu ontolojinin amaçlarından biri, mevcut veri setleri arasında karşılaştırmalar yapmak amacıyla veri setleri için kalite sonuçlarını sunmaktır. Sonuç olarak tutarsız verileri içeren bir kalite raporu üretilir.

Fürber (2015), Semantik Veri Kalitesi Yönetim Çerçevesi (Semantic Data Quality Management Framework, SDQM) adı verilen bir sistem önermektedir. Bu çerçeve temel olarak üç bileşenden oluşmaktadır. Kullanıcıların önceden tanımlanmış veri kalitesi boyutları arasından kalite gereksinimlerini tanımlamaları için bir kullanıcı wiki'si oluşturulur. İkinci bileşen, iki temel amaç -kullanıcıların gereksinimlerini temsil etmek ve kalite sonuçlarını veri kalitesi unsurlarıyla ilişkilendirmek- için tasarlanmış Veri Kalitesi Yönetimi (dqm) sözlüğüdür (Fürber ve Hepp, 2011). dqm:DataRequirement sınıfı, tüm veri gereksinimlerinin üst sınıfı olarak tanımlanır. Göreve bağlı uygulamalar durumunda dqm:appliesFor özelliği, belirli gereksinimleri dqm:Task sınıfına bağlar. Ontolojide tanımlanan kurallar, metriklerin değerlerini hesaplamak için ilgili SPARQL kurallarıyla birlikte kullanılır. Bu sistemle uygulanan metrikler özniteliklerle ilgilidir. Metriklerden

bazıları; eksik değerler (missing values), sözdizimi ihlali (syntax violation) (yalnızca harflere izin verilir), aralık dışı değer (out of range value), üst sınır, benzersiz değer ihlali (unique value) dir. Fürber ve Hepp (2010), bu sistemde kuralları SPARQL sorguları ile uygulamaktadır.

Geisler vd. (2016), veri akışlarının (data streams) veri kalitesi yönetimi için ontoloji tabanlı bir yaklaşım önermektedir. Veri akışları için otomatik kalite değerlendirmesinde ontolojileri, konfigürasyon olarak kullanmıştır. Bu, her iki çalışmanın temel ortak noktasıdır. İlgili çalışmada veri akışlarındaki kesintilerden kaynaklanan sorunlara odaklanılmıştır. Kalite kavramlarını ve aralarındaki ilişkileri tanımlamak, kalite değerlendirme sürecini modellemek ve kalite değerlendirme için yeniden kullanılabilir hale getirmek için ontoloji oluşturmuşlardır. Bu sistem için oluşturulan ontolojide kalite metrikleri, kalite unsurları ve fonksiyonlar gibi önerdiğimiz ontolojiyle bazı ortak kavramlar bulunmaktadır. Kalite değerlendirmesine yönelik işlemler, karşılık gelen SPARQL sorgu yapılarıyla birlikte uygulanır (Geisler vd., 2016). Kullanıcıların belli bir alana özel kurallarını tanımlayabilecekleri bir sistem yoktur. Metrikler ve metriklerin işlevleri bellidir, bunun yerine sistem yalnızca kalite değerlendirme için önceden tanımlanmış kuralları uygular. Konumsal veriye yönelik herhangi bir uygulama kuralı önerilmemiştir.

Debattista vd. (2016), ontoloji tabanlı bir veri kalitesi değerlendirme çerçevesi olan Luzzu'yu önermiştir. Luzzu, bağlı verinin kalitesini test etmek için geliştirilmiştir ve kalite değerlendirmek için tasarladıkları alana özgü dil olan Luzzu Kalite Metrik Dili'ni (Luzzu Quality Metric Language-LQML) kullanmaktadır (Debattista vd., 2015). LQML'de 27 metriğe karşılık gelen terimler vardır. Temsili Kısıklık, Birlikte Çalışabilirlik, Yorumlanabilirlik ve Çok Yönlülük, Luzzu'da uygulanan kalite boyutlarıdır. URI'leri kısa tutmak, Mevcut Terimlerin Yeniden Kullanımı, bu boyut için uygulanan metriklerden bazılarıdır. Kullanıcılar bu metrikleri ilgili LQML terimleriyle çağırabilir.

Mobasheri (2017), OSM veri kalitesini artırmak için ontoloji tabanlı bir yöntem önermiştir. Çalışmada örtülü bilgileri veya verilere açıkça dahil edilmeyen fakat mevcut olanlardan çıkarım yapılabilecek ayrıntıları bulmak için konumsal çıkarsama kullanılmıştır. Önerilen yaklaşım konumsal çıkarsama için bir dizi kural kullanan bilgi aracısı (information broker) adı verilen bir sistemi içermektedir. Uygulamada, yaya geçitleri ve bina girişleri ile ilgili eksik bilgileri arayarak, yöntemlerini test etmişlerdir. Geliştirilen sistem, yollar ve binalarla ilgili mevcut verileri analiz ederek “yaya geçidi” ve “bina girişi” verilerinin olası konumlarını çıkartmıştır ve bu veriler daha sonra yönlendirme ve navigasyon amacıyla

bütünlüğü artırmak için OSM'ye eklenmiştir. Detay çıkartma aşamasında sistemde ontoloji ve SQWRL kuralları kullanılmıştır.

X. Xu ve Cai (2020), yeraltı tesisatlarının güvenlik yönetmeliklerine uygunluğunu test etmek için ontoloji tabanlı bir yöntem önermiştir. Önerilen yöntemde yeraltı tesisat verileri (konumsal veri ve malzeme verisi vb.) ve yönetmelikteki kurallar ve kavramlar ontoloji kullanılarak tanımlanmıştır. Birçok yaklaşımda olduğu gibi olası yönetmelik ihlali belirlemek için ontolojinin içerisinde alana yönelik kurallar üretilmiştir. Çalışmada üretilen yazılım bu kurallara karşı yeraltı tesisat verisini otomatik olarak test ederek uyumsuz verileri belirlemiştir. Bu çalışmada ontoloji tabanlı yöntemin hedefi kurumsal kaynaklı veri olup tezde OSM verisini test etmeye yönelik bir sistem önerilmiştir.

Homburg ve Boochs (2019), konumsal verilerin kalitesini değerlendirmek için verilerin kullanılacağı bağlamı vurgulayarak geleneksel yaklaşımlardan farklılaşan bir yöntem önermiştir. Çalışmada, verinin kullanıldığı bağlamla; tematik haritalar kastedilmektedir. Kalite değerlendirmesi yapılacak verilerin kullanım amaçlarına yani temalarına göre test edilmesi gereken kalite elemanlarının farklılık gösterdiği gerekçesiyle bu gereksinim için bir yöntem geliştirmeyi hedeflemiştir. Çalışmada SWT, verilerin kullanım senaryolarını tanımlamak ve bu senaryoları özelleştirilmiş veri kalitesi modelleriyle ilişkilendirmek için kullanılmıştır. Önerdikleri yaklaşımı, OSM verisiyle “afet planlaması” temasına uygulayarak örneklemiştir.

Homburg vd. (2019), BKG'nin verilerini kullanarak açık bağlı veriyi güncellemeye yönelik bir veri kalite test çerçevesi önermişlerdir. Bu açıdan bu tezde önerilen çalışmaya benzerlik göstermektedir. Yöntemleri Homburg ve Boochs (2019) yayınında bahsedilen yönetime dayanır.

Homburg (2020), tematik haritaların amaçlanan kullanımlarına uygun olmasını sağlamak amacıyla ontoloji tabanlı otomatik kalite değerlendirme sistemi için bir çerçeve önermiştir. Çalışmanın ana fikri, ilgili durumları veri kalitesi ölçümleriyle (doğruluk, tamlik gibi) ilişkilendirmektir. Böylece, temeldeki verilerin belirli bir durumun ihtiyaçlarını karşılayıp karşılamadığına dayalı olarak tematik haritaların değerlendirilmesine olanak tanımaktadır. Önerilen çerçevede, alana ait kavramlar arasındaki ilişkileri kurmak için bu tez kapsamında önerilen sistemde de olduğu gibi ontolojiler kullanılmıştır. Bu yaklaşım, bir haritanın belirli bir tematik durumun ihtiyaçlarıyla ne kadar uyumlu olduğu konusunda çıkarsama yapma olanağı sunar. Bahsedilen tematik haritalar için kaynak veri olarak GCB'den elde edilen konumsal veri kullanılmıştır. OSM verisi kullanılarak elde edilen

verinin öznitelik açısından zenginleştirilmesi amacıyla Vikiveri kullanılmıştır (Vikiveri, 2022). Çalışmada, kullanıcı geri bildirimlerinin çerçeveye nasıl entegre edilebileceği araştırılarak, kullanıcı katkısı ile çerçevenin doğruluğunun ve kapsamının artırılması amaçlanmıştır. Çalışma veri kalitesinin değerlendirilmesi için kullandığı teknik temelinde ontoloji tabanlı bir yaklaşım sunmasıyla bu tez kapsamında önerilen sistemle benzerlik göstermektedir. Çalışmalardaki temel farklılık odaklandıkları veri kalitesi elemanları ve önerdikleri sistemlerin hizmet ettiği kullanıcılar olmuştur. Bahsedilen çalışma, OSM verisini tematik doğruluk açısından otomatik olarak test etmeyi odağında tutarken, tez çalışmasında diğer kalite elemanları (özellikle konumsal doğruluk) dikkate alınarak, her iki haritanın birbirini ne kadar güncelleyebileceği araştırılmıştır.

Qiu vd. (2021), otonom araçların güvenli bir şekilde otomatik sürüş yapabilmesi için zengin semantik veriye ihtiyacından yola çıkarak harita veri kalitesini artırmak ve test etmek için ontoloji tabanlı bir yöntem önermiştir. Önerilen yöntemde; harita verileri, özel olarak yüksek çözünürlüklü haritalar için tasarlanmış bir ontoloji ve bir dizi kural kullanılarak ek bilgilerle zenginleştirilir. Bu işlem, hataları belirlemek ve düzeltmek için gerekli konumsal bilgiyi sağlar. Zenginleştirilmiş verideki tutarsızlıkları bulmak için bir “harita kalite ihlalleri ontolojisi” kullanılmıştır. Ontolojide, ihlalleri algılamak için kurallar üretilmiştir. Yüksek önem taşımayan hataları otomatik olarak düzeltmek için stratejiler önermektedir. Bu sayede, otonom sürüş modunu kapatmadan otonom arabanın güvenli bir şekilde çalışmaya devam etmesi sağlanmış olur.

1.2.4. OSM Veri Kalitesi

İlgili literatürde OSM verilerinin en çok ele alınan yönü veri kalitesidir. OSM verilerinin kullanımı için atıf verilmek şartıyla kullanılabilir olması, kapsamı, veri içeriği gibi özellikleri açısından kitle kaynak verilerin kalitesinin değerlendirilmesi için en çok kullanılan veri kaynağıdır. Bu bağlamda OSM veri kalitesinin değerlendirilmesinde yapılan çalışmalar incelendiğinde üç grupta toplanabilir. Birincisi, OSM verilerinin yetkili kurum verileri ile karşılaştırıldığı ve yetkili kurum verilerinin referans veri seti olarak doğru kabul edildiği çalışmaları içerir. Önemli örnekler arasında M. Haklay (2010), Girres ve Touya (2010) ve Neis vd. (2011) sayılabilir. Arsanjani vd. (2015) gibi ikinci grup çalışmalar OSM kalitesine odaklanırken, üçüncü grup çalışmalar OSM verisinin geçmişine odaklanır ve bu

geçmiş veri ile birlikte OSM veri tabanında saklanır. Üçüncü gruba örnek olarak Keßler ve De Groot (2013) ve Keßler vd. (2011) verilebilir.

OSM Konumsal Doğruluk

OSM verisi üretilmeye başlandığı zamandan günümüze en çok çözüm üretilmeye ve değerlendirilmeye çalışılan kalite kriterlerinden biri “konumsal doğruluk” olmuştur. Konumsal doğruluğun değerlendirilmesi için literatürde iki temel yöntem mevcuttur. Birincisi doğru kabul edilen yetkili veriler referans veri seti olarak kullanılarak, OSM verisinin referans veri setiyle kıyaslanmasıdır (Helbich vd., 2012; Antunes vd., 2015a; El-Ashmawy, 2016; Brovelli ve Zamboni, 2018; H. Zhang ve Malczewski, 2019; W.-B. Zhang vd., 2019).

İkinci yöntemde ise OSM verisinin konumsal doğruluğunun içsel parametreler kullanılarak hesaplanmasıdır. Bu yöntem genellikle verinin ait olduğu bölgeye yönelik referans veri setinin olmadığı durumlarda OSM verisinin konumsal olarak kullanılabilirliği hakkında karar verme sürecinde kullanılmaktadır (M. (Muki) Haklay vd., 2010; Van Exel vd., 2010; Arsanjani vd., 2013; Barron vd., 2014; Jacobs, 2018; Paiva ve Camboim, 2022).

1.3. Temel Tanımlar Ve Kavramlar

1.3.1. Gönüllü Coğrafi Bilgi

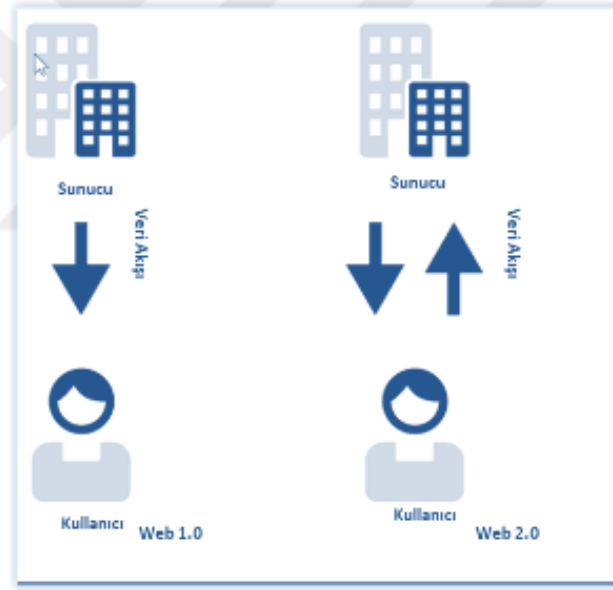
Gould (1999), konumsal bir yüzyılın gelişini öngörmüş ve “çoğu insanın içinde bir coğrafya uzmanı bulunduğunu” ileri sürmüştür. Son yirmi yıldaki çok sayıda yeni teknolojik gelişme gerçekten de herkesin içindeki coğrafyacının potansiyelini ortaya çıkarmıştır. Yirmi birinci yüzyılın ilk on yılında ortaya çıkan GCB olgusu, Gould (1999)’un kaleme aldığı konumsal yüzyılın birçok tezahüründen biridir (D. Sui vd., 2013).

GCB, konumsal bilgi içeriği ve yapısında, konumsal bilgiyi kullanma, paylaşma ve konumsal bilgi üretme yöntemlerinde büyük bir değişimi temsil etmektedir (D. Sui vd., 2013). GCB, coğrafi uzmanların yanı sıra, uzman olmayan kişilerin de katkıda bulunduğu geleneksel KVA kavramına alternatif olması açısından “yeni” olarak nitelendirilebilecek konumsal veri kaynağıdır. Bahsedilen yeniliğe sebep olan değişim, temel olarak Web 1.0 teknolojisinin Web 2.0 teknolojisine evrilmesidir.

Web 2.0, kullanıcı katılımı, herkesin erişimine açık olması ve kullanıcı sayısının artması sonucu ortaya çıkan ağ etkisi ile karakterize edilen bir takım ekonomik, toplumsal

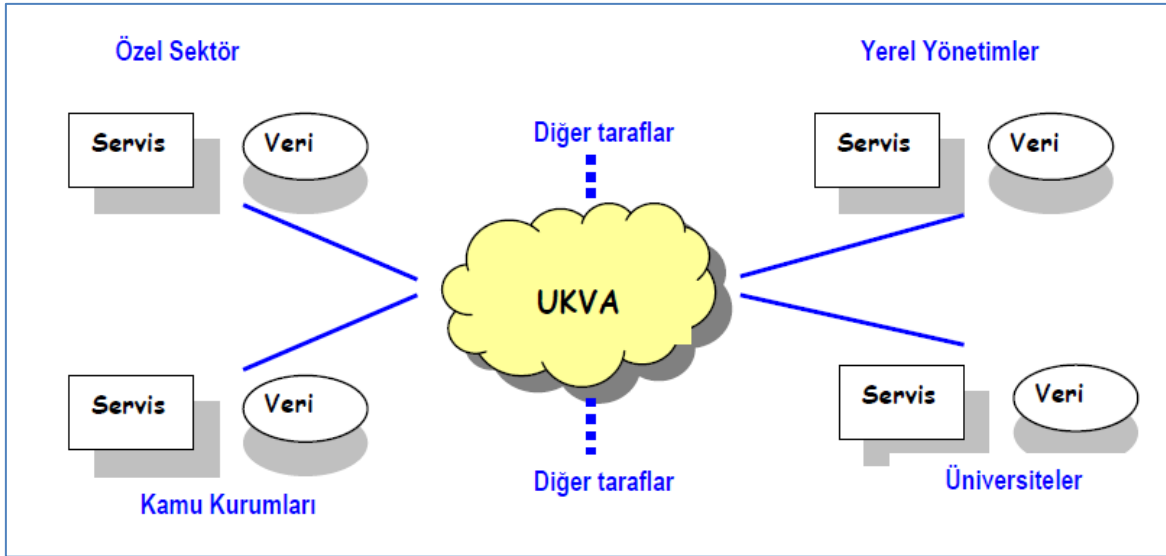
ve teknolojik yönelimin bir araya gelerek oluşturduğu yeni nesil internettir. Web 1.0'dan Web 2.0'a geçişin sonucu olarak artık kullanıcılar da Web'de veri paylaşabilmektedir. Bu yeni eğilimi farklı topluluklar tarafından tanımlamak için kitle kaynak kullanımından “kullanıcı tarafından oluşturulan içeriğe”, Geoweb'den Semantik Web'e, GCB'den Neocoğrafya'ya, PostGIS'e, Vatandaş Bilimi'ne ve e-Bilime kadar farklı kavramlar ortaya çıkmış olsa da, genel fikir şu şekilde birleşmektedir: Birden fazla bilgi işlem cihazı/platformu (geleneksel masaüstü bilgisayarlar, iPad'ler veya akıllı telefonlar) aracılığıyla konumsal bilgileri oluşturmak, paylaşmak ve analiz etmek için internetin kullanılması anlamına gelmektedir (Turner, 2006; D. Sui vd., 2013).

Şekil 1 ile gösterildiği gibi, GCB, Web 2.0'ın kullanıcıların da veri paylaşımına olanak veren altyapısı nedeniyle, konumsal alanda uzman olmayan kişiler de konum verisi paylaşarak veri üretimi kısmında yer almaya başlamıştır.



Şekil 1. Web 1.0 ve Web 2.0'da veri akışı.

Bu değişimden KVA dahil birçok alan etkilenmiştir. Geleneksel KVA'da, veri sadece yetkili kurumlar tarafından üretilip son kullanıcılarla paylaşılmaktadır. Şekil 2 üzerinde gösterildiği gibi, Ulusal KVA'nın gelişimi ile, modern çağın SWT gibi getirileri kayda geçirilmiş, güncellenebilirlik ve birlikte işlerlik gibi kavramlar dahil edilmiştir (Çetin Cömert ve Akinci, 2004; Ç Cömert vd., 2008).



Şekil 2. Ulusal KVA'nın kapsamı ve yapısı (Akıncı ve Cömert & Akinci, 2004).

GCB kavramı CBS alanında ilk olarak OSM projesiyle gündeme gelmiştir. Günümüzde gönüllü kullanıcıların web üzerinden konumsal etiketlerle paylaşımlar yaptığı OSM ve Wikimapia sitelerindeki çevrimiçi haritalar, Flickr gibi web sitelerinde konumla ilişkili paylaşılan fotoğraflar, eBird projesinde kuş gözlemi için paylaşılan gerçek zamanlı veriler GCB'ye örnek verilebilir. Kullanıcıların gönüllü katkı yapabildiği konumsal veri odaklı siteler arasında OSM, Google Maps, Wikimapia, LinkedGeoData, Google Maps ve GeoNames yer alır (Vyron Antoniou, 2017).

Google Maps, ilgi çekici yerler ve yollar gibi özelliklerin haritalandırılmasına olanak tanır. Wikimapia, dünyadaki yerleri tanımlamak için araçlar sağlar ve bir coğrafya sözlüğü (gazetteer) hizmeti olarak kabul edilebilir (Goodchild, 2011). Wikimapia, gönüllü kullanıcılara fotoğraf yükleme ve kategorize etme, açıklama ekleme, yerleri işaretleme olanağı verir (Wikimapia, 2012).

LinkedGeoData, OSM verilerini kullanarak SWT'ye konumsal boyut ekler (Stadler vd., 2012; LinkedGeoData, 2023). GeoNames, kullanıcıların sitesine yer adlarını ekleyebileceği veya düzenleyebileceği başka bir coğrafya sözlüğü hizmetidir (Vyron Antoniou, 2017).

1.3.2. Kitle Kaynak Kullanımı

Kitle kaynak sosyal medya, harita, biyoçeşitlilik, sağlık gibi çok çeşitli alanlarda kullanılmaya başlanan bir kavramdır. Geleneksel çalışanlardan veya tedarikçilerden farklı

olarak genellikle çevrimiçi olmak üzere geniş bir kullanıcı grubundan katkı isteyerek ihtiyaç duyulan hizmetleri, fikirleri veya içeriği elde etmeye yönelik bir uygulamadır. Estellés-Arolas ve González-Ladrón-de-Guevara (2012), kitle kaynak kavramının farklı alanlardaki tanımlamalarını inceleyerek, ortak bir kavram tanımı önermiştir. Bu tanıma göre kitle kaynak kullanımı, “bir bireyin/kurumun, kâr amacı gütmeyen bir kuruluşun veya şirketin, farklı bilgi, heterojenlik ve sayıya sahip bir grup kişiye, esnek bir açık çağrı aracılığıyla gönüllü olarak bir projeyi üstlenmeyi teklif ettiği bir tür katılımcılı çevrimiçi etkinliktir”. Değişken karmaşıklık ve modülerlikteki ve kullanıcıların iş, para, bilgi ve/veya deneyimlerini getirerek katılması gereken görevin üstlenilmesi her zaman karşılıklı faydayı gerektirir. Sistemi kullanıcılara sunan taraf, kullanıcının kullanım türüne göre sisteme aktardığı her ne ise yalnızca ona erişebilir ve ondan yararlanabilirken; kullanıcı da gerek ekonomik gerekse takdir görme, kendine güven duyma veya kişisel becerilerinin gelişimi olsun bir ihtiyacın giderilmesinden doğan memnuniyeti elde edecektir (Estellés-Arolas ve González-Ladrón-de-Guevara, 2012).

Papadopoulou ve Giaoutzi (2014), kitle kaynak kullanımı sürecini bir dizi adımla tanımlamıştır. Bu adımlar; bir sorunun çevrimiçi olarak yayımlanması, kalabalık katılımcılar tarafından alternatif çözümler üretilmesi, önerilen çözümlerin değerlendirilmesi, sağlanan en iyi çözümün seçilmesi ve seçilen çözümün, sorunu çevrimiçi olarak ilk yayımlayan şirket veya kurum tarafından kullanılmasıdır.

Kitle kaynak kullanımı ile geliştirilen en temsili uygulamalar arasında böyle bilgilerin destekçileri olarak OSM, Wikipedia, Waze, Arcbazar ve Facebook yer almaktadır (Harvard, 2015). Wikipedia, ortaklaşa bir ansiklopedi oluşturmak için Wiki yazılımını kullanan uluslararası bir projedir. Maddeleri her kullanıcı (vandalizm nedeniyle engellenmemiş) tarafından düzenlenebilir ve düzenlemeler kaydedilir. Tüm maddelerin sürüm geçmişi herkese açıktır (Wikimedia, 2024).

Waze, mobil cihazlarının konumuna bağlı olarak kullanıcılar ve sürüş rotaları hakkında gerçek zamanlı bilgi sağlayan uydu tabanlı bir navigasyon uygulamasıdır (Waze, 2024). Waze uygulamasında, bir arkadaş çevresi ve mesaj alışverişi platformu olarak işleyen sosyal ağı ve sürüş rotaları ve gerçek zamanlı trafik güncellemeleri sunmak için kullanıcılarının sürüş izlerini öğrenir; bu da diğer GPS tabanlı uygulamalara göre avantaj sağlar (Laor ve Galily, 2022).

Arcbazar, mimari tasarım hizmetleri için bir kitle kaynak platformudur (Arcbazar, 2023).

1.3.3. OpenStreetMap

OSM, güncellenen dünyanın konumsal verilerini gönüllülerin katkısıyla ücretsiz olarak üretmek amacıyla 2004 yılında başlatılan bir projedir. Dünya çapındaki gönüllüleri konumsal verilerin toplanması yoluyla katkıda bulunmaya teşvik eden, CBS'nin vikileştirilmesi olarak da adlandırılan kitle kaynak kullanımı kavramına dayanmaktadır (Sui, 2008). OSM, kullanıcılara haritayı düzenlemek, özellik eklemek ve kaldırmak için araçlar sağlar (Haklay ve Weber, 2008). Proje başladığı zamandan günümüze veri tabanında depolanan bilgi miktarı bakımından sürekli olarak büyümektedir. OSM'nin kurulmasının ve büyümesinin ardındaki iki ana itici güç, dünyanın büyük bir kısmında harita bilgilerinin kullanımı veya kullanılabilirliği üzerindeki kısıtlamalar ve ucuz taşınabilir uydu navigasyon cihazlarının geliştirilmesidir (Sehra vd., 2014).

OSM verisi birçok fayda sağlamaktadır. Öncelikle, veriler açık içerik lisansı ile tamamen ücretsizdir. Kendileri için önemli olan ilgi çekici noktaları da ekleyebilen kayıtlı kullanıcılar tarafından sürekli olarak güncellenmektedir. Ayrıca, dünyanın her yerinden gönüllü oluşturma potansiyeline sahiptir; böylece çoğu ticari haritalama şirketi için veri elde etmenin zor olabileceği bölgelerde de gönüllülerin katkısıyla konumsal veri elde edilebilmektedir (Ather, 2009).

OSM'nin açık ve kitle kaynaklı yapısı ile dünya çapında gönüllüler, GPS cihazları, uydu görüntüleri ve diğer kaynaklardan elde ettikleri verileri kullanarak OSM haritasına katkıda bulunur. Bu, OSM'nin sürekli güncellenmesini ve genişlemesini sağlar, ancak aynı zamanda veri kalitesi ve tutarlılığı konusunda zorlukları da beraberinde getirir; veri kalitesi ve güvenilirliği değişkenlik gösterebilir. Kullanıcılar ayrıca, haritadaki hataları düzeltmek ve veri tutarlılığını sağlamak için birbirleriyle iş birliği yapar.

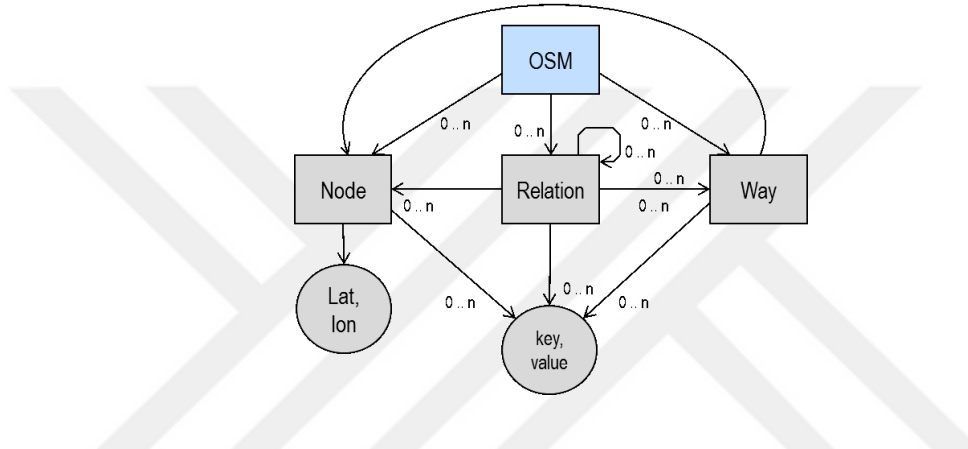
OSM verilerinin katkı kaynağındaki çeşitlilik, kullanımındaki çeşitliliğe de yansımıştır. OSM verileri, navigasyon sistemlerinden kentsel planlamaya, afet yönetiminden turizme kadar geniş bir yelpazede değerlendirilmektedir. OSM'nin esnek yapısı, bu çeşitliliğin ana unsurlarındandır; özelleştirilmiş haritaların ve uygulamaların geliştirilmesini kolaylaştırır.

OSM'nin temel veri yapısı üç ana bileşenden oluşur: düğümler, çizgiler ve ilişkiler. Kavramsal veri modeli, Şekil 3 ile gösterilmiştir. Konum bilgileri düğümlerde tutulur. Dolayısıyla, çizgilerin köşeleri olarak kullanılan düğümler de bağımsız birer OSM

elemanlarıdır. Örneğin ESRI Shapefile dosyalarında bu geçerli değildir; poligon ve çizgilerin köşe noktaları ayrı olarak değerlendirilmez.

OSM verileri XML tabanlı olan *.osm formatında saklanır, "anahtar:değer" çiftleri şeklinde etiketlenir. Bu etiketler, harita üzerindeki herhangi bir öğenin özelliklerini ve niteliklerini tanımlar. Örneğin, bir yola "highway=secondary" etiketini eklemek, bu yolun ikincil bir karayolu olduğunu belirtir.

*.osm formatının *.osm ve *.pbf gibi daha sıkıştırılmış alternatifleri mevcuttur. MATLAB yazılımı 2023 versiyonları ile bu formatları kullanabilir hâle gelmiştir.

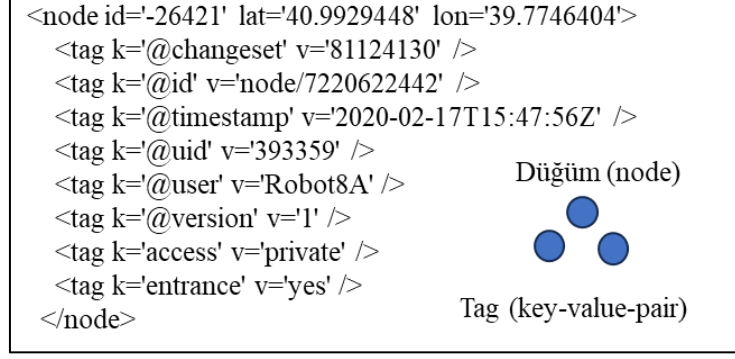


Şekil 3. Basitleştirilmiş OSM kavramsal veri modeli (OpenGeoEdu, 2024).

Düğümmler

OSM düğümleri en temel veri öğeleridir ve belirli bir enlem ve boylam koordinatına sahip noktaları temsil eder. Bu düğümler, tek başlarına (örneğin, bir şehirdeki su kuyusu gibi) ya da daha karmaşık yapıların parçası olarak (örneğin, bir yolun ucundaki nokta) kullanılabilir.

Şekil 4 örneğinde görüldüğü gibi konum (enlem, boylam, lat, lon) yanında anahtar-değer (key-value) çiftleri halinde çeşitli etiketler (tag) bulunmaktadır. @ ile başlayan anahtarlar temel anahtarlardır. OSM düğüm sayısı 2022 Kasım ayı itibarıyla 8 milyarı geçmiştir (OSM, 2022). Bu düğümlerin sadece dörtte birinin etiketi vardır.



Şekil 4. OSM XML örneği, düğüm (nokta, node).

Çizgiler

OSM çizgileri düğümlerden oluşur ve onları bağlar. Yollar, demiryolları, nehirler gibi çeşitli çizgisel özellikleri temsil eder. Düğümler “nd” XML ögesi ile belirtilir, dolayısıyla ayrıca konum bilgisi vermeye gerek kalmaz. Çizgiler aynı zamanda poligonlar olarak, örneğin bir parkın sınırlarını çizmek için kullanılabilir; bu durumda ilk düğüm ve son düğüm Şekil 5 örneğinde görüldüğü gibi birbirine eşittir. Bu örnek HGM verisinden uyarlanmıştır ve HGM öznelikleri OSM anahtarlarına dönüştürülmüştür. OSM çizgi sayısı 2022 Kasım ayı itibarıyla 900 milyonu geçmiştir (OSM, 2022). Düğüm sayısının % 11’i kadar çizgi vardır.



Şekil 5. OSM XML örneği, çizgi (way).

İlişkiler

OSM ilişkileri, sadece düğümler, çizgiler ve hatta ilişkiler arasındaki ilişkileri tanımlar. Örneğin, bir otobüs güzergâhı ilişkisi, güzergâhtaki durakları (düğümler) ve

güzergâhın kendisini (çizgiler) içerebilir. Şekil 6 örneğinde bir kısıtlama (sağa dönüş mecburiyeti) ilişkisi görülmektedir.

```
<relation id=«373395» visible=«true» timestamp=«2010-01-13T19:35:29Z»
Version=«5» changeset=«3612328» user=«JosmJo»uid=«118197»>
  <member type=«way» ref=«48242678» role=«to»/>
  <member type=«node» ref=«27270924» role=«via»/>
  <member type=«way» ref=«47229081» role=«from»/>
  <tag k=«restriction» v=«only_right_turn»/>
  <tag k=«type» v=«restriction»/>
</relation>
```

İlişki (Relation)

Şekil 6. OSM XML örneği, ilişki (relation)

OSM ilişki sayısı 2022 Kasım ayı itibarıyla 10 milyonu geçmiştir (OSM, 2022).

Devlet Kurumları ve Özel Sektörde OSM Verisinin Kullanımı

OSM, hem konumsal veri ile iş yapan özel sektör firmaları tarafından hem de devlet kurumları tarafından kullanılmaktadır. OSM verilerini kullanarak hizmet veren birçok özel sektör firması bulunmaktadır. Bu firmalar genellikle OSM verisini işleyerek yine OSM projesine veya OSM kullanıcılarına hizmet sunmaktadır. Web haritalama, navigasyon ve veri çıkartma gibi işlemler firmaların sağladıkları hizmetlere örnek verilebilir. MapBox, MapQuest, GeoFabrik, Mapzen; OSM Vikipedi sayfasında belirtilen ve OSM ile ilgili hizmet veren firmalardandır (OSM Wiki, 2018).

OSM'nin yetkili kurum tarafından kullanımı, ancak ilgili kurumun bilgilendirmesiyle anlaşılır. Diğer yönden bakıldığında OSM'de yetkili verilerin yıllara yayılan süreçte önemli ölçüde kullanıldığı görülmektedir. Dünya çapında farklı ülkelerden farklı devlet kurumları ürettikleri veriyi OSM'nin kullanımına açarak OSM verilerinin üretilmesine ve güncellenmesine katkıda bulunmuştur. ABD'de Topolojik Olarak Entegre Coğrafi Kodlama ve Referanslama Sistemi (Topologically Integrated Geographic Encoding and Referencing, TIGER) verileri, Kanada'nın tamamını kapsayan bir harita olarak GeoBase verisi, Fransa'daki Arazi Örtüsü haritası verilerini içeren CORINE verileri farklı kurumların OSM'ye katkılarına örnektir.

2013 yılında New York City Belediyesi, bina ve adres noktalarını içeren “yüksek değerli” veri setlerinin OSM'yi geliştirmek için kullanılmasına izin vererek, MapBox yardımıyla OSM'ye yüklemiştir (Barth, 2013; OSM Wiki, 2024). Bunun karşılığında New York City'nin CBS grubu, OSM'de kendi veri setleriyle ilgili yapılan değişikliklerden

haberdar edilerek harita verilerinin güncel kalmasını sağlamak yönünde bir adım atmıştır. Bu durum, New York Belediyesi'ni ABD'de OSM'nin hem katılımcısı hem de katkı sağlayıcısı yapmıştır.

MapGive, küresel haritalama topluluğuna katılımı teşvik ederek artıran ve insani yardım ve kalkınma programlarını desteklemek için açık konumsal verilerin oluşturulmasını kolaylaştıran bir ABD Dışişleri Bakanlığı girişimidir (MapGive, 2014). Portland'ın Trafik Birimi çok-modlu trafik planlayıcılarını güçlendirmek için OSM'yi kullanmaktadır (TriMet, 2009).

Fransa Ulusal Adres Veri Tabanı (Base Adresse Nationale, BAN), Fransa'da listelenen her adresi (25 milyon adres) coğrafi koordinatlarıyla ilişkilendirmektedir (Chausson, 2017). BAN, ekonomi, toplum ve kamu hizmetleri için temel bir referans oluşturmak üzere Fransa'daki kamu yetkilileri ile OSM Fransa arasında yenilikçi bir iş birliği modelinin sonucudur.

1.3.4. Yetkili Konumsal Veri Kaynakları ve Özellikleri

Yetkili veri kaynakları, devlet kurumları ve resmi harita kurumları tarafından sağlanan konumsal bilgiler ve haritalar içeren veri setleridir.

Yetkili veri kaynakları, uluslararası ve ulusal standartlara uygun olarak oluşturulur ve sürdürülür. Avrupa INSPIRE Direktifi ile 2007 yılında oluşturulan INSPIRE temaları uluslararası standartlara örnektir (INSPIRE, 2014). Türkiye'de Bakanlıklararası Harita İşlerini Koordinasyon ve Planlama Kurulu'nun gündeme aldığı Büyük Ölçekli Harita ve Harita Bilgileri Üretim Yönetmeliği, Harita Genel Müdürlüğü'nün Topografik Veri Tabanı (TOPOVT) Veri Sözlüğü, ulusal standartlara örnektir (Resmi Gazete, 2005; HGM, 2019).

Üretilen verinin kalitesi, konumsal doğruluk, mantıksal tutarlılık ve güncellik gibi kriterler açısından değerlendirilir. Örneğin, HGM tarafından üretilen TOPOVT verilerinin kalitesinin test edilmesinde, yine kurum tarafından üretilen "kıymetlendirme kurallarına" "TARLABAHCE içinde AGAC detayı olamaz", "MEZARLIK içinde MEZAR detayı olamaz" gibi kurallara göre verilerin uygunluğu değerlendirilmektedir. Bu süreç, verilerin kullanımının çeşitli uygulama alanları için uygun ve güvenilir olmasını sağlamaya yardımcı olur.

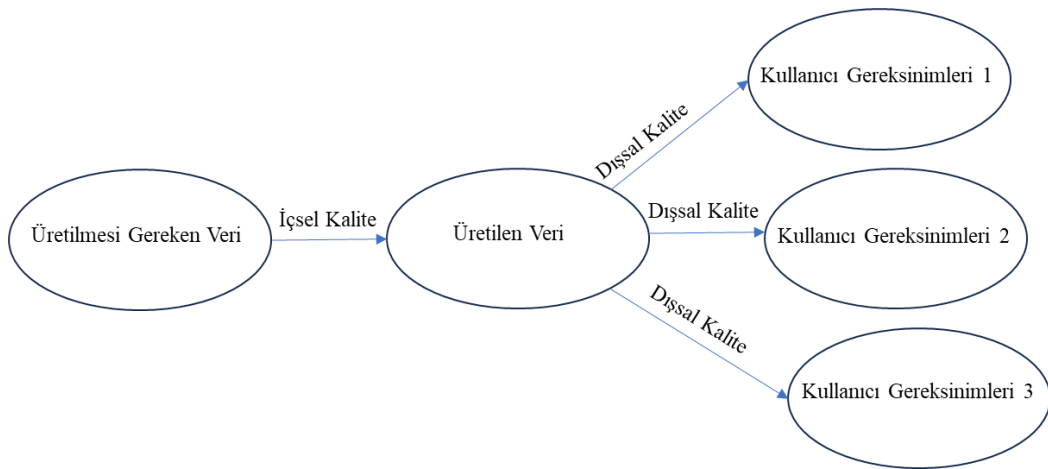
Yetkili konumsal veri kaynakları, kentsel ve bölgesel planlama, altyapı yönetimi, çevresel analizler ve acil durum hizmetleri gibi birçok alanda kullanılır. Bu veriler, resmi ve

hükûmet düzeyinde karar verme süreçlerinde temel bir rol oynar ve genellikle politika yapımı ve kamu hizmetlerinin planlanması için kullanılır. Yetkili veriler, kurumlar tarafından alanında uzman kişiler görevlendirilerek, sıkı bir veri üretme ve kontrol yönetmeliği sonucunda oluşturuldukları için hassas ve güvenilir veri kaynakları olarak kabul edilir (Külcü, 2020).

1.3.5. Konumsal Veri Kalitesi

Konumsal veri kalitesi, konumsal verilerin kullanım değerini belirleyen önemli bir unsurdur. Veri kalitesi literatürde, bakış açısına bağlı olarak konumsal veri üreticileri ve konumsal veri kullanıcıları açısından ele alınarak temelde iki farklı şekilde tanımlanmaktadır (Herzog vd., 2007; R. Devillers vd., 2007; ISO, 2023). Kullanıcı açısından ele alınan kalite kavramı dışsal (extrinsic) veri kalitesi olarak, üretici tarafını ilgilendiren kalite kavramı ise içsel (intrinsic) veri kalitesi olarak adlandırılmaktadır. Juran (1967)'de dışsal (extrinsic) veri kalitesi, “kullanım amacına uygunluk” olarak tanımlanmaktadır. Şekil 7 ile gösterildiği gibi bunu da kullanıcı gereksinimleri belirler. İçsel veri kalitesi ise üretilen veri ve ilgili spesifikasyonlar arasındaki tutarlılıktır (Wand ve Wang, 1996; Batini ve Scannapieca, 2006).

Bu tez çalışmasının konusu kapsamında içsel kaliteye ve literatürde onun bir alt elemanı olarak kabul edilen kavramlara, özellikle “konumsal doğruluk” kavramına odaklanılmıştır.



Şekil 7. İçsel ve dışsal veri kalitesi (R. Devillers vd., 2007).

Veri kalitesi elemanları etki alanına yani veri türüne göre ele alınmaktadır. GCB, KVA ve Bağlantılı Veriler'in kalite değerlendirme sürecinde ortak kalite elemanları olduğu gibi, verinin yapısından kaynaklanan farklı veri kalite elemanları da bulunmaktadır (Vinet ve Zhedanov, 2011; Zaveri vd., 2015; Fonte vd., 2017; ISO, 2023). Uluslararası Standardizasyon Örgütü (International Standardization Organization, ISO), konumsal veri kalitesi için ilk olarak 2001 yılında ISO:19113 ISO (2001) standardı ile başlayarak farklı birçok kalite standardı yayımlamıştır, son yayımlanan standart, ISO 19157:2023'tür (ISO, 2023). Çeşitli kurumlar tarafından yayınlanan kalite standartları veri üreticilerinin verileri hassas bir ölçüm metoduna göre elde ettiği varsayılarak hazırlanmıştır. Hassas GPS ölçüsü, fotogrametrik ölçüm veya yersel ölçümler sonucunda elde edilen konumsal verinin değerlendirilmesinde izlenecek yöntem ve hata kaynakları veri üretme yöntemine göre değişiklik gösterir, fakat bellidir. GCB'nin doğası gereği, kurumlarca üretilen veriden daha farklı hata kaynakları barındırmaktadır, bu durum GCB için yeni kalite kriterlerini araştırmaya sebep olmuştur (Goodchild ve Li, 2012a; Jackson, 2014; Antoniou ve Skopeliti, 2015; Bordogna vd., 2016; Senaratne vd., 2017; Antoniou, 2017). Antoniou (2017), GCB kalitesini değerlendirmek için ISO 19157:2013 standardının yeterliliğini tartışarak, GCB'ye özgü ek olarak, güvenilirlik, sosyo-ekonomik kalite göstergeleri ve katılımcı göstergeleri önermiştir.

Tarih boyunca konumsal veri üreticileri tarafından en yüksek kalitede, daha doğru, daha tutarlı, daha az eksikli haritalar ve konumsal veriler üretilmeye çalışılmaktadır. Modern çağda, 1980'lerin sonlarında ve 1990'larda, komiteler konumsal verilerin kalite unsurlarını tanımlamak için bir araya gelmiştir. USGS Veri Standartları Komitesi 1983 yılında kalite ile ilgili standartlar geliştirmiştir (McEwen vd., 1983). Kartografik veri setlerinin veri kalitesinin belirtilmesi gerekliliği ve ortak kalite açıklamasına yönelik standartlar geliştirilmiş olan standartta belirtilmektedir. 1987 yılında "Sayısal Kartografik Veriler için Önerilen Standart Taslağı" yayımlanmıştır ve bu standart veri kalitesi unsurlarını tanımlamaktadır (Moellering, 1987).

Uluslararası Kartografya Birliği (International Cartography Agency, ICA), veri kalitesi kriterlerini, kalite değerlendirme yöntemlerini geliştirmek için birkaç kez toplanmıştır ve ICA Konumsal Veri Kalitesi Komisyonu'nu kurmuştur (Raines, 1997). Bu öneriler ve ISO 9000 Kalite Yönetim Sistemi - Temelleri ve Sözlüğü standardı, veri kalitesi değerlendirmesine rehberlik etmiştir. 2000'li yıllarda, konumsal veri kalitesi, değerlendirme ve ölçüm ilkelerini standartlaştırmak için birçok kalite standardı yayınlanmıştır. ISO

19157:2023, kalite elemanları, değerlendirme yöntemleri ve kalite ölçümleri için günümüzde geçerli olan kalite standardıdır (ISO, 2023).

Veri kalitesinin değerlendirilmesine yönelik temel kavramlar, “veri kalitesi elemanı”, “veri kalitesi kapsamı” ve “veri kalitesi raporu” aşağıda tanımlanmıştır.

Konumsal Veri Kalitesi Elemanı: Konumsal verilerin kalitesini belirleyen ayrı bir ögedir (ISO, 2013; 2023). Bir veri setinin “tamlığı” veya “mantıksal bütünlüğü” konumsal veri kalitesi elemanlarına örnektir.

Konumsal Veri Kalitesi Kapsamı: Konumsal veri kalitesinin kapsamı (scope), test edilecek verinin kapsamıdır. Konumsal verilerin veya ilişkilerin herhangi bir ortak özelliği, konumsal veri kalitesi kapsamı için tanımlanabilir (ISO, 2013).

Konumsal Veri Kalitesi Raporu: Kalite değerlendirmesinin bir sonucu vardır ve sonuç belgelenmelidir. Raporda niceliksel, uygunluk ve tanımlayıcı sonuçlar bulunabilir (ISO, 2013; 2023).

Teknolojik gelişmeler doğrultusunda haritalar sayısal hale gelmiştir. Bunun sonucunda haritalar için yeni üretim yöntemleri gelişmiştir. Bu gelişmeler konumsal verinin üretim araçlarını, yöntemlerini, depolanmasını ve yeniden kullanılabilirliği gibi özelliklerini etkilemektedir. Üretim yöntemlerine göre sınıflandırılan pek çok konumsal veri türü bulunmaktadır. KVA, yetkili kurumlar tarafından üretilen veriler, GCB ve açık veriler bunlara örnektir. Ancak her tür konumsal veri için aynı anlam taşıyan benzer kalite elemanları, bazen farklı anlamlarda, aynı zamanda ortak kalite elemanları veya tamamen yeni kalite elemanları da bulunmaktadır. Bu farklılıkların bir sonucu olarak, akademik çevreler ve Uluslararası Standartlar Teşkilatı (International Organization for Standardization, ISO) ve Açık Coğrafi Bilgi Konsorsiyumu (Open Geospatial Consortium, OGC) gibi kuruluşlar, kalite alanına ilişkin kavramları sürekli olarak standartlaştırma arayışındadır. Bu çalışmalardan elde edilen ortak kalite elemanları KVA’lar için aşağıda listelenmiştir.

KVA için Konumsal Veri Kalite Elemanları

Ulusal Dijital Kartografik Veri Standartları (National Committee for Digital Cartographic Data Standards, NCDCCDS) Komitesi başlangıçta beş veri kalitesi ögesini tanımlar (Moellering, 1987). 1995 yılında NCDCCDS, standardı zamansal doğruluk ve semantik tutarlılık olmak üzere iki kalite unsuruyla daha güncellemiştir.

Avrupa Standardizasyon Komitesi (Comité Européen de Normalisation, CEN), konumsal veri kalitesi için bir standart oluşturmuştur ve kalite unsurlarını tanımlamıştır

(CEN, 1991). ICA komisyonu konumsal veri standardını 1995 yılında; ABD Federal Coğrafi Veri Komitesi, Konumsal Veri Transfer Standardı'nı (Spatial Data Transfer Standard, SDTS) 1999 yılında yayımlamıştır. Tüm bu standartların kalite elemanları Tablo 3 ile özetlenmiştir ve sonrasında açıklanmıştır. ISO/TC 211, (2023) standartlarda yer alan veri kalitesi elemanlarını tanımlar. Bu standartlar veri kalitesi elemanlarını beş ila yedi sınıfta incelemektedir: Köken, konumsal doğruluk, öznitelik doğruluğu, semantik doğruluk, zamansal bilgi/doğruluk, mantıksal tutarlılık ve bütünlük (CEN, 1991; IGN, 1997; FGDC, 2000; Guptill ve Morrison, 2013; ISO, 2013).

Tablo 3, Devillers vd. (2005) kaynağından kalite elemanlarının neyi ifade ettiğine göre düzenlenmiştir, isimlerinde belli farklar olsa da içerikleri benzeyenler birlikte gruplandırılmıştır.

Tablo 3. Konumsal veri kalitesi sınıflandırması.

Veri Kalite Elemanları	CEN	ICA	ISO	SDTS	IGN
Köken	X	X	X		
Konumsal Doğruluk	X	X	X	X	X
Öznitelik Doğruluğu			X	X	
Semantik Doğruluk	X	X	X		X
Tamlık	X	X	X	X	X
Mantıksal Bütünlük	X	X	X	X	X
Zamansal Bilgi/Doğruluk	X	X	X		X

Köken (Lineage): Üretim tarihi, güncellemeler, verilerin nihai ürüne kadar olan dönüşümleri, veri kaynağının açıklamaları ve türetme yöntemleri dâhil olmak üzere verinin geçmişi hakkında bilgi verir (Clarke ve Clark, 1995).

Konumsal Doğruluk (Positional Accuracy): Veri kümesindeki koordinat değerlerinin referans veri kümesindeki değerlere uygunluğudur (Drummond, 1995).

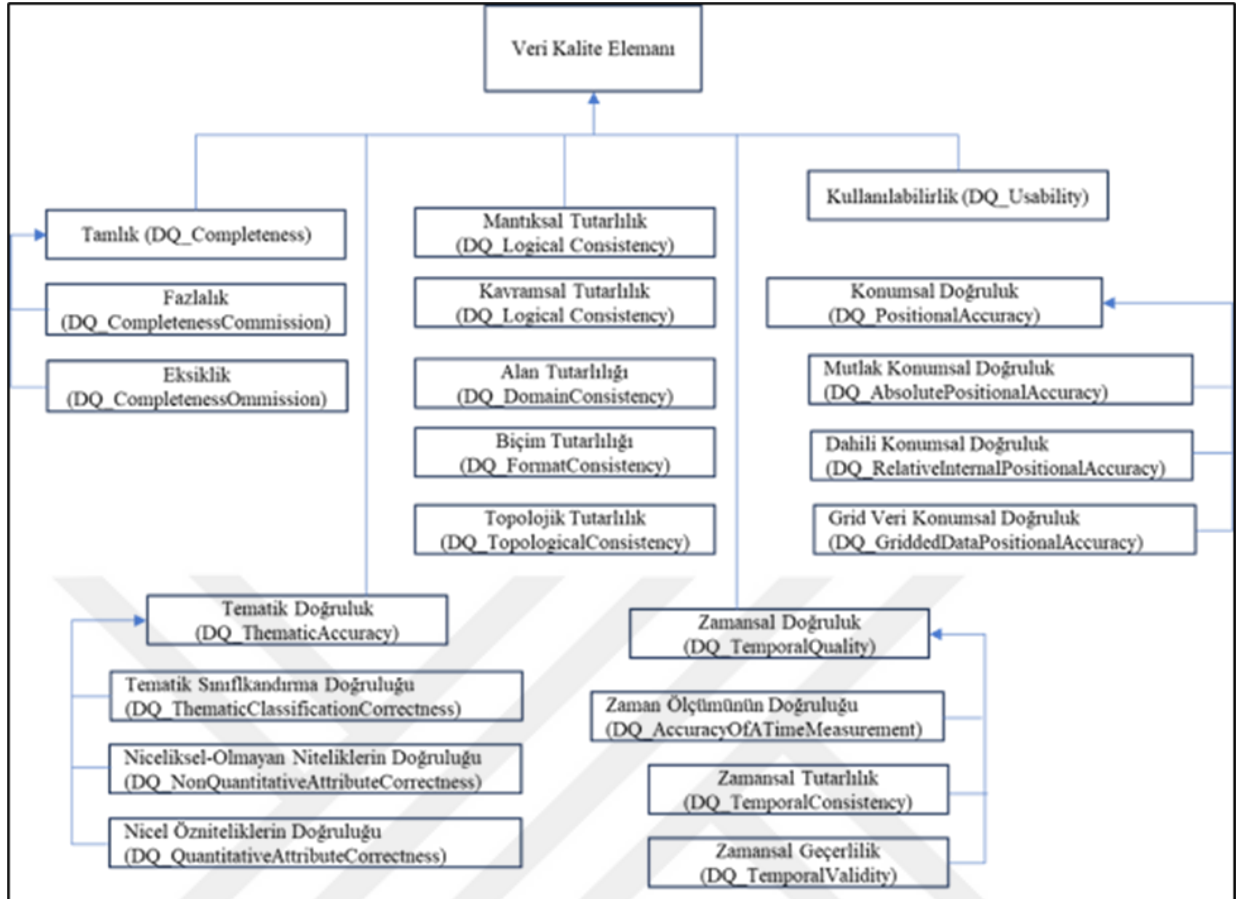
Öznitelik Doğruluğu (Attribute Accuracy): Öznitelik türleri nicel ve nitel olmak üzere iki ayrı sınıfta ele alınmaktadır. Detay adları, bina veya yol türleri niteliksel özniteliklere örnektir. Niceliksel öznitelikler, herhangi bir ölçüm veya analiz sonucunda ortaya çıkan sayısal değerleri, bir şehrin nüfusunu, bir yolun genişliğini veya bir ülkedeki şehir sayısını içerir. Öznitelik doğruluğu, niceliksel özniteliğin doğruluğunu ve niteliksel özniteliğin doğruluğunu ifade eder (Servigne vd., 2006).

Semantik Doğruluk (Semantic Accuracy): Bu kriter tamlık, nitelik doğruluğu gibi diğer bazı kalite kriterleriyle ilgilidir. Algılanan gerçekliğin ve modellenen gerçekliğin doğruluğudur. Dolayısıyla semantik doğruluk, spesifikasyonların modelleme doğruluğu olarak tanımlanabilir (Salgé, 1995).

Tamlık (Completeness): Bu kalite unsuru modele, verilere, nesnelere ve niteliklere uygulanabilir. Veri tamlığı, belirli nesnelerin ihmal edilmesi (veri yokluğu) veya beklenenden fazla veri bulunması hataları olarak tanımlanır. Modelin bütünlüğü, sağlanan temsilin kullanıcıların gereksinimlerine uygunluğu olarak tanımlanır (Brassel vd., 1995).

Mantıksal Tutarlılık (Logical Consistency): Mantıksal tutarlılık, konumsal veriler için mantıksal yapı kuralları ve öznelilik kuralları ile ilgilidir; verinin, veri setindeki diğer verilerle uyumluluğunu tanımlar. Bir veri setinin yapısal bütünlüğü ile ilgilenen konumsal veri kalitesi unsurudur. Bu bütünlük, konumsal verilerin ve detaylar arasındaki ilişkilerin modellendiği bir çerçeveye dayanır (Kainz, 1995).

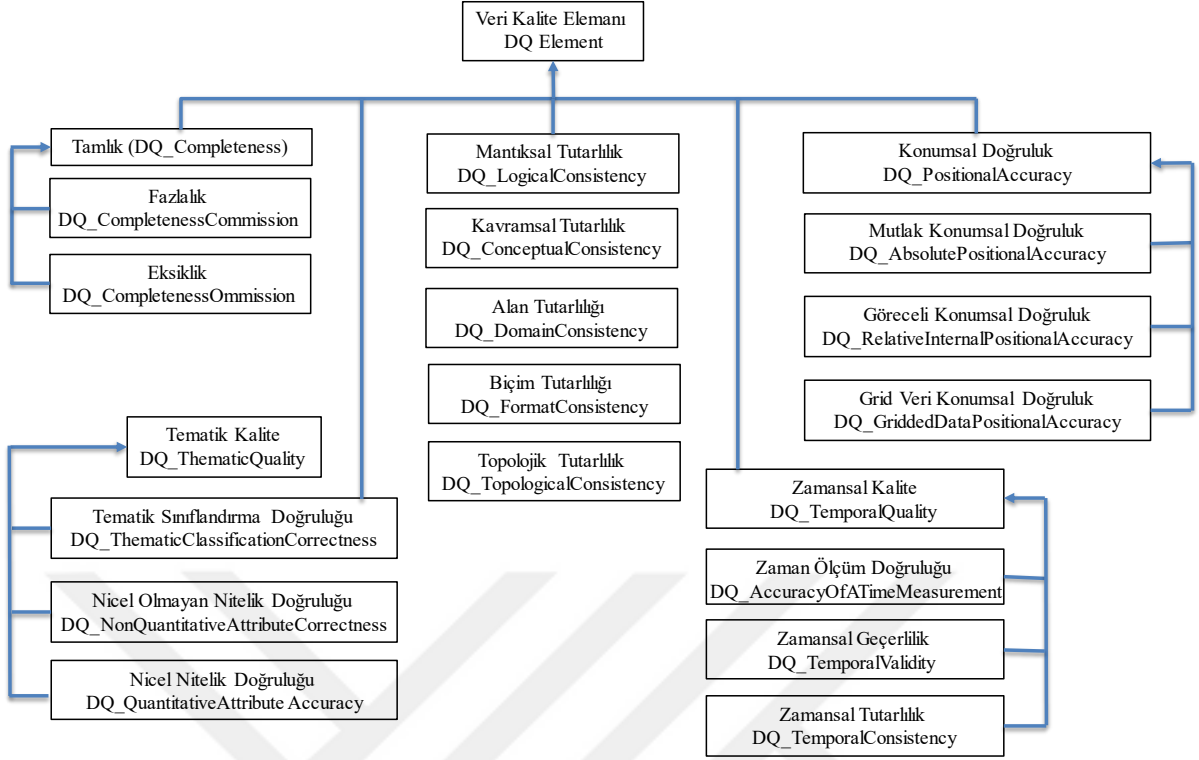
Zamansal Bilgi/Doğruluk (Temporal Information/Accuracy): Guptill ve Morrison (2013) konumsal verilerin zamansal yönleri hakkında bilgi sahibi olmanın önemini vurgulamıştır. Buna örnek olarak, bir veri detayının ne zaman toplandığı veya revize edildiği ile ilgili bilgilerin veri kalitesinin değerlendirilmesinde önemli bir faktör olduğunu belirtmiştir. ISO 19157: 2013 şeması Şekil 8 ile gösterilmiştir.



Şekil 8. ISO19157 2013 şeması (ISO, 2013).

1.3.6. ISO 19157:2023 Konumsal Veri Kalite Elemanları

ISO 19157’de konumsal veri kalite elemanları; tamlık (completeness), mantıksal tutarlılık (logical consistency), konumsal doğruluk (positional accuracy), tematik kalite (thematic quality), zamansal kalite (temporal quality) olarak beş ana başlıkta sınıflandırılmıştır. Ayrıca meta kalite elemanları diye bir grupta da vardır. ISO 19157 standardı kapsamında tanımlanan veri kalite elemanları Şekil 9 üzerinde gösterilmiştir.



Şekil 9. Veri Kalitesi Elemanları (ISO, 2023).

Tamlık: Öznitelikler, detaylar ve ilişkiler de dâhil olmak üzere doğru miktarda verinin varlığı olarak tanımlanır, “fazlalık” ve “eksiklik” alt elemanlarını içerir (ISO, 2023).

Mantıksal Tutarlılık: Bir veri seti oluşturmak için kullanılan mantıksal kurallara uygunluk düzeyidir. Kavramsal tutarlılık (conceptual consistency), etki alanı tutarlılığı (domain consistency), biçim tutarlılığı (format consistency), topolojik tutarlılık (topological consistency) olmak üzere dört alt kalite elemanı vardır (ISO, 2023) .

Kavramsal Tutarlılık: Bir veri setinin kavramsal şemasına uygunluğu olarak tanımlanır (ISO, 2023). Etki alanı tutarlılığı, öznitelik değerlerinin tanımlanan değer alanına uygunluğudur. Format tutarlılığı, veri setinin fiziksel yapısı için tanımlanan formatın uygunluğudur. Topolojik tutarlılık, veri seti için bildirilen topolojik kuralların doğruluğu olarak tanımlanır (ISO, 2023).

Konumsal Doğruluk: Konumsal verilerin gerçek dünyadaki konumları ne kadar doğru yansıttığını belirtir (ISO, 2023). Konumsal doğruluk üç alt kalite elemanı ile tanımlanmıştır; mutlak veya harici (absolute/external) doğruluk, göreceli veya dâhili (relative/internal) doğruluk, gridli (gridded) veri konum doğruluğu (ISO, 2023).

Mutlak doğruluk, konum değerlerinin bir referans veri setine göre yakınlığı olarak tanımlanır. Göreceli doğruluk, konum değerlerinin bir veri setindeki dâhili referansa göre

yakınlığıdır. Gridli veri konum doğruluğu ise raster formatındaki verinin piksel değerlerinin kabul edilmiş veya gerçek değerlerine yakınlığı olarak ifade edilmektedir (ISO, 2023).

Tematik Kalite: Veri setindeki niceliksel ve niceliksel olmayan özniteliklerin doğruluğu ve detayların mevzuata ya da kullanım amacına uygun sınıflandırılması olarak tanımlanır. Üç alt kalite elemanı vardır; tematik sınıflandırma (thematic classification) doğruluğu, niceliksel-olmayan (non-quantitative) niteliklerin doğruluğu, nicel (quantitative) öznitelik doğruluğu.

Tematik Sınıflandırma Doğruluğu: Detayların ve özniteliklerin sınıflandırılmasının bir referans veri seti ile karşılaştırmanın ne kadar doğru olduğudur. Niceliksel-olmayan öznitelik doğruluğu, niceliksel olmayan bir özniteliğin doğru olarak verilip verilmediğini ölçmek için tanımlanır. Nicel öznitelik doğruluğu, verinin öznitelik değerinin, özniteliğin referans değerine ne kadar yakın olduğunun belirlenmesi olarak tanımlanır.

Zamansal Kalite: Zamansal özniteliklerin ve detayların zamansal ilişkilerinin doğruluğu olarak tanımlanır. Üç alt kalite elemanı vardır; zaman ölçümünün doğruluğu (accuracy of a time measurement), zamansal tutarlılık (temporal consistency), zamansal geçerlilik (temporal validity). Zaman ölçümünün doğruluğu, bildirilen zaman ölçümlerinin doğru olarak kabul edilen veya bilinen değerlere yakınlığı ile ifade edilmektedir. Zamansal tutarlılık, sıralanan olayların doğruluğudur. (ISO, 2023).

Konumsal veri ile ilişkilendirilen iki ana doğruluk türü vardır; konumsal doğruluk ve tematik kalite. Konumsal doğruluk, verinin coğrafi konumunun hassasiyetini ve doğruluğunu ifade eder (Drummond, 1995; ISO, 2023). Konumsal doğruluk, GPS verileri, haritalama teknikleri ve coğrafi referanslama sistemleri kullanılarak değerlendirilir. Tematik doğruluk, verinin temsil ettiği detayların (yollar, binalar, su yolları vb.) doğruluğunu ifade eder; verinin detay seviyesi ve doğru sınıflandırılması ile ilgili bir değerlendirmedir.

1.3.7. Gönüllü Coğrafi Bilgi Kalitesi

Web 2.0'ın ortaya çıkışı, kullanıcıların da internet üzerinden içerik üretmesine ve paylaşmasına olanak sağlamıştır. Bu içeriğin önemli bir kısmı konumsal veridir. Son yıllarda internet ve diğer teknolojilerin gelişmesiyle kurumların konumsal veri üretiminin yanında kullanıcılar da bu sürece dâhil olmuştur (Goodchild ve Li, 2012b). Geleneksel yöntemle konumsal veri, belirli kuralları ve yönergeleri izleyen kurumlar tarafından üretilmektedir (Qi Zhou, 2018). Harita üreten ve kullanan kurumlarda üretilen verinin kalitesi, genel olarak

alan uzmanları tarafından mevcut yönergedeki kurallara göre değerlendirilmektedir. Bunun yanında GCB, yetkili veri setlerini ve ürünleri zenginleştirebilir, tamamlayabilir veya güncelleyebilir, hatta yenilerini oluşturmak için tek kaynak olabilir (V. Antoniou, 2011).

Ancak günümüzde GCB yaklaşımının ortaya çıkmasıyla birlikte, konumsal verinin büyük bir kısmı, CBS ve konumsal veri kalitesi konularında uzman olmayan kişiler tarafından toplanmaktadır. GCB veri kalitesine yönelik herhangi bir yönergenin takip edilmemesi, veri toplama yöntemlerinin farklı olması, mobil cihazların konumsal doğruluklarının farklılığı, kullanıcı algılarının değişkenliği ve Zielstra ve Hochmair (2013)'te vurgulanan homojenlikten yoksun olması gibi problemler, GCB ve kitle kaynak yaklaşımları ile üretilen verinin kullanılmasının önündeki en büyük engeldir. Bu nedenle GCB'yi kullanmak ve sunulan verinin kalitesini değerlendirmek için araçlara ihtiyaç vardır.

GCB'de veri toplama yönteminde bir standart bulunmadığından belirli bir kalite değerlendirme yöntemi de mevcut değildir. GCB'yi veri kalitesinin heterojenliğinden kaynaklanan belirsizlikten dolayı, GCB için kalite elemanları ve test yöntemleri bu yaklaşım ortaya çıktığı zamandan beri araştırma konusu olmuştur (Haklay vd., 2010; Goodchild ve Li, 2012b; Touya, 2012; Esmaili vd., 2013; Neis vd., 2013; Arsanjani ve Bakillah, 2015; Albakri, 2016; Jacobs, 2018; Chow vd., 2023; Yamashita vd., 2023). Hem literatürde hem de GCB gönüllü projeleri tarafından yapılan çalışmalarda KVA'lara ek olarak farklı kalite elemanları, ölçümleri, metrikleri ve test yöntemleri ortaya konulmuştur. Bazı çalışmalar GCB'ye özgü yeni kalite elemanları önermişlerdir. Antoniou ve Skopeliti (2015), ISO 19157 Standardının GCB verisine uygulanabilirliğini araştırarak standardın GCB'ye uygulanabilmesi için ek kalite göstergeleri önermiştir. Antoniou ve Skopeliti (2015), ISO 19157'de belirlenen kalite elemanlarının her biri için literatürde yapılan araştırmaları ve çözüm önerilerini ortaya koymuştur.

ISO 19157 standardına 2013 yılında standardın güncellenmesiyle eklenen "kullanılabilirlik" kalite elemanı, GCB literatüründe "kullanıma uygunluk" terimiyle ilişkili olduğu için GCB için en uygun kalite elemanıdır (Antoniou ve Skopeliti, 2015). Fakat bu kalite elemanı 2023 güncellemesindeki beş ana kalite elemanı arasında yer almamaktadır.

GCB için konumsal doğruluk veya tamlık gibi kalite elemanlarının değerlendirilmesinde genellikle yetkili kurumların ürettiği referans konumsal veri seti kullanılmaktadır. Goodchild ve Li (2012b), üç alanda veri analizi yoluyla GCB veri kalitesini sağlamak için içsel kalite değerlendirme elemanlarının ve yöntemlerinin uygulanabileceğini belirtmektedir. Bu göstergeler: a) Kitle kaynaklı revizyon; veri kalitesi

birden fazla katılımcı tarafından sağlanır, b) Sosyal ölçümler; katkıda bulunanların kendilerinin katkılarının kalitesinin değerlendirilmesine odaklanır ve c) Konumsal tutarlılık; katkıda bulunulan varlıkların tutarlılığının analizi yapılır.

GCB kalite değerlendirmesi ile ilgili çalışmaların çoğunda ortak olarak değerlendirilen kalite elemanları “Köken, konumsal doğruluk ve tamlık” tır (Sehra vd., 2013). Bu kriterlere ek olarak, “öznitelik doğruluğu, mantıksal tutarlılık, semantik doğruluk, zamansal doğruluk” kriterleri de çalışmalarda yer almaktadır. OSM’ye özgü önerilen bir diğer kalite kriteri de “etiketlerin doğrulanabilirliği”dir.

Van Exel vd. (2010), kitle kaynak için standartlarda mevcut olan kalite kriterlerine ek farklı kriterler olması gerekliliğini belirtmiştir. Ayrıca kitle kaynak verisinin kalite açısından temelde “kullanıcıyla ilgili kalite unsurları” ve “detaylarla ilgili kalite unsurları” olmak üzere iki ana kategoride ele alınması gerektiğini vurgulamıştır. Bu unsurlar, kitle kaynaklı verilere özgü kalite unsurlarıyla genişletilmiş mevcut kalite unsurlarından oluşabilir. Kullanıcı boyutu, bireysel bir katılımcının bakış açısından bilgi katkılarının kalitesini gösterir. Kullanıcı kalitesini belirlemek için üç bileşen önerisi sunmaktadırlar: Yerel bilgi, Deneyim ve Tanınma.

Kullanıcının Yerel Bilgisi: Kullanıcının bir bölgeye aşinalığı, katkısının konumsal ve zamansal modeliyle ilişkilendirilebilir. Bir katkının kalitesi, kullanıcının en aşina olduğu alanlar için daha yüksektir.

Deneyim: Bir kullanıcının katkısının kalitesi projeye katkıda bulunma konusundaki genel deneyimiyle ilişkilidir. Deneyim, kullanıcının projeye kayıtlı olduğu süre, projeye kaydettiği GPS izlerinin miktarı, eklenen veya düzenlenen detayların sayısı ve aynı zamanda proje bağlamında sanal ve gerçek hayattaki forumlardaki etkinliği kullanılarak ölçülebilir (Van Exel vd., 2010).

Tanınma (Recognition): Kitle kaynak veri kalitesinin Van Exel vd. (2010) tarafından önerilen üçüncü kullanıcı unsurudur. Çevrimiçi sosyal ağlarda ve kullanıcı katkılarına izin veren çevrimiçi bağlamlarda, genellikle itibar ile ilgili göstergeler belirlenir. Bu göstergeler kullanıcının projede ne kadar tanındığını belirtir.

Ballatore ve Zipf (2015), yaptıkları çalışmada GCB için kalite değerlendirme kriterlerini kapsamlı bir şekilde sınıflandırmıştır.

Doğruluk (Accuracy): Kavramsallaştırma ve alan bilgisi arasındaki farktır. Detayların sınıflara ayrılmasındaki doğruluk derecesi olarak görülebilir. Göstergeler; birden fazla

sınıflandırmaya sahip detayların sayısı, katkıda bulunanların sayısıdır (Ballatore ve Zipf, 2015).

Ayrıntı Düzeyi (Granularity): Çok soyut kavramlardan çok spesifik kavramlara doğru ilerleyen, verilerde mevcut tematik tanımlama seviyesidir. Göstergeler; sınıf hiyerarşisindeki sınıfların derinliği (Ballatore ve Zipf, 2015).

Tamlık (Completeness): İlgilenilen özelliklerin kavramsallaştırılmasındaki kapsamdır. Sınıf tamlığı ve öznitelik tamlığı arasında bir ayrım vardır. Göstergeler; sınıf sayısı ve öznitelik sayısıdır (Ballatore ve Zipf, 2015). Dolayısıyla ISO19157 veri kalitesi elemanı olan tamlıktan farklılıkları vardır.

Tutarlılık (Consistency): Detayların tanımlanmasındaki homojenlik derecesidir. Göstergeler: Aynı niteliklerle tanımlanan bir sınıftaki detayların sayısı; tutarlı detayların veya özniteliklerin diğerlerine oranı, mutlak nitelik sayısına karşı ağırlık (Ballatore ve Zipf, 2015).

Uyumluluk (Compliance): Bir özniteliğin, bir detayın veya bir dizi detayların belirli bir S kaynağına uyumsuzluktan tam uyumluluğa kadar değişen değerlerle bağlılık derecesidir. Gösterge; harici bir S kaynağında tanımlanan sınıf ve öznitelik sayısı ile toplam sınıf ve öznitelik sayısı arasındaki oran (Ballatore ve Zipf, 2015).

Zenginlik (Richness): Gerçek dünya varlığının tanımı için dâhil edilen detayların miktarı ve çeşitliliğidir. Gösterge; bir detayı tanımlayan özniteliklerin sayısı (Ballatore ve Zipf, 2015).

1.3.8. Gönüllü Coğrafi Bilgide Konumsal Doğruluk

Veri kalitesinin sürekli olarak izlenmesi ve değerlendirilmesi, yüksek kaliteli ve güvenilir konumsal veri setlerinin oluşturulmasında kritik bir rol oynar. GCB'nin veri toplama tekniği KVA'dan farklıdır. KVA verileri kurumlar tarafından ve geleneksel yöntemlerle üretilmektedir. GCB kalite değerlendirme yöntemleri ve kalite elemanları sunan birçok çalışma bulunmaktadır (Van Exel vd., 2010; Helbich vd., 2012; Antunes vd., 2015b; Zhou, 2017; Antoniou, 2017; Elias vd., 2021; Paiva ve Camboim, 2022). Veri türlerindeki farklılık nedeniyle, GCB'nin bazı kalite boyutları ve değerlendirme yöntemleri KVA'dan farklıdır (Goodchild, 2011).

Aynı zamanda GCB'nin konumsal doğruluğunu değerlendirmek için çeşitli çalışmalar yapılmıştır. Cipeluch vd. (2010), OSM'nin Google Maps ve Bing Maps'e göre konumsal

doğruluğunun analizini İrlanda'daki alanlar için yapmıştır. Çalışmada, bazı yerlerde bu kaynaklar arasında 10 metreye varan (Google Maps için) farklar olduğu sonucu ortaya çıkmıştır. Bu farkların, bazı detay türlerinin sadece düşük çözünürlüklü görüntüler üzerinden sayısallaştırılmasından kaynaklandığını belirtmiştir (Antoniou, 2017).

Haklay (2010), OS verileriyle OSM yol detaylarını karşılaştıran bir çalışma yapmıştır ve ortalama hatayı 5,8 metre bulmuştur. Bu hata düzeyi; çoğu arazi örtüsü haritası için ciddi sorun teşkil etmesi muhtemel olmayan, ancak küçük veya dar detayların (göletler, çitler, nehir kıyısı habitatları, vb.) gözden kaçmasına veya yanlış konumlandırılmasına neden olabilecek bir mesafe olarak yorumlanmıştır (Antoniou, 2017).

Canavosio-Zuzelski vd. (2013), vektörel veri kalitesinin bir parçası olarak OSM'nin konumsal doğruluk değerlendirmesini gerçekleştirmiştir. Ancak bu çalışmada, yetkili ölçme verileri referans olarak kabul edilmemiş, hem yetkili veriler hem de OSM verilerinden bağımsız stereo görüntülere göre değerlendirme yapılmıştır. Dolayısıyla GCB'nin yetkili veriden daha iyi olduğu konular tespit edilebilmektedir (Antoniou, 2017).

Zielstra ve Hochmair (2013) çalışmalarında, Flickr ve Panoramio fotoğraflarıyla ilişkilendirilen konumu, fotoğrafta neyin temsil edildiğini analiz eden yazarlar, belirlenen fotoğrafın konumuyla karşılaştırmıştır. Konumsal kaliteyi etkileyebilecek çeşitli hususlar tespit edilmiştir. Örneğin, bazı fotoğraflara atanan konum, fotoğrafın çekildiği konum iken, diğerleri için konumun neyi temsil ettiğine dair herhangi bir ek bilgi bulunmamaktadır ve muhtemelen biraz uzakta fotoğrafta belirtilen detayın konumudur. Konumsal kaliteyi etkilediği belirlenen diğer bir unsur ise, bölgede bulunan benzer detayların birbirine karıştırılmasıdır; bu durum fotoğrafların konumu bir uydu görüntüsü veya dijital harita üzerinde görüntülendiğinde ortaya çıkmaktadır (Antoniou, 2017).

Antoniou ve Skopeliti (2015) literatürde konumsal doğruluğu test etmek için kullanılan yöntemleri sınıflandırmıştır.

- Tampon bölge (Goodchild ve Hunter, 1997; Kounadi, 2009; Haklay, 2010),
- Yol ağının kesişimleri arasındaki mesafe (Antoniou, 2011),
- Noktasal detaylar için Öklid mesafesi, (Girres ve Touya, 2010; Jackson vd., 2013; Mashhadi vd., 2015),
- Çizgisel detaylar için ortalama Öklid uzaklığı ve Hausdorff uzaklığı (Girres ve Touya, 2010),
- Alan detaylar için ayrıntı düzeyi, kompaktlık ve yüzey mesafesi (Girres ve Touya, 2010),

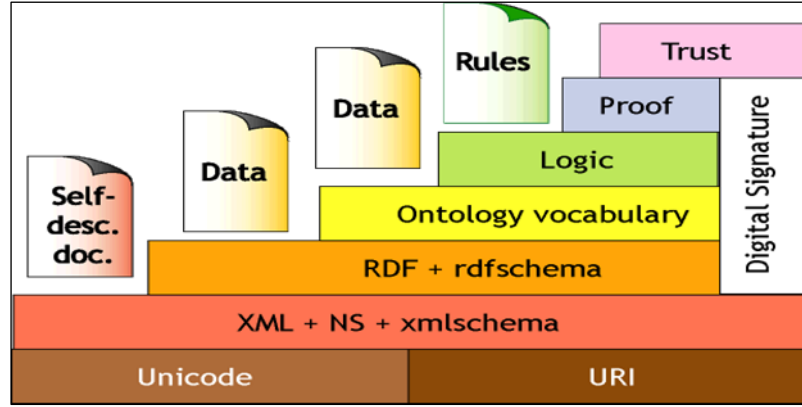
- Şekil benzerliği (döndürme fonksiyonu) (Mooney vd., 2010; Fan vd., 2014; Kalantari ve La, 2015),
- x ve y hata mesafesi (Stark, 2011),
- Grid tabanlı minimum sınırlayıcı geometri ve yönlü dağılım (Forghani ve Delavar, 2014),
- Köşe sayısı, ortalama köşe mesafesi ve poligon merkezleri arasındaki mesafeler (Kalantari ve La, 2015),
- Yönsel ve metrik mesafe ilişkilerini dikkate alan çoklu temsilde konumsal benzerlik (Hashemi ve Abbaspour, 2015).

1.3.9. Semantik Web ve İlgili Kavramlar

SWT ve ontolojiler, CBS ve veri yönetiminde önemli rol oynar. Bu bölümde SWT ile ilgili temel kavramlar tanımlanmıştır.

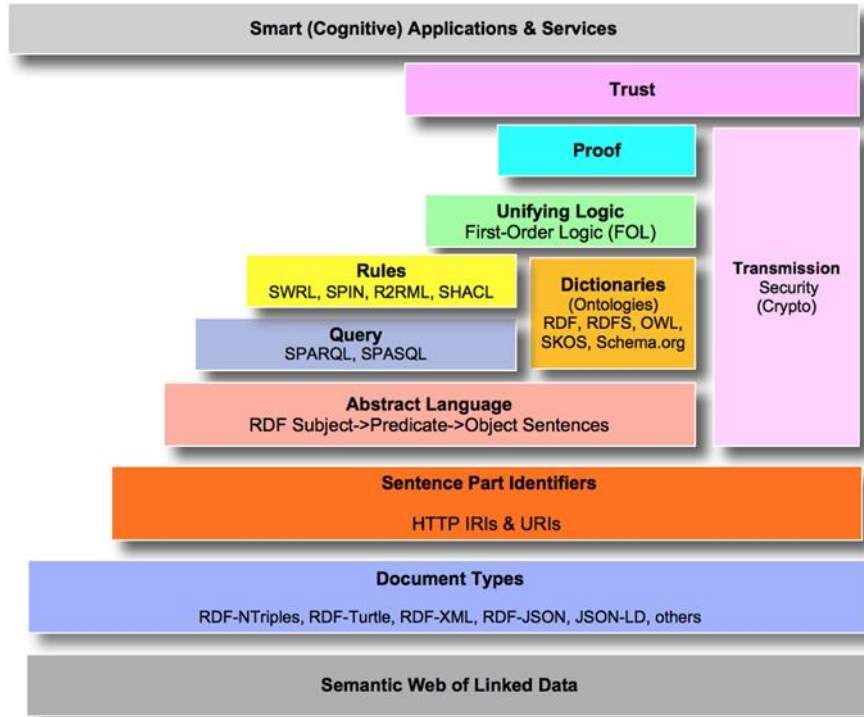
Günümüzde internetteki bilgilerin çoğu insanların yorumlayabileceği şekilde kodlanmıştır. Web'deki bilginin büyüklüğü ve bilgiye olan ihtiyacın hızla artması, bilgilerin sadece insanlar tarafından değil bilgisayarlar tarafından da yorumlanabilmesini zorunlu kılmaktadır. Bu gerekliliği fark eden Berners-Lee ve Hendler (2001), "Semantik Web"i ortaya atmışlardır. Tim Berners-Lee, "Semantik Web"i "Veri Ağı" olarak tanımlar. "Semantik Web"in altında yatan fikir, web üzerinden erişilebilen verileri ontolojiler aracılığıyla semantik olarak süslemektir. Bu, web belgesinin içeriğinin çok fazla insan müdahalesine gerek kalmadan bilgisayarlar tarafından okunmasına ve anlaşılmasına yardımcı olacaktır. Semantik Web, web sayfalarının semantik içeriğinin web tabanlı olarak geliştirilmesidir; sayfalar arasında gezinen yazılım araçları, karmaşık görevleri kolaylıkla gerçekleştirebilir (Berners-Lee ve Hendler, 2001). SWT, yalnızca web sayfalarını veya görsellerini değil aynı zamanda insanları, yerleri ve kuruluşları da tanımlar; bilginin, çeşitli ortamların ve otomasyonun keşfini ve yeniden kullanımını daha etkili hale getirir (Guha ve McCool, 2003).

2001 yılı SWT mimarisi Şekil 10 ile gösterilmektedir.



Şekil 10. Semantik Web katmanlı yapısı

Bu katmanlı yapının daha yeni versiyonu Şekil 11 ile gösterilmektedir. Bu iki şekilde görüldüğü gibi ontolojiler SWT'nin kalbinde yer almaktadır. Yeni sürümde sorgular ve kurallar için sırasıyla SPARQL ve SWRL gibi seçenekler bulunmaktadır (W3C, 2004).

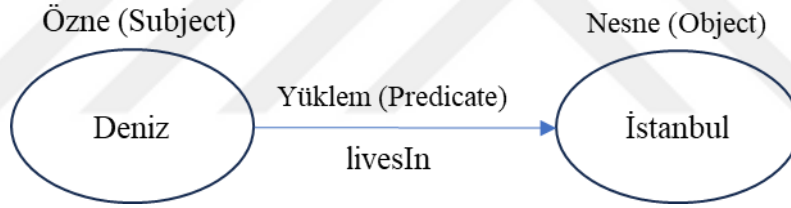


Şekil 11. SWT katmanlı yapısı (Idehen, 2017).

Ontoloji, modelleme alanında standart olarak kullanılır. Gerçek dünyadaki nesneler veya kavramlar, özellikleri ve ilişkileriyle birlikte ifade edilebilir. Ontolojilerle kavramlar

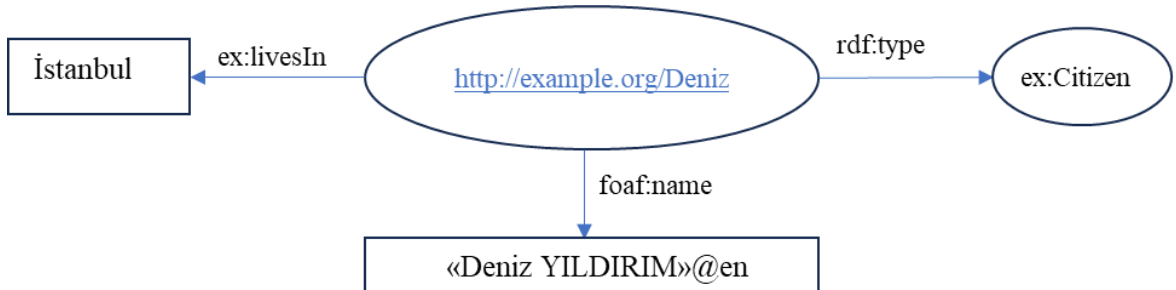
arasında semantik ilişkiler kurulur; bu şekilde kavramlar standart bir şekilde ifade edilir. Ontolojiler hiyerarşileri alt sınıf-üst sınıf ilişkilerini kullanarak tanımlar. “Erkek” ve “Kadın”, “İnsan” adlı bir sınıfın alt sınıfları olan ontolojik sınıflar olabilir.

Kaynak Açıklama Çerçevesi (Resource Description Framework, RDF): Semantik Web'i geliştirmek için ana temsil çerçevesidir ve web kaynakları arasındaki bağlantıyı sağlayan verileri temsil etmek için grafik yapısını kullanır (W3C, 2014). RDF, (yönlü) grafiklerin soyut bir versiyonunu tanımlayan üçlülere kullanır (Antoniou ve Harmelen, 2008; W3C, 2014). RDF'de özne-yüklem-nesne üçlüsü ile yüklemelere göre öznenen nesneye bir “kenar” vardır. “ex:Deniz ex:livesIn ex:Istanbul” gibi bir ifade, Şekil 12 ile gösterildiği gibi RDF özne-yüklem-nesne üçlülere olarak temsil edilebilir. Her bir düğümü ve kenarı temsil etmek için RDF, Tekdüzen Kaynak Tanımlayıcısı (Uniform Resource Identifier, URI) kullanır. Çeşitli RDF serileştirme formatlarında örneklerle birlikte kullanılır. URI'ler kaynakları tanımlayan karakter dizileridir (Berners-Lee vd., 2005). Web adresleri gibi lokasyonla tanımlayan URL veya isimle tanımlayan URN gibi çeşitleri vardır.



Şekil 12. RDF üçlü örneği ex:Deniz ex:livesIn ex:Istanbul.

Yazara ilişkin bir bilgi tabanı Şekil 13 üzerinde gösterilmiştir. Bu şekilde Deniz Yıldırım'a ilişkin örnek bir bilgi tabanı çizge ile gösterilmiştir.



Şekil 13. RDF çizgesi.

Web Ontoloji Dili

Web Ontoloji Dili (Web Ontology Language, OWL), ontolojileri tanımlamak için kullanılan bir Semantik Web dilleri ailesidir (W3C, 2012). Birincil OWL dilleri OWL Full, OWL DL ve OWL Lite'tır. OWL ontolojileri, bir etki alanındaki kavramlar ve varlıklar hakkında açık mantıksal karşılık gelen aksiyom kümesidir. OWL ontolojileri, RDF formatlarında serileştirilir ve geçerli bir OWL dokümanı geçerli bir RDF dokümanıdır, ancak bunun tersi her zaman doğru değildir. RDF dokümanları genel olarak owl:Ontology gibi OWL yapılarını içermez.

OWLAPI

OWLAPI, OWL ontoloji tabanlı uygulamalar için açık kaynaklı bir JAVA API'sidir. Ontoloji düzenleme, oluşturma ve serileştirme gibi ontolojilerle yapılacak ana işlevler OWLAPI kullanımı ile uygulanabilir. Pellet ve FACT++ gibi popüler ontoloji çıkarsama araçları desteklenmektedir (Horridge ve Bechhofer, 2011).

2. YAPILAN ÇALIŞMALAR

Yetkili kurum verilerinin açık veri kaynakları ile güncellenebilirliğinin değerlendirilmesi için bu çalışmada ontoloji tabanlı veri kalitesinin değerlendirilmesine ve iyileştirilmesine yönelik bir metodoloji geliştirilmiştir. Ayrıca, yetkili verilerin daha iyi hale getirilmesini hedefleyen bu tez çalışmasında, bu verileri kendi içinde ve diğer yetkili verilerle birlikte işler bir şekilde kalite değerlendirmesini yapan, temel geometrik hata düzeltmelerini yapmaya ön ayak olan SDQO ontolojisi ön değerlendirmede kullanılmıştır (Yılmaz, 2018). Yetkili verilerin, kitle kaynak veriler ile iyileştirilebileceği görülmüş, daha kullanılabilir, doğru ve güncel olabileceği değerlendirilmiş, bunun sonucunda verilerin daha da geliştirilmesi için bir metodoloji ve çerçeve geliştirilmiştir.

2.1. Genel Çerçeve

Bu çerçevede öncelikle SDEO ontolojisi geliştirilmiştir. İşlemler bu ontoloji güdümünde ilerler. Metodoloji kapsamında SDEO ile belirlenen çok sayıda seçenek vardır, örneğin yöntemlerin kendileri ve parametreleri SDEO ile belirlenir. Literatürde mevcut yöntemler ya da benzerleri kullanımdadır. Ayrıca iki boyutlu yapılar için Sanal Ara Poligon (SAP) yöntemi geliştirilmiştir.

SDQO ontolojisinde olduğu gibi veri ontolojileri de kullanılmaktadır. Ayrıca, poligon HGM katmanları için GeoJSON dosyaları oluşturulup geometrik veriler burada tutulmaktadır.

Ontoloji tasarımında, yetkili kurum personelinin gerektiğinde bu ontolojiler üzerinde düzenlemeler yapacağı göz önüne alınmıştır. Bazı işlemler için bu personelin ontoloji konusunda uzman olan araştırmacı ile, seyrek de olsa, beraber çalışması gerekebilecektir (özellikle notasyon (annotation) özellikleri).

Verilerin yaygınlığı ve ulaşılabilirliği göz önüne alınarak, yetkili veri kaynağı olarak HGM ve kitle kaynak veri olarak OSM seçilmiştir. Çalışma bölgesi Trabzon olarak belirlenmiştir. İyileştirme işlemi sırasında HGM ve OSM veri setlerinin entegrasyonu ve analizi gerçekleştirilmektedir.

Önerilen metodolojinin mimarisi Şekil 14 ile gösterilmiştir. Ayrıca bu çalışmada kullanmak için MATLAB üzerinden çalışan bir temel GUI oluşturulmuştur. Çalışma

Türev vektör veriler (kitle kaynak ve yetkili) raster bantlara dönüştürülür. Bu bantlar kurum personelinin önerisi üzerine üretilmiş olur. Örneğin, bir bantta kitle kaynak veride eşleniği olan yetkili verilerin gösterimi yapılabilir.

Gösterim için bantlar seçilir ve kurum personeli incelemelerini yaparak kurum içerisindeki sonraki adımlar için noktaları belirler. Örneğin, tavan-arası tarihinden sonra kitle kaynak veride gerçekleşen değişiklikler, kurum uzmanları tarafından değerlendirilir.

Doğal afetlerin olması gibi durumlarda daha kaliteli veriler ulaşılabilir hale gelir. Kurum personeli tarafından SDEO ontolojisi üzerinden seçilen yöntemler bu değişime uygunsa, örneğin yapı boyutlarının (yerel) çözünürlüğe oranı yöntemlerin çalışma kalitesini çok etkilemiyorsa, başka bir değişikliğe gerek kalmadan metodoloji uyarlanabilir. Çalışmada geliştirilen yöntemlerde bu nokta da göz önüne alınmıştır; örneğin çözünürlük (yarı) bağımsızlığı önemsenmiştir.

Çalışmanın esnek yönlerinden biri de yöntemlerin çeşitliliğidir. Çalışma sonrasında çıkan yöntemler ya da hâlihazırda uygulamada GUI uygulamasıyla ulaşılamayan yöntemler için kurum personeli araştırmacıya başvurabilir. Mevcutta literatürde sıkça karşılaşılan, örneğin ağırlık merkezi tabanlı yöntemlerin geliştirilmişlerine ve yeni ortaya koyduğumuz yöntemlere GUI uygulaması ile ulaşılabilir.

2.2. SDEO Üst Hiyerarşik Yapısı

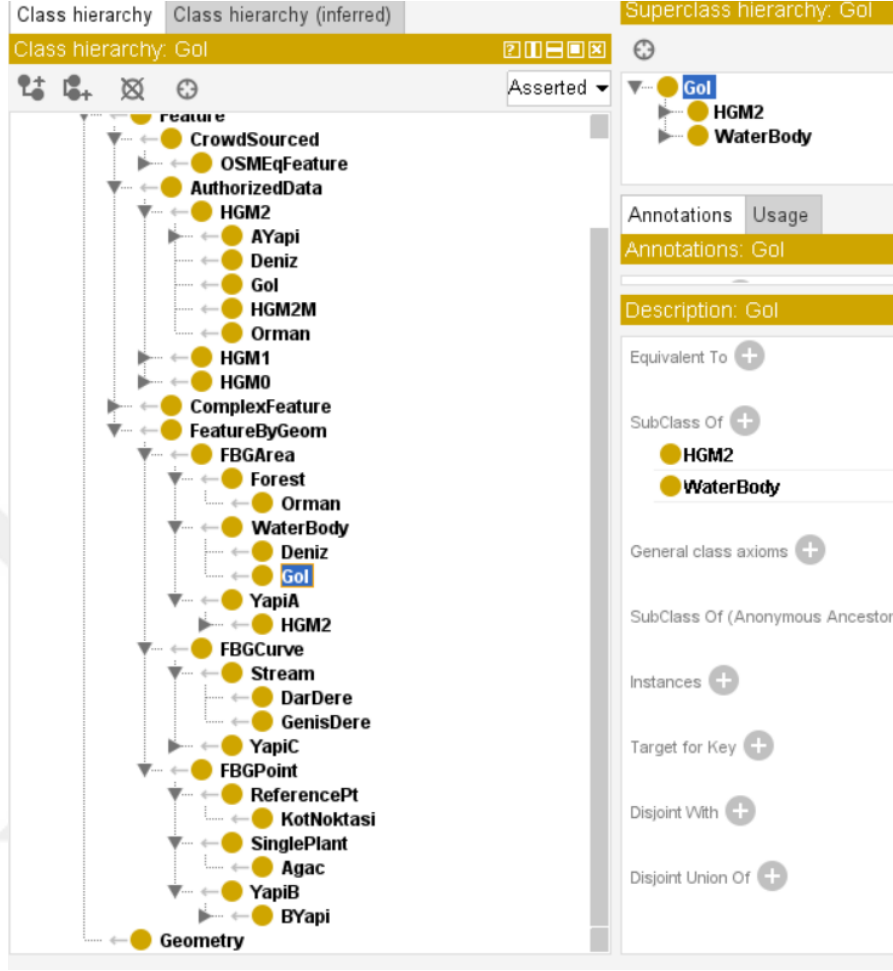
SDEO üst hiyerarşik yapısına bu bölümde değinilmiştir. İleriki bölümlerde ontoloji hakkında daha ayrıntılı bilgiler sunulmaktadır.

2.2.1. Üst Sınıflar

geo:SpatialObject, OGC GeoSPARQL ontolojisi kaynaklıdır ve geo:Feature ve geo:Geometry sınıflarını içerir. geo:Feature alt sınıfları Şekil 15 üzerinde gösterilmiştir. Şekilde görüldüğü gibi HGM ile ilgili sınıflar sdeo:AuthorizedData sınıfı altındadır. sdeo:OSMEqFeature sınıfı elemanları ilişkilendirilmiş OSM detaylarına karşılık gelir. Farklı kitle kaynak ya da yetkili veri katmanları bu sınıfların alt sınıfları olarak eklenebilir. sdeo:ComplexFeature alt sınıflarında geometri tipleri homojen değildir.

sdeo:Aspect, SDQO veri kalitesi sınıflarını içerir.

sdeo:Process, sdeo:Method, sdeo:SpaceTime ve SDQO kaynaklı sonuç ve süreçleri içeren sdqo:DataQualityResult sınıfını içerir.



Şekil 15. Detay Sınıfları.

2.2.2. Veri Tipi Üst Özellikleri

sdeo:Attribute; özniteliklere karşılık gelen veri tipi özelliklerinin sınıflarını içerir, soyut bir üst özelliktir. Alt özellikleri sdeo:HGMAttribute, sdeo:MethodAttribute, sdeo:MixedAttribute ve sdeo:OSMAttribute gibi veri kaynağına göre ayrılmış somut özelliklerdir. Bunlar da özniteliklere karşılık gelen özellikleri içerir.

sdeo:ID üst özelliği; sdeo:HGMGlobalID, sdeo:OSMID ve sdeo:shortID özelliklerini içerir. Bu özellikler detayları belirleyici kimlik verilerdir.

sdeo:LocationProp üst özelliği, enlem, boylam ve yükseklik özelliklerini içerir. sdeo:hasCentroid alt özelliklerden biridir.

2.2.3. Nesne Özellikleri

geo:hasGeometry, OGC GeoSPARQL kaynaklıdır, detayları geometrilerine bağlar.

sdeo:hasAnchor, yöntemleri çapa noktalarına bağlar.

sdeo:isAlike, HGM detaylarını ilişkilendirildikleri OSM detaylarına bağlar. Aynı HGM detayı birkaç defa ilişkilendirilebilir.

sdeo:pointToPoly, nokta tipindeki HGM detaylarını, kapalı çizgi tipindeki OSM detaylarına bağlar. Bir SWRL kuralı ile eğer bir HGM noktası sdeo:isAlike ile OSM kapalı çizgisi ile ilişkilendirilmiş ise, aralarında otomatik olarak sdeo:pointToPoly özelliği ile bağlanır.

sdeo:hasNextNode, sıfır boyutlu yapıları bağlar, sıra oluşturur.

sdeo:hasPart, yöntemleri ve detayları parçalarına bağlar.

2.2.4. Notasyon Özellikleri

sdeo:notasyonHGM, sdeo:notasyonOSM özellikleri ile bunları içeren sdeo:notasyon üst özelliği oluşturulmuştur. Ayrıca rdfs:isDefinedBy özelliğinin kullanımı uygulamada işlevseldir.

2.3. HGM Verilerinin Hazırlanması

Tez çalışmasında, yetkili veri kaynağı için konumsal veri üreten ve/ya sağlayan çeşitli resmî kurumlar incelenmiştir, HGM kurumunun veri kullanımı ve paylaşımında kolaylık sağladığı görülmüştür.

Türkiye'nin tamamına ait TOPOVT vektör verileri 2019 sonu itibarı ile üretilmiştir. Tüm ülke sınırlarını kapsayan 76 adet 1:250.000 ölçekli alandan yıllık ortalama 7-8 adedi bölgesel olarak, Yol, Koprugecit, Tunnel, Altgecit detay sınıfları ise tüm ülke sınırlarını kapsayan alanda 2 yıllık periyotlarla güncellenmektedir; haritalar 1:25000 ölçekli olarak sunulmaktadır (HGM, 2019).

Trabzon iline ait katmanlar tez çalışması kapsamında HGM'den temin edilmiştir. 1:25000 ölçekte yarım milimetre 12,5 metreye karşılık gelir, bu da Trabzon civarında enlem boylam cinsinden yaklaşık olarak 0,0001 dereceye karşılık gelir. Bu değer yüze bölünmüş ve tüm koordinat verileri bu çözünürlüğe göre düzenlenmiştir. Uygulamada iki farklı çözünürlük vardır, diğeri sdeo:resolution veri tipi özelliği ile belirlenir.

Veri çeşitli katmanlara ayrılmış bir şekilde ESRI Shapefile formatlarında (shp, shx, prj, dbf, cpg, sbx, sbn, shp.xml) kaydedilmektedir. Örneğin Trabzon'daki çeşitli bölgeler için, “büyük bina” sınıfı için buyukbina.shp ve yedi tane daha (metin tabanlı veya ikili) diğer uzantılı dosya vardır. Bu da işlem kolaylığını ve performansı azaltabilir. Trabzon ili 23 farklı bölgeye ayrılmıştır.

HGM (2020a) incelenip işlemler yapıldığında 129 tane katman olduğu görülür (fotogrametrik katmanların önemli bir kısmı alt detaylara dönüştürülmüştür) ama bunların sadece 84 tanesinin Trabzon verisinde kullanıldığı tespit edilmiştir. 23 bölgede ortalama olarak 45 katman vardır. SDEO ontolojisinde katmanların 26 tanesine karşılık gelen sınıflar mevcuttur. Dosya boyutları incelenmiş, toplam dosya boyutunun % 98,5 gibi bir kısmının karşılandığı görülmüştür. Pareto ilkesine göre % 80 gibi bir oran beklenirdi. Ontolojideki katmanların varlığı HGM personelinin kontrolünde olacaktır.

HGM verisinde katmanlar geometrik tiplerine göre ayrılmış ama farklı klasörlere konmamıştır; aynı bölge klasöründe hem 8 tane küçük bina (nokta tipinde), hem de 8 tane büyük bina (poligon tipinde) dosyası vardır. Her katmana ait bir Topo detay sıra numarası (topodetays) vardır, örneğin büyük binalar için bu numara 66'dır (HGM, 2020b). Alt tipler için de “topodetaya” değeri vardır. Bina kullanım amacı (bikula) ve yapı fonksiyon kategorisi (yafonkka) değerleri için Tablo 4 incelenebilir. SDEO'da bu özniteliklere karşılık gelen veri tipi özellikleri bulunmaktadır.

Tablo 4. Bina Kullanım Amacı ve Yapı Fonksiyon Kategorisi.

Değer	bikula	yafonkka
0	Ağırlık Kontrol İstasyonu	Ambar
1	Anaokulu	Bina
2	Anten Haberleşme Binası	Cami
3	Arıtma Tesisi	Cezaevi
4	Bekçi Kulübesi	Demiryolu Makas Binası
5	Demiryolu İstasyonu	Demiryolu Yakıt İkmal İstasyonu
6	Dispanser	Deniz Feneri
7	Gümrük Binası	Eğitim Kurumu
8	Haberleşme	Fabrika
9	Hastane	Gayri Müslim Türbesi
10	İbadethane	Gözlemevi

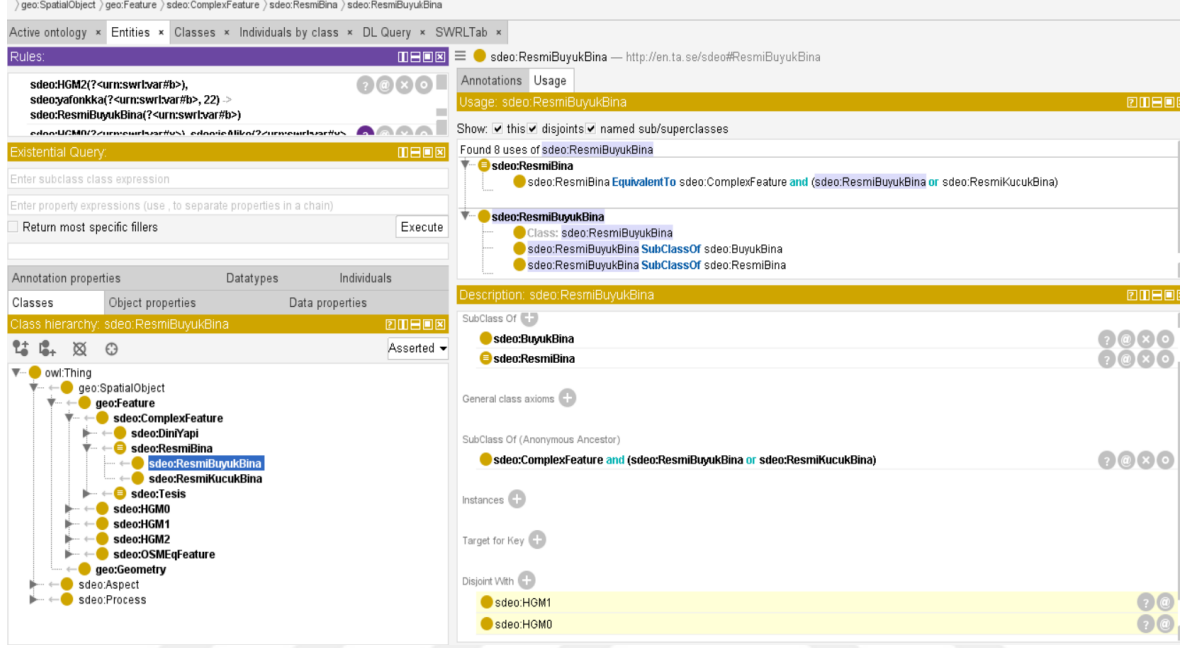
Tablo 4'ün devamı

11	İlkokul	Haberleşme Tesisi
12	İlköğretim Okulu	Hangar
13	İlk Yardım İstasyonu	Hükûmet Binası
14	Karakol Binası	İslami Türbe
15	Konut	İtfaiye
16	Kreş	Kilise
17	Kule Kontrol	Kubbe
18	Lise	Manastır
19	Meslek Lisesi	Meteoroloji İstasyonu
20	Meslek Yüksek Okulu	Minaresiz Cami
21	Meteoroloji	Müze
22	Ortaokul	Resmi Bina
23	Poliklinik	Sağlık Kurumu
24	PTT	Saray
25	Radar İstasyonu	Sera
26	Radar Telsiz İstasyonu	Sinagog
27	Resmi	Tapınak
28	Sağlık Evi	Toplantı/Konser Salonu
29	Sağlık Ocağı	Yayla Evi
30	Tarım	
31	Telefon Kulübesi	
32	Trafik Kontrol Noktası	
33	TV İstasyonu	
34	Üniversite	
999	Diğer	Diğer

HGM yapıları ve eşdeğer OSM yapıları ilişkilendirildikten sonra, SWRL kuralları ile, iki taraftaki eksik öznitelikler ve diğer bilgiler tamamlanabilir. Aşağıdaki SWRL kuralında yafonkka değeri 22 olanların resmi olması durumu kullanılmıştır. ResmiBuyukBina adında bir HGM katmanı yoktur, ontolojide yeni oluşturulmuştur.

sdeo:HGM2(?b), sdeo:yafonkka(?b, 22) → sdeo:ResmiBuyukBina(?b)

Ayrıca, Şekil 16 incelendiğinde sdeo:ResmiBuyukBina için çıkarsama sonrası tanımlaması görülür. sdeo:HGM2 kaynaklı olarak sdeo:HGM1 ve sdeo:HGM0 ile ayrıklığı ortaya çıkmıştır. Dolayısıyla sdeo:ResmiKucukBina ile de ayrıktır, ayrıca ikisi, sdeo:ResmiBina'yı tanımlar.



Şekil 16. Çıkarsama sonrası sdeo:ResmiBuyukBina.

Çalışmada ortaya konan yöntemde ilişkilendirmenin ön aşaması kimlik değerlerinin atanmasıdır. Bunun için ESRI Shapefile dosyalarında var olan bir kimlik değeri kullanılacaktır.

2.3.1. ESRI Shapefile HGM Verisine sdeo:shortID Değer Atanması

Shapefile, Environmental Research Institute (ESRI) tarafından ortaya konmuş açık kaynak ama ticari bir formattır. Geliştirildiği 1980'li yıllar ve öncesinde yaygın olduğu gibi coğrafi ilişkisel veri tabanı türünde bir dosyalar bütünüdür. PoligonZ kayıt verisi bilgisi orijinal halinde (1988 yılı) Tablo 5'te gösterilmiştir (ESRI, 1988). Modern alternatifleri OGC GeoPackage ve GeoJSON formatlarıdır.

MATLAB Harita araç kutusunda ESRI Shapefile dosyalarını okumak için shaperead fonksiyonu kullanılır. HGM'nin sunduğu dosyalara uygulandığında, Well Known Text (WKT) formatındaki geometrik veri sütunlarını okuyamamaktadır. MATLAB readgeotable

komutu ile tablo olarak okunduktan sonra, geotable2table komutu ile WKT veriye yönelik yeni bir sütun eklenerek tablo olarak geometrik veri de kullanılabilir hale gelmiştir ve bu sorun aşılmıştır. readgeotable komutu GeoJSON dosyalarını da okur. GeoJSON dosyalarının bir dezavantajı öznitelik sayısı çoğaldığında yaşanan performans problemidir. Uygulamada öznitelikler veri ontolojisinde saklandığından bu dezavantaj da bertaraf edilmiştir.

Tablo 5. ESRI Shapefile teknik tanım PoligonZ kayıt verisi bilgisi.

Position	Field	Value	Type	Number	Byte Order
Byte 0	Shape Type	15	Integer	1	Little
Byte 4	Box	Box	Double	4	Little
Byte 36	NumParts	NumParts	Integer	1	Little
Byte 40	NumPoints	NumPoints	Integer	1	Little
Byte 44	Parts	Parts	Integer	NumParts	Little
Byte X	Points	Points	Point	NumPoints	Little
Byte Y	Zmin	Zmin	Double	1	Little
Byte Y+8	Zmax	Zmax	Double	1	Little
Byte Y+16	Zarray	Zarray	Double	NumPoints	Little
Byte Z*	Mmin	Mmin	Double	1	Little
Byte Z+8*	Mmax	Mmax	Double	1	Little
Byte Z+16*	Marray	Marray	Double	NumPoints	Little

Note: $X = 44 + (4 * \text{NumParts})$, $Y = X + (16 * \text{NumPoints})$, $Z = Y + 16 + (8 * \text{NumPoints})$

ESRI Shapefile verilerinde GlobalID isminde bir değişken bulunur. Diğer evrensel belirleyiciler gibi 128 bitliktir ama 16'lık sistemde ve 32bit + 16bit + 16bit + 16bit + 48 bit şeklinde 8+4+4+4+12 karakter halinde gruplanmıştır (örneğin {E28FE7B7-7B80-4341-94D2-4C647D4F79C6}). Ticari bir format olmasının da etkisiyle ne ölçüde rastgele atandığı bilinmemektedir. Bu rastgeleliği ölçmek ve yapı komşulukları ile olası ilişkisine bakmak için MATLAB corrcoef fonksiyonu kullanılarak bir deneme yapılmıştır. HGM verisinde rastgele 1000 yapı seçilmiştir. GlobalID değerlerinin ilk dört karakterleri alınmıştır ve onluk sistemde sayılara çevrilmiştir. Yapıların ağırlık merkezleri hesaplanmıştır. Bu ağırlık merkezleri arasındaki Manhattan Mesafesi ile ilk dört karakterin sayısal karşılıkları arasındaki mutlak mesafeler arasındaki korelasyon, -0.0015 olarak hesaplanmıştır. Tekrar test edildiğinde tekrar sıfıra yakın ve genelde negatif korelasyonlar çıktığı görülmüştür. Dolayısıyla GlobalID oluşturulurken muhtemelen rastgele oluşturulmaktadır, rastgele oluşturulmuyorsa da komşu binaların GlobalID değerleri uzaktır.

SDEO ontolojisinde tüm 128 bitlik GlobalID yerine daha az bir kısım (örn: 16 bit) (sdeo:shortID) kullanılacaktır. Bir notasyon özelliği ile hangi kısmın olacağı belirlenir. Raster dönüşümünden sonra o katmandaki her yapıya bu sdeo:shortID kimlik değeri atanmıştır ve bu değerler veri ontolojisinde ve GeoJSON dosyalarında tutulacaktır. Tüm sınırlayıcı dörtgen (bounding box) içindekiler için olmasa bile, birbirine çok yakın yapılarda bu belirleyici değeri farklı olmalıdır. Bu şekilde, raster aşamasında OSM yapılarının HGM eşdeğerlerinin bulunması ve ilişkilendirilmesi hedeflenmektedir.

Az Bit Kullanımı Riski

GlobalID için 128 bit yerine 16 bit kullanılması durumunda farklı yapılara aynı ID atanma riski artar. Riski hesaplamak için çeşitli işlemler yapılmıştır. Bu riskin gerçekleşme olasılığı; M tane yapı ve N bitte tüm yapılara farklı değeri verilmiş olma olasılığı $(1 - 2^{(-N)})^M$ dir. Bu olasılık da yaklaşık olarak $e^{(-M \cdot 2^{(-N)})}$ olur. Kaç kez çalıştırıldığında en az bir aynı değeri atanma riskinin formülü de

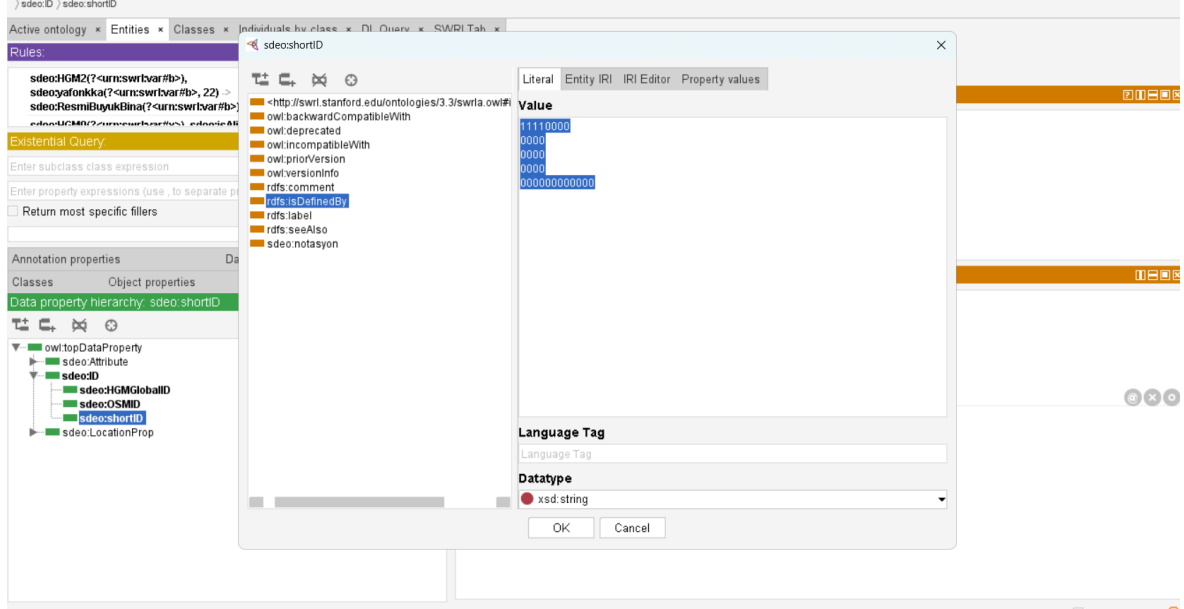
$$\frac{1}{(1 - e^{-M \cdot 2^{(-N)}})}$$

olur. Bu formül karmaşık gözüktüğü de bir M, N çifti için hesaplandığında yakın değerler kolayca tahmin edilebilir. Tüm sınırlayıcı dörtgen için incelenecekse yapı sayısı için M=400 değerinin atanabileceği varsayılmaktadır. Birbirine yakın yapılar değerlendirilecekse M=10 yeterli olacaktır. Deneme sayısı, M ile yaklaşık olarak ters orantılıdır. Bit sayısı bir artınca (örn:15 bitten 16 bite) da kaç deneme olacağının beklentisi takriben iki katına çıkar.

16 bit durumu için hesaplanan değerler M=400 için 164, M=10 için 6550 denemedir. Mesela her hafta kalite değerlendirme uygulaması yapılırsa, tüm yapılar için yaklaşık 3 sene sonra ontolojide sdeo:shortID değerinin hesaplanmasında GlobalID için hangi bitlerin kullanılacağını değiştirmesi gerekecektir.

Uygulamada HGM yapısının OSM deki ilgili yapı ile ilişkilendirilmiş olup olmama ayrımı yapıldığı için 16 bit yerine 15 bit kullanılmaktadır. 15 bit durumu için bu sonuçlar ikiye bölünmelidir, yani bu durumda 18 ay sonrasında sdeo:shortID mask değerini düzenleme gereksinimi olacaktır. Kurum yetkilisi SDEO ontolojisinde bu düzenlemeyi her zaman yapabilir. Birbirine yakın yapılar için risk daha da küçüktür. Yine de uygulamada kontroller yapılacaktır ve duruma göre kullanıcı uyarılacaktır. Burada rdfs:isDefinedBy notasyon özelliği kullanılır. İlk dört bit durumunu Şekil 17 göstermektedir. Burada yapılacak

değişikliklerin sonuçları için ontoloji uzmanı araştırmacı, HGM personeli ile iş birliği içinde incelemeler yapmalıdır. Takriben beş seneden önce gerçekleşmeyecektir.

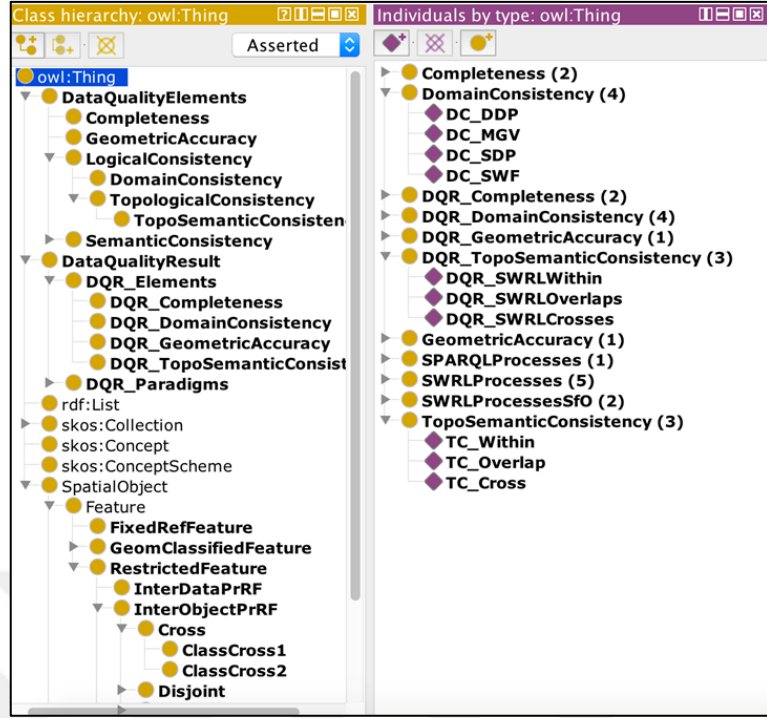


Şekil 17. sdeo:shortID için notasyon özelliği ayarlama.

2.3.2. SDQO ile Kalite Denetimi

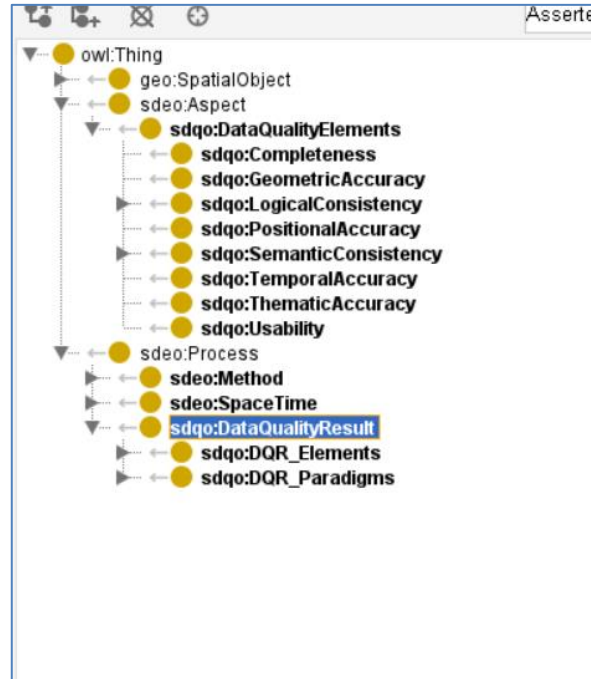
SDQO ontolojisi Yılmaz (2018) tarafından geliştirilmiştir. SDQO ontolojisi, kalite değerlendirmesi ile ilgili gerekli kuralları ve kavramları içerir. Veri kalitesi için uygun ve seçilmiş elemanlara bağlı kurallar setine göre, SDQO ayrıca veri kalitesi elemanlarını ilgili prosedürlerle işlemek, entegre etmekten ve prosedürleri konumsal veriler ile kalite elemanlarına uygun olarak uygulamaktan sorumludur.

SDQO'da, owl:Thing'in hemen altında, biri veri, biri veri kalitesi elemanları ve biri de veri kalitesi sonuçları ve süreçleri için olmak üzere üç üst sınıf vardır. Bunlar sırasıyla ogc:SpatialObject, sdqo:DataQualityElement ve sdqo:DataQualityProcessResult sınıflarıdır.



Şekil 18. SDQO hiyerarşisi.

SDEO ontolojisiinde kullanılan SDQO sınıfları Şekil 19 ile gösterilmiştir.



Şekil 19. SDEO'daki SDQO sınıfları.

Toplanan veriler, SDQO kullanılarak ön değerlendirme sürecinden geçirilir. Bu süreçte, veri setlerindeki topolojik tutarlılık ve geometri tiplerinin doğruluğuna yönelik eksiklikler, tutarsızlıklar, hatalar tespit edilir.

2.4. OpenStreetMap Verilerinin Hazırlanması

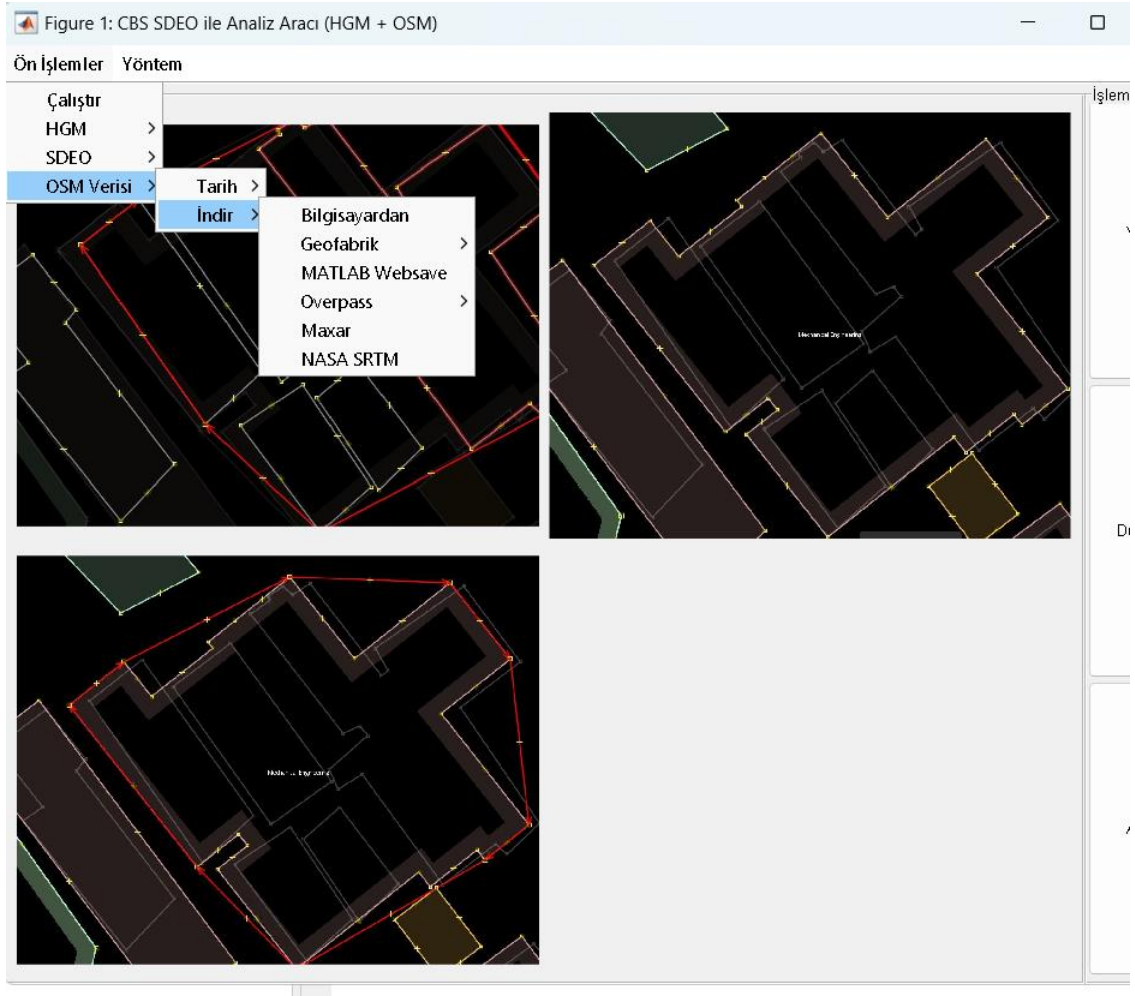
2.4.1. Veri indirme

OSM verisi için belirtilmiş sdeo:BBoxOSM ve bağlı olan veri tipi özelliklerine göre OSM'den veri indirilir. Protégé üzerinden düzenleme ve kontroller yapıldıktan sonra, GUI'de ilgili tercih doğrultusunda indirme işlemi Matlab üzerinden yapılabilir. Overpass Turbo sitesinden OverpassQL komutları kullanılarak da yapılabilir.

Aşağıda örnek bir Overpass sorgusu gösterilmiştir. Hem tavan-arası hem de güncel veri indirilir.

```
[bbox:40.991, 39.771, 40.995, 39.779][date:"2023-10-02T00:00:00Z"];
//[timeout:180];
(
way[!"highway"];> -> .yerler;
//nwr.yerler["entrance=""] -> .bin;
);
// */
out meta;
out count;
```

Şekil 20 ile GUI ve indirme menüsü gösterilmiştir. OSM verisinin hem uygulamada temel olarak kullanılan sdeo:DateOSM'de verilen tarihteki eski verisi (tavan-arası veri), hem de karşılaştırmak ve kullanıcıyı (kurum personeli ya da araştırmacı) bilgilendirmek için güncel hali indirilir. HGM verisine eşlenik tarihi belirten sdeo:DateOSM verisini kullanıcı doğru olarak tutmalıdır. Gerekli düzenleme Protégé ile yapılır.



Şekil 20. OSM verisi indirmek için hazırlanan kullanıcı arayüzü.

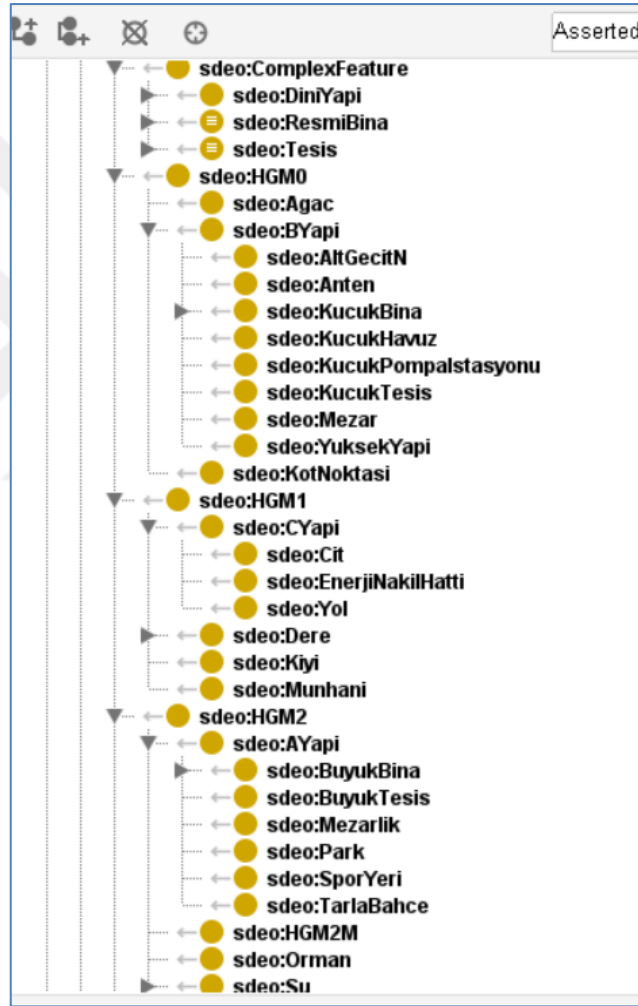
Veri İndirme Uygulamaları ve Servisleri

OSM ile üretilebilecek uygulama ve görselleştirme türleri ne olursa olsun, OSM veri tabanı içindeki verilere erişimi sağlayan uygulama ve hizmetler, OSM'nin veri mimarisinin önemli bir parçasıdır. Geofabrik, OSM verilerine erişimde en iyi bilinen sağlayıcılardan biridir ve kıtasal, ulusal ve bölgesel boyuttaki veri setlerine erişim sağlar; veriler en az saatte bir sıklıkta yüklenir ve farklı formatlarda (*.osm, *.o5m, *.pbf) sunulur.

OSM ayrıca ham verilerin OSM veri tabanından çıkarılmasına ve veri tabanına kaydedilmesine olanak tanıyan bir API de sağlar. Overpass Turbo ile Overpass API hizmeti, OSM harita veri tabanının seçilen bölümlerine erişime izin veren salt okunur bir yöntem sağlar; istemciler sorguları özel bir API sorgulama dili kullanarak veya Overpass Turbo tarafından sağlanan grafik arayüzü kullanarak gönderir.

2.5. HGM Verisinin Ontolojiye ve Dosyalara Aktarılması

Yılmaz (2018), tarafından geliştirilen SDQO ontolojisi kullanılarak kalite incelemesi yapıldıktan sonra çalışma için ortaya konmuş SDEO üzerindeki değişiklikler, Protégé ontoloji editörü ile yapılabilir. Ontoloji üzerindeki HGM sınıfları ve üretilen hiyerarşik yapı Şekil 21 üzerinde gösterilmektedir. Veriler HGM'deki geometrik yapılarına göre sınıflandırılmıştır; sdeo:HGM0, sdeo:HGM1, sdeo:HGM2 ve çoklu poligon tipi veri için sdeo:HGM2M.



Şekil 21. Ontolojide HGM sınıfları hiyerarşisi.

Çalışmada, HGM'den alınan veriler, uygulama için GeoJSON formatına dönüştürülüp kullanılmıştır. Veri dönüşümü için MATLAB kullanılmıştır (GeoJSON); alternatif olarak QGIS kullanılabilmektedir.

HGM veri düzenleme algoritması aşağıda gösterilmiştir.

```

Başla
! shaperead fonksiyonu çalışmıyor
! Geometrik verileri tablo olarak oku:
geoTable ← readgeotable(dosya);
! WKT formatındaki veriyi tabloya ekle:
geoTable ← geotable2table(geoTable, ...);
! Yeni eklenen sütunu "WKTData" olarak adlandır
WKTData ← geotable.WKT;
Bitir;

```

Sıfır ve bir boyutlu HGM verileri koordinatları ve öznitelikleri ile veri ontolojisinde tutulur. Burada OSM benzeri bir sistem devreye sokulmuştur. Çoklu noktalar ve çoklu çizgiler önce tekli parçalara ayrılır. GlobalID değerleri bu parçalara aktarılır, dolayısıyla her parçanın GlobalID değeri aynı olur. sdeo:coeffValue ile çoklu yapıdaki sıra korunur, böylece çoklu yapıdan veri kaybı olmamış olur.

Tekli noktalar için koordinat değerleri sdeo:hasCentroid ile tutulur. GlobalID değerlerinin eşitliğine karşılık gelen anonim sınıflar kurallarda değerlendirilebilir.

Her tür detay için koordinat değerleri, veri ontolojisinde grid yapısında tutulur; katı (hard) çözünürlük değerine göre yuvarlamalar yapılır. Ağırlık merkezi değerleri girildiğinde de direkt enlem boylam değerleri girildiğinde de durum böyledir.

Tekli çizgilerde her doğru parçasının uç noktaları veri ontolojisinde koordinatları ile oluşturulur (N41000000E39000000 gibi). GlobalID bunlara aktarılmaz. OSM çizgilerinde olduğu gibi tekli çizgiler bu düğümlerle tanımlanır (sdeo:partValue veri tipi özelliği). Ayrıca düğümler sdeo:hasNextNode nesne özelliği ile bağlanır. SWRL kuralı ile birden çok düğüme bağlananlar tespit edilir. Ağırlık merkezi değerleri de hesaplanıp saklanır.

Poligon yapıların öznitelik değerleri kimlik değerleri ile veri ontolojisinde tutulurken, katmanın GeoJSON dosyasında kimlik değerleri ve geometrik yapıları tutulur. OSM benzeri sistem burada uygulanmamıştır.

2.5.1. Öznitelik Verileri

HGM verisinin kimlik değer dışında tutulan etiket verileri olarak Şekil 22 ile gösterildiği gibi, isim, tesistipi, topodetays, bikula ve yafonkka seçilmiştir.

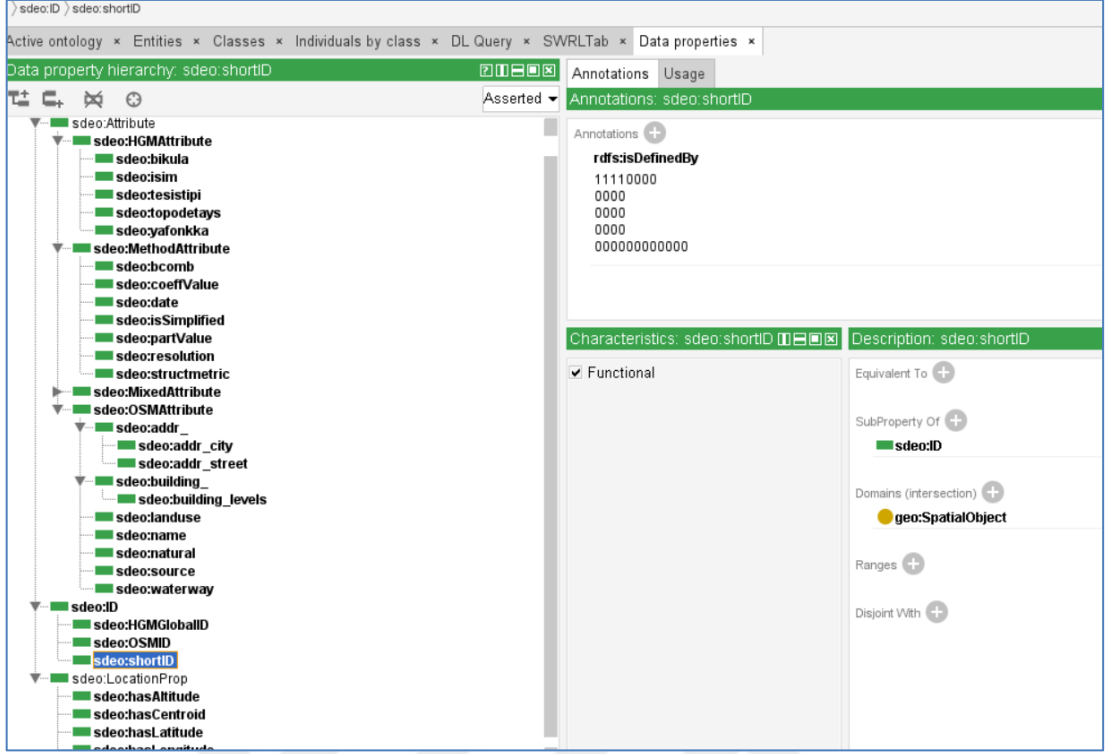
İsim değeri “Bilinmiyor” olanların isimleri boşa “” çevrilmiştir. Kullanıcı bir SWRL kuralı ile HGM deki sdeo:isim değerleri ile OSM deki sdeo:name değerlerini karşılaştırıp kullanabilir.

isim(?x, ?i), name(?y, ?n), isAlike(?x, ?y), stringConcat(?ni, ?n, ", ", ?i) -> nameisim(?x, ?ni)

Sıfır ve bir boyutlu yapılar için ID olarak GlobalID kullanılır. İki boyutlu yapılar için GlobalID değeri kayıtlı tutulur ama shortID değerleri de hesaplanıp kullanılır.

sdeo:shortID, rdfs:isDefinedBy notasyon özelliğinde tanımlanmış maskeye göre hazırlanır. İlişkilendirme olup olmamasına göre bir bit kaybının olacağı, maske düzenlenirken göz önüne alınmalıdır. Burada ilk 16 bitlik kısım seçimine göre açıklama yapılmıştır.

İlk onaltılık basamak 7'den büyükse (8, 9, a, b, c, d, e, f), 8 çıkarılarak 15 bite çevrilir. (0->32767 ilişkilendirme yok) Daha sonra eğer HGM verisi OSM verisi ile ilişkilendirilmişse (sdeo:isAlike), bu değerlere 32768 (2^{15}) eklenir. Zaten ilişkilendirilmiş HGM yapılarının sdeo:shortID değerleri 32768 sayısından büyük-eşit olurken, henüz ilişkilendirilmemişlerin değeri daha küçük olacaktır.



Şekil 22. HGM ve OSM etiket verileri ve sdeo:shortID.

2.5.2. Geometrik Veriler ve Yapı Geometri Tipine Göre Uygulanan Yöntemler

Tez çalışmasının ana odak noktası yapılardır. İki boyutlu verilerin geometrik verileri GeoJSON dosyalarında tutulurken kimlik değerleri (sdeo:shortID) veri ontolojisinde tutulur. sdeo:isAlike nesne özelliği (object property) HGM verileri ile OSM verileri arasında kurulmuş bağlantıyı (ilişkilendirme) gösterir. sdeo:pointToPoly nesne özelliği ise 0 boyutlu yapılar ve 2 boyutlu yapılar arasındaki bağlantı içindir. SWRL kuralı ile 0 boyutlu HGM yapıları, sdeo:isAlike ile eşleştirilmiş OSM 2 boyutlu yapıları (kapalı çizgi) sdeo:pointToPoly ile ilişkilendirilir.

$$\text{sdeo:HGM0(?x)} \wedge \text{sdeo:isAlike(?x, ?y)} \wedge \text{sdeo:ClosedWay(?y)} \rightarrow \text{sdeo:pointToPoly(?x, ?y)}$$

sdeo:BYapi

Sıfır boyutlu yapılar (sdeo:BYapi) için iki yöntem sunulmuştur; sdeo:MCentroid0 ve sdeo:MInPolygon. Kurum personeli ya da araştırmacı tercihi ontoloji üzerinden yapar, önceki tercihinin korur ya da ontolojiyi düzenleyerek değiştirir.

sdeo:MCentroid0 yöntemi OSM'deki ilgili binaları köşelerine göre ağırlık merkezleri ile ilişkilendirir ve bu HGM verisi ile karşılaştırılır. Ağırlık merkezi hesaplanmasında yakın

ya da uzak tüm köşelere eşit ağırlık verilmiştir. Örneğin kenar uzunluklarına bağlı bir hesaplama yapılmamıştır. sdeo:coeffValue eşik değerini belirler.

sdeo:MInPolygon yöntemi sdeo:isAlike nesne özelliği ile ilişkilendirmiş verileri değerlendirebilir. Karşılık gelen OSM verisi de sıfır boyutlu ise sdeo:MCentroid0 yöntemi gibi davranır ama iki boyutlu ise noktaların OSM’de karşılık gelen poligonların yaklaşık olarak içinde olup olmadığına bakar.

sdeo:HGM1 ve sdeo:Yol

Münhani verisi OSM’de NASA verileri ile oluşturulabilir; bunun yerine HGM Münhani verisi ile NASA verisinin direkt karşılaştırılması önerilir. Gelecek çalışmalarda bu üç veri kaynağının birlikte edinimi, değerlendirilmesi ve daha iyi veri üretimi planlanmıştır.

Yol verileri için sdeo:MCentroid1 yöntemi Öklid mesafesi ile sdeo:MCentroid1M yöntemi ise Manhattan mesafesi ile ağırlık merkezleri arasındaki mesafeye bakar.

sdeo:MWhite1 ve sdeo:MWhite1M yöntemlerinde doğu batı doğrultusunda hesaplanan farklar sdeo:coeffValue özelliğinin değeri (örneğin 1,3) ile çarpılır. Whitening tipi bir dönüşüm yapılmış olur, dolayısıyla Mahalanobis tipi bir mesafe hesaplanmış olur. Koordinat sistemi enlem boylam cinsinden olmaya devam etse de mesafe hesaplamaları yeryüzü üzerindeki (metre cinsinden), Trabzon civarında benzemiş olur.

Bu yöntemler için sdeo:isSimplified değeri var ve 1 ise Ramer-Douglas Peucker son nokta uyum algoritması ile çizgiler basitleştirilir. Yöntemler için sdeo:resolution değeri Ramer-Douglas Peucker algoritması için kullanılırken Trabzon için 1 metrenin yaklaşık 0,00001 dereceye tekabül ettiği kayda alınmalıdır. Varsayılan değer de budur.

Xiao-li ve De (2010), optimal eşik değeri için yöntem sunmuştur.

Bu yöntemler çizgileri düğümlere indirgemektedir, örneğin uzunluk bilgisi kullanılmaz. Alternatif bir metrik olarak ağırlık merkezi tabanlı yöntemi genelleştirmek düşünülmüştür. Karşılaştırılan iki çizgi unsuru $2N+1$ (örneğin 11’er) tane eşit uzunlukta parçaya bölündükten sonra düğümler karşılaştırılmıştır. Bu düğümler farklı doğru parçaları üzerinde olabilecektir.

Gelecekteki çalışmalara bırakılan bir başka yöntemde önce orta noktaları aynı olan iki doğru parçası için bir metrik, uzunluk ve eğime bağlı olarak türetilir. Sonra tüm doğru parçalarına, orta nokta arasındaki mesafe kullanılarak genelleştirilir. Fakat çizgiler genelde birden çok doğru parçasından oluşur.

sdeo:HGM2 ve sdeo:HGM2M

sdeo:HGM2 hem sdeo:HGM2M'yi (çoklu poligonal yapı) hem de sdeo:BuyukBina sdeo:Gol gibi HGM bazlı sınıfları kapsar. İlgili sınıflardaki verinin belirleyici değerleri ve diğer etiket verileri veri ontolojisinde tutulur. MATLAB readgeotable ile ilgili tüm 2 boyutlu yapıların iç boşluk sayıları araştırılır. İç boşluğu olmayan yapıların geometrik ve diğer verileri sınıf (katman) değerlerine bağlı olarak ilgili GeoJSON dosyasında saklanır.

Çoklu poligon yapılar içinse ana poligon verisi ve iç poligon sayısı diğer etiket türü veriler ile ilgili GeoJSON dosyalarında tutulur. İç poligonlarının geometrik verileri "icpoligon.geojson" dosyasında poligonun kimlik değeri ile birlikte saklanır. Raster aşamasında bu iç poligonların çevreledikleri alandaki piksel değerleri sıfıra (siyah) çevrilir.

Katmanların GeoJSON dosyalarındaki geometrik veriler tekli poligon verisidir. sdeo:MCentroid2 ve sdeo:MCentroid2M yöntemleri bu ana poligonların ağırlık merkezlerine göre, sırasıyla Öklid Mesafesi ve Manhattan Mesafesi kullanarak karşılaştırma yaparken çalışma için üretilmiş ana yöntem olan SAP (sdeo:MSAP) yöntemi daha ayrıntılı bir karşılaştırma yapar. SAP hem HGM verisinde hem de OSM verisinde poligon (kapalı yol) tipindeki yapıların karşılaştırılması için tez çalışması kapsamında ortaya konulmuş bir yöntemdir. Bu yöntemin kullanılıp kullanılmaması ve temel parametreleri SDEO üzerinden Protégé ile düzenlenir.

2.5.3. OSM ve HGM Verilerinin Karşılaştırılması İçin Ön İşlemler

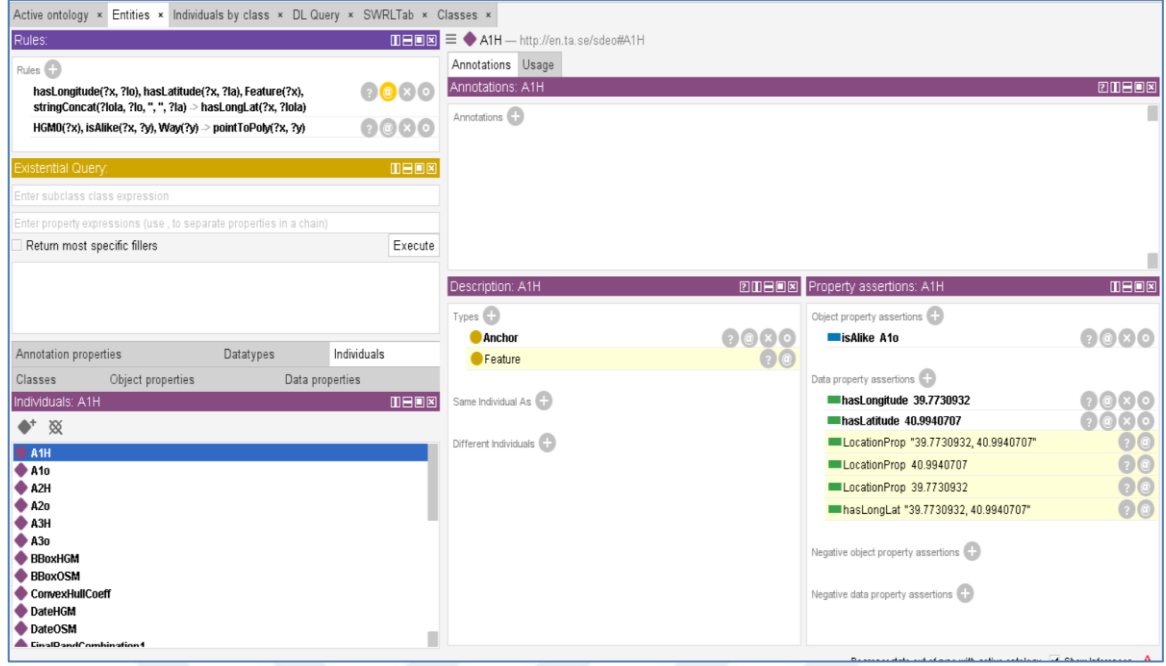
OSM ve HGM verilerinin karşılaştırılması ve ilişkilendirilmelerin yapılmasından önce OSM verisine dönüşüm uygulanır ve HGM verisine daha çok benzeştirilir. Dönüşümler hem global olarak tüm veriye birden, hem de lokal olarak uygulanır. Bunun için çapa (anchor) noktaları kullanılır. Bu noktalar koordinat dönüşümleri için kullanılan kontrol noktalarıdır.

Global dönüşümler için, dört farklı seçenek vardır.

- Öteleme
 - OSM çapa noktaları koordinat ortalamaları HGM'ninkilere eşit olacak şekilde öteleme yapılır.
- Afin Dönüşüm
 - İlişkilendirilmiş çapa noktaları arasında dönüşüm matrisi en küçük kareler yöntemine göre hesaplanır ve tüm veriye uygulanır.
- Tanjantlı Afin Dönüşüm

- Boylam değerleri ikiye bölünür; sonra hem enlem hem boylam değerlerinin tanjantı alınarak reel düzleme varılır. Reel düzlemde afin dönüşümü yapılır. Arctan fonksiyonu ile geri dönülür; sonrasında boylama karşılık gelen değerler iki ile çarpılır.
- Homeomorfizm (diffeomorfizm, tersi olan spline)
 - Bu seçenek, önceki seçenekleri genelleştirir. Onlarda olduğu gibi bir homeomorfizm (tersi sürekli olan sürekli fonksiyon) seçilir ve uygulanır. Ontolojide anonim fonksiyon şeklinde yazılır. Örneğin “@(x,y) ($x+1, y+1$)”
 - Türevin tanımlandığı yerlerde türev sıfır olamamalı.
 - Monoton fonksiyonlar tercihte önceliklidir.

Ontolojide çapa noktaları sdeo:Anchor sınıfında tutulur. Bir HGM sınıfı bu sınıfın alt sınıfı yapılırsa o sınıf elemanlarından ilişkilendirilmiş olanları çapa noktası olur. sdeo:Anchor sınıfındaki ya da alt sınıflarındaki detayların sadece sdeo:isAlike ile bir OSM (sdeo:OSMEqFeature) detayı ile ilişkilendirilmiş olanları çapa noktalarıdır. Uygulamada olası alt sınıflardakilere (cami gibi) ek olarak üçer tane eşlenik çapa noktası kullanılmıştır. Şekil 23 ile sdeo:A1H gösterilmiştir.



Şekil 23. Çapa noktası A1H ve bilgileri.

Çapa noktalarının koordinat değerleri kullanıcı tarafından sağlanır ve buna göre OSM verisinde, değişiklikler bir GeoJSON dosyasında saklı tutularak, öteleme, döndürme, afin dönüşüm gibi işlemler yapılabilir. HGM vektör verisi sabit tutulur. Hiç çapa noktası yoksa dönüşüm gerçekleşmez.

Global dönüşüm için, aşağıdaki RDF üçlülerinden biri üretilebilir (uygun f ve g fonksiyonları için ve yöntem SAP olmak zorunda değil).

sdeo:MSAP sdeo:coeffValue “affin”,
sdeo:MSAP sdeo:coeffValue “translation”,
sdeo:MSAP sdeo:coeffValue “tanaffin”,
sdeo:MSAP sdeo:coeffValue “@(x,y) ($f(x,y)$, $g(x,y)$)”

Homeomorfizm seçeneklerinde MATLAB anonim fonksiyonları tanımlanır ve başta @ işareti olur. Mevcut uygulamada üçüncü RDF üçlüsü (tanaffin) eklenmiş ve afin dönüşümden önce ve sonra trigonometrik fonksiyonlar uygulanmıştır.

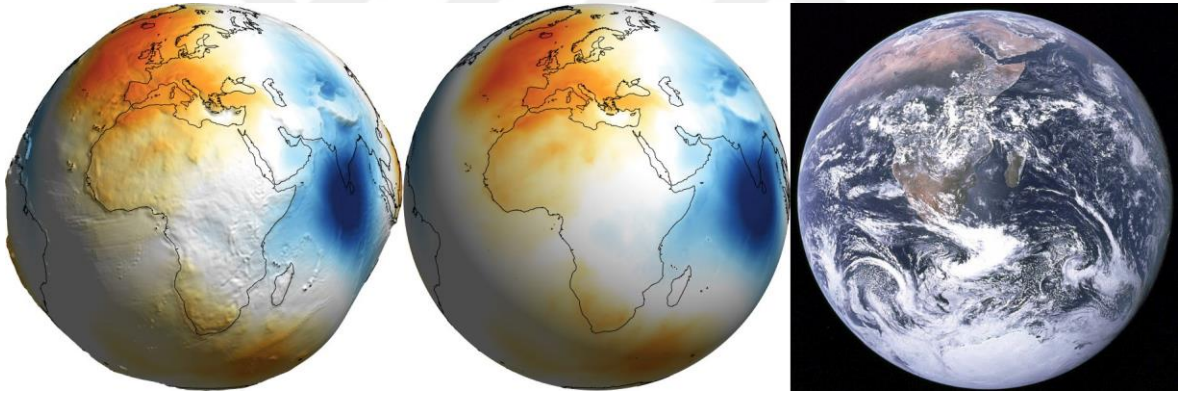
360 tane boylam ve 180 tane enlem vardır. 179. batı boylamından 2 derece batıya gidilince 181. batı boylamına değil 179. doğu boylamına ulaşılır.

Reel düzleme ulaşmak için boylam değerlerinin yarısının ve enlem değerlerinin tanjant değerleri alınabilir. Tanjant fonksiyonu bir diffeomorfizmdir, her yerde türevi vardır,

tersi vardır (arctan) ve tersinin de türevi vardır. Radyan değil de derece tabanlı tanjant fonksiyonu seçildiği için -90 derece ile 90 derece arasında tanımlanır.

Yerkürede negatif dereceler batıya ve güneye karşılık gelmektedir. Boylam derecesi ikiye bölününce istenilen tanım kümelerine ulaşılmış olur. WGS84 elipsoidin basıklık değeri 1/300 gibi küçük bir sayıdır. Çalışmadaki bölgelerin küçüklükleri de düşünülerek, jeoid ya da elipsoid yerine lokal olarak küresel düşünülmüştür (düzlemsel değil). Haversine formülü böylece geçerli olur. Verilen trigonometrik fonksiyonlar bu yerküreyi reel düzleme aktarır. Afin dönüşüm reel düzlem içindir. Reel düzlemde dönüşüm yapıldıktan sonra yerküreye geri dönülür.

Bezděk ve Sebera (2013), Potsdam patatesi için bir betik yazmış, bu betiğin çeşitli abartı faktörleri ile elde edilen sonuçları Şekil 24 ile gösterilmiştir. Soldaki görüntüde abartı faktörü 6161, ortadaki görüntü yerküre ve sağdaki görüntü ise “mavi misket”tir (Wuebbles, 2012).



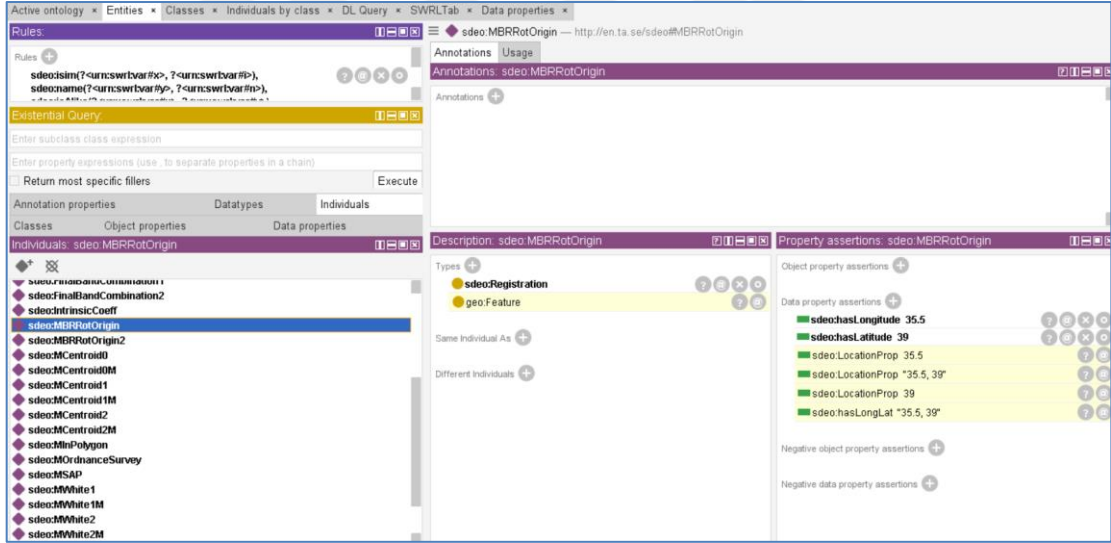
Şekil 24. Potsdam patatesleri ve dünya.

Lokal dönüşüm için farklı çapa noktası seti kullanılmaz, global dönüşüm için kullanılan çapa noktaları kullanılır. Her bir yöntem için en fazla bir tane global dönüşüm ve en fazla bir tane aynı tip (öteleme veya döndürme) lokal dönüşüm RDF üçlüsü SDEO ontolojisinde yer almalıdır. Lokal dönüşüm seçenekleri aşağıdaki gibidir.

- Öteleme (sdeo:coeffValue yüklemi ile)
 - “localtranslation” en yakın çapa noktası çiftine göre
 - “localtranslation2” sondaki sayı çapa noktası çifti sayısıdır
 - “localtranslationw3” en yakın 3 çapa noktası, uzaklığa göre ağırlıklı
- Döndürme

- Seçilen döndürme ögesi yönteme sdeo:hasPart ile bağlanır.
 - sdeo:MSAP sdeo:hasPart sdeo:MBRRotOrigin2
 - sdeo:MSAP sdeo:hasPart sdeo:RotAroundCenter.
- Döndürme ögesi parametreleri tanımlanır.
 - sdeo:MBRRotOrigin2 sdeo:hasLongLat “41,39”.
 - sdeo:RotAroundRegionCenter sdeo:coeffValue “0.1”.
- Ölçekleme değeri 1 alınır.

Şekil 25 ile gösterildiği gibi sdeo:MBRRotOrigin için enlem boylam değerleri verilmişse (sdeo:hasLongLat), bu nokta etrafında dönüklük yapmak için sınırlayıcı dikdörtgenler (minimum bounding rectangles) ve ortalama doğrultuları hesaplanır. Var olan SWRL kuralı sayesinde, eğer enlem ve boylam değerleri ayrı ayrı verilmişse, sdeo:hasLongLat değeri de bir çıkarsama sonucu oluşturulmuş olur.



Şekil 25. sdeo:MBRRotOrigin.

sdeo:RotAroundRegionCenter ile sınırlayıcı dörtgen (bbox) ağırlık merkezi (orta noktası) etrafında verilen derece kadar döndürülür. Bu dönüşümler ile orijinal OSM verisinin, HGM verisine yaklaştırılması hedefi gerçekleştirilmiştir.

Uygulamada yapılan dönüşümde, dönüklük değerleri 0 dereceye ve global ölçekleme değerleri de 1'e yakın çıkmıştır. Poligonların köşe noktalarının yerleşimlerinin uyumlu olduğu görülmüştür. Örneğin, beşgen bir binanın en kısa kenarı HGM verisinde güneye denk geliyorsa, dönüştürülmüş OSM verisinde de güneye denk gelir. SAP (sdeo:MSAP)

yönteminde bu varsayımla hareket edilir, yani homolog denilen düğümlerin yönelimlerinin örtüşmesi beklenir. Örneğin dışbükey kabukta da olan düğümler için, bir HGM poligon dışbükey kabuğundaki saatin tersi yöndeki hareket, OSM kapalı çizgi dışbükey kabuğundaki saatin tersi yöndeki hareketle bağdaşır.

sdeo:MSAP yöntemi yoğun kentsel bölgeler düşünülerek tasarlanmıştır, yapıların seyrek dağıldığı bölgelerde de çalışır. Fakat böyle bölgelerde;

sdeo:MSAP sdeo:coeffValue “docentroidandstop”

RDF üçlüsünü koymak tercih edilebilir, böylece dönüşümden sonraki işlemler yapılmaz ve ilişkilendirmeler hemen gerçekleştirilir.

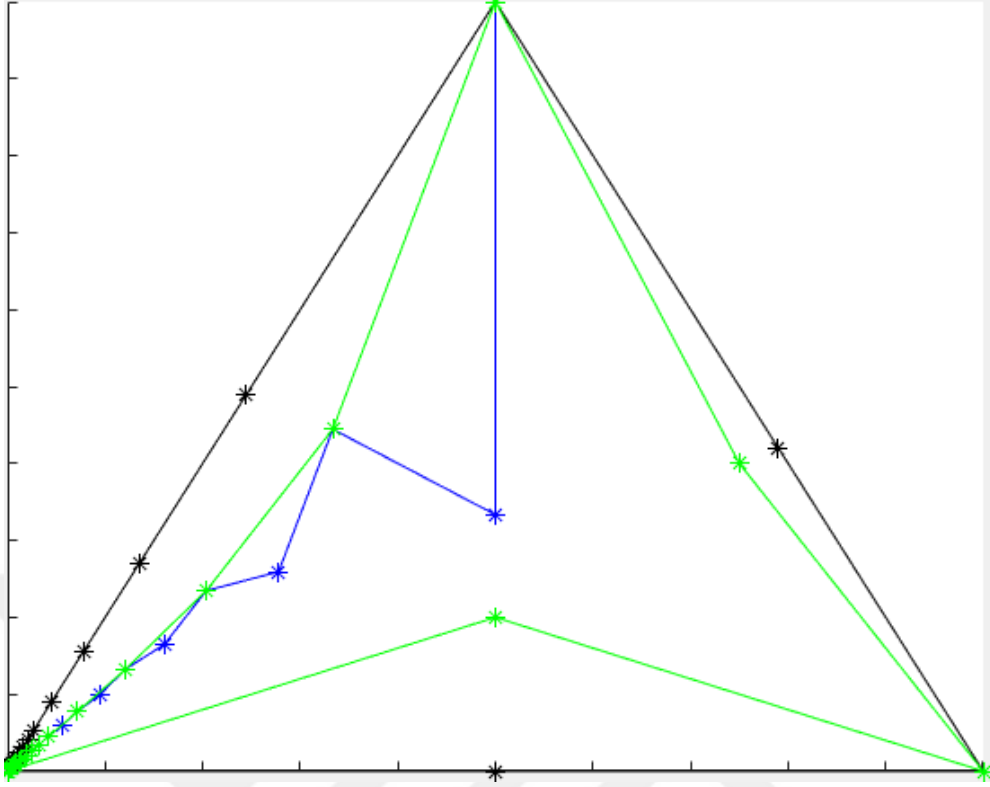
GUI’de yetkili kurum personeli dönüşümden sonra kullanacağı yöntemi, boyutuna göre seçer. Seçeneklerden biri iki boyutlu yapılar için SAP yöntemidir. Ayrıca tüm işlemleri varsayılan seçeneklerle peş peşe yapma olanağı GUI’de verilmiştir.

2.5.4. SAP Yöntemi

SAP yönteminde diğer yöntemlerde olduğu gibi önce SDQO test ve dönüşüm işlemleri yapılır. Homolog düğümlerin yerleşimlerinin benzeşmesi sağlanır.

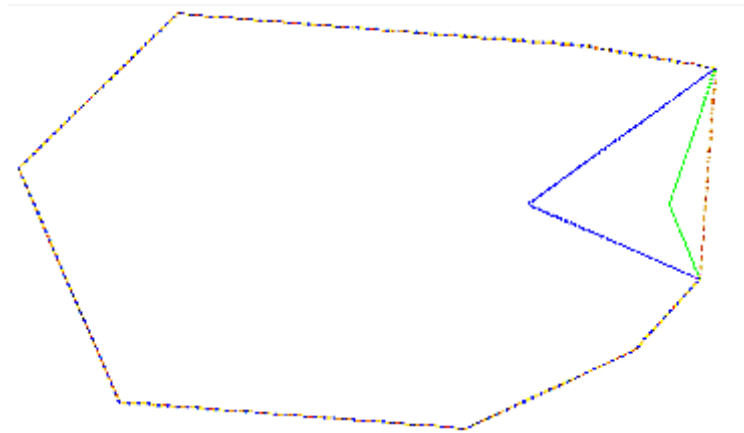
Bu yöntemde, OSM kapalı çizgileri ve HGM poligonları karşılaştırılıp ilişkili olanlar tespit edilir. Bunun için dışbükey kabuklar çok genel bulunmuştur ve ara poligonlar incelenmiştir. Bu poligonların köşeleri, aynı dışbükey kabukta olduğu gibi, orijinal poligonun köşeleri arasındadır. Ara poligon için köşelerin elenmesi, dışbükey kabuğun dışına taşılmayacak şekildedir. İç açılara ve “komşu” dışbükey kabuk köşeleri ile yapılan üçgenlerin alanlarına bakılır. Bayazit (2009) ve Keil ve Snoeyink (2002) dışbükey alt çokgenlere ayırma algoritmalarına, ayrıca Alpha Shape işlemlerine, ontoloji parametrelerine göre bakılır (Edelsbrunner, 2011).

Şekil 26 ile bir ara poligon örneği çokgenlerle gösterilmektedir. Dışbükey kabuk siyah üçgendir. Tüm üçgenler dışbükeydir (üçgenlere ayırma bir dışbükey alt çokgenlere ayırma işlemidir). Siyah üçgenin tüm köşeleri aynı zamanda yeşil çokgenin (ara poligon) köşesidir. Yeşil çokgenin tüm köşeleri aynı zamanda mavi orijinal çokgenin köşeleridir. Orijin etrafında mavi çokgenin çok (ama sonlu) sayıda köşesi vardır. Siyah yıldızla gösterilmiş noktalar, yeşil çokgen köşelerinin üçgen kenarlarına en yakın olduğu noktalardır (kenara düşey inme).



Şekil 26. Ara poligon örneği.

İç açı ve ilgili üçgen arasındaki ilişki Şekil 27 ile gösterilmiştir. Mavi ve yeşil poligonların dışbükey kabukları aynıdır (kırmızı poligon), hatta sadece bir tane içbükeylik köşesine sahiptir ve bu köşeler aynı dışbükey kabuk (kesikli kırmızı poligon) köşeleri ile komşudur. Dolayısıyla dışbükey kabuk üzerindeki kenar üçgenlerde ortaktır. Üçgenin alanı yüksekliğe, bu da iç açıya ve noktanın konumuna bağlıdır.

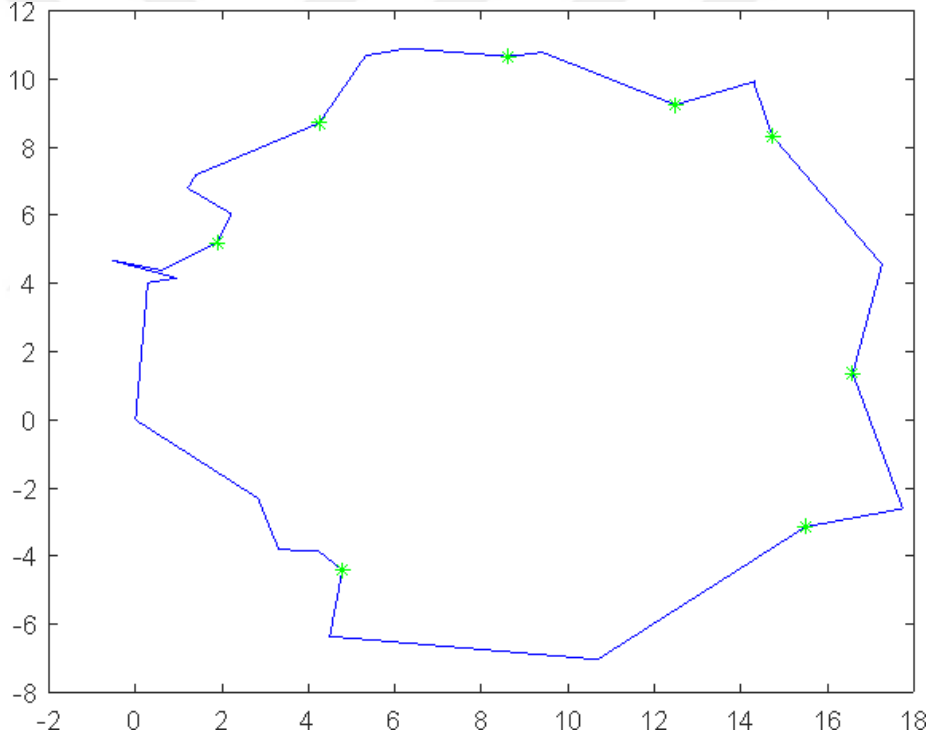


Şekil 27. İçbükey köşede iç açı ve alan ilişkisi.

Bu ara poligonların bazı köşelerinin konumları değiştirilerek, SDEO ontolojisi güdümünde, SAP poligonları üretilmiştir. Dönüşüm seçenekleri ve diğer seçenekler için öznesi sdeo:MSAP, yüklemi sdeo:hasPart nesne özelliği olan tüm üçlüler ve çıkarsama sonrasında öznesi sdeo:MSAP ve yüklemi sdeo:MethodAttribute olan üçlüler incelenir.

sdeo:MethodAttribute, yöntemlerle ilgili veri tipi özelliklerinin üst özelliğidir. Çıkarsama öncesinde yüklemi sdeo:MethodAttribute olan bir üçlü konulmamıştır. Ontolojide bir üçlü varsa, yüklem üst özelliklerle değiştirilirse çıkarsamada o üçlüler de eklenir. Üretilen noktalar Steiner (1971) noktaları ile benzeşir, ayrıca içbükeyliği azaltır (Lien ve Amato, 2004).

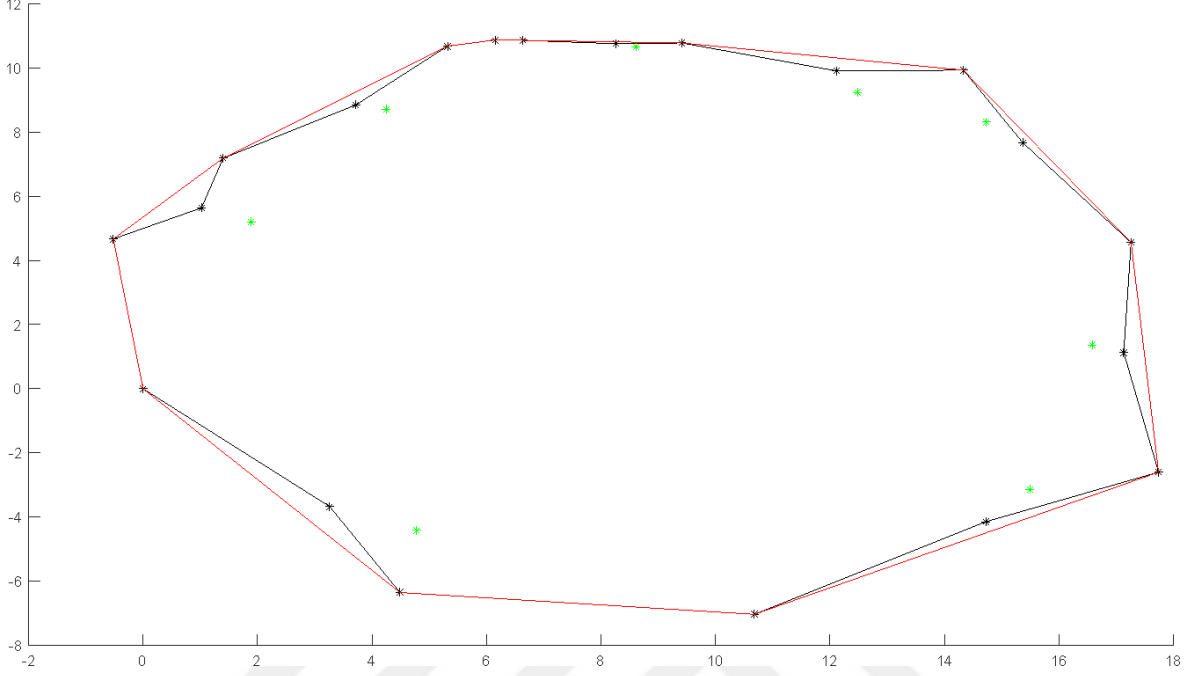
Bir poligon ve ara köşeleri Şekil 28 ile gösterilmiştir. Aynı poligon için SAP ve dışbükey kabuk ara köşelerle beraber Şekil 29 ile gösterilmiştir.



Şekil 28. Poligon (mavi) ve ara köşeler (yeşil yıldız).

Orijinal poligonun dışbükey kabuğu hem ara poligonu hem de SAP poligonunu kapsar, dolayısıyla hepsinin dışbükey kabuğu aynıdır ve eğer orijinal poligon dışbükey ise hepsi birbirine eşittir; dışbükey poligonlar için SAP poligonu kendisine eşittir. Her durumda, OSM ve HGM poligonlarının kendileri değil de SAP poligonları karşılaştırılır. Burada “poligon

yakınlığı” kavramı devreye girmektedir. sdeo:ConvexHullCoeff SAP poligonunun dışbükey kabuğa ne kadar yakın olacağını belirler.



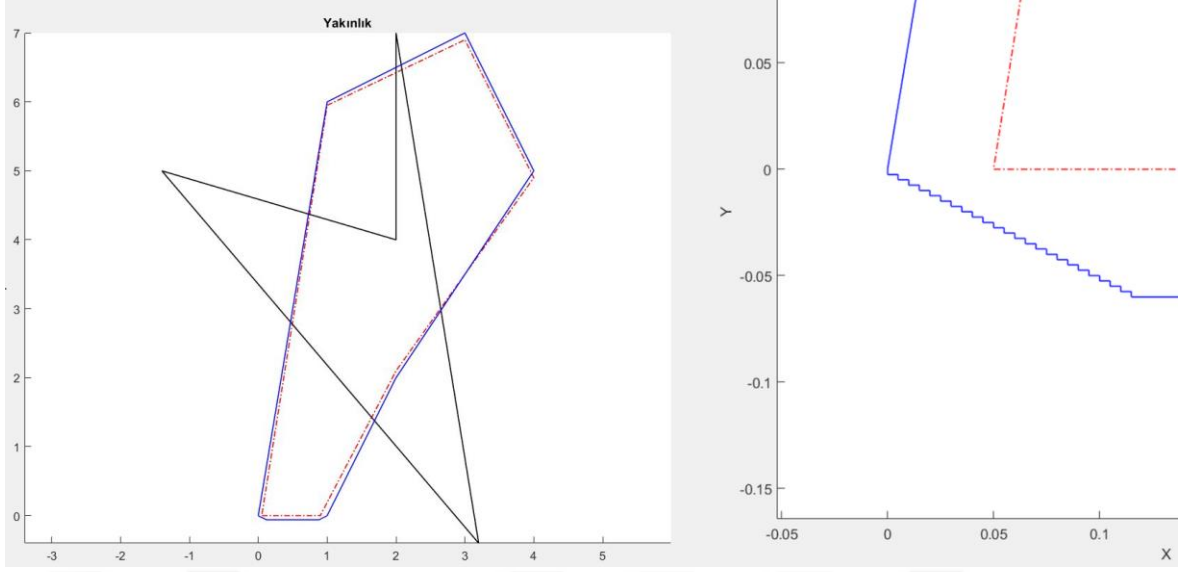
Şekil 29. SAP (siyah), dışbükey kabuk (kırmızı) ve ara köşeler (yeşil).

Poligon Yakınlığı

Eşdeğer iki yapının poligonlarında ağırlık merkezleri çok farklı yerlerde olabilir (dolayısıyla ağırlık merkezi yöntemleri yanlış çalışabilir), köşe sayıları farklı olabilir, kesişme varsa kesişen nokta sayısının az ya da çok olması gerektiği konusunda yorum yapılamayabilir. Bu vektörel bilgiler hangi çiftlerin eşdeğer olduğunu bulma konusunda yardımcı olmayabilir. Gene de bir insan örneğin uydu görüntülerinde bunlara baktığında, kolayca bunların eşdeğer olup olmadığına karar verir.

SAP yöntemi geliştirilirken bu “poligon yakınlığı” düşünülmüştür. Poligonların dışbükey kabuklarını karşılaştırmak faydalı olur ama birçok zaman fazla genelleme yapar. Poligonları karşılaştırmak yerine onların (sanal) ara poligon versiyonları karşılaştırılır. Bu (sanal) ara poligonların tasarımındaki hedef, ancak ve ancak eşdeğer yapıların poligonları için bunların benzeşmesidir.

Şekil 30 örneğinde kırmızı poligon ile mavi poligonun benzeştiği düşünülür. Halbuki siyah poligon; ağırlık merkezi, köşe sayısı, kesişme sayısı açılarından kırmızı poligona yakındır. Mavi poligonun sol alt köşesinde çok sayıda “basamak” vardır.



Şekil 30. Poligon yakınlığı.

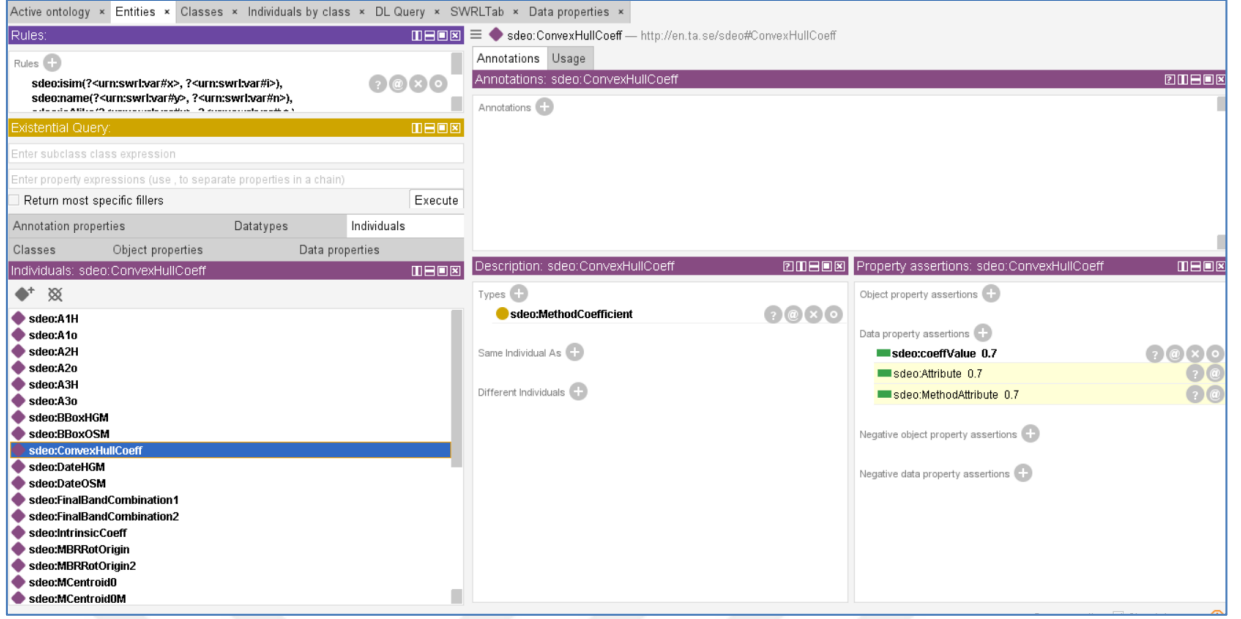
sdeo:ConvexHullCoeff için Katsayı İlk Tahmini

Uygulamada MATLAB’da rastgele olarak on binlerce sayıda, kenarları çok kısa olmayan ama muhtemelen eşkenar da olmayan dışbükey çokgenler oluşturulmuştur. Birer köşenin yeri, çokgen içbükey hale gelecek şekilde ama bunun dışında rastgele bir şekilde değiştirilmiştir. Dolayısıyla tek köşesinden dolayı dışbükeyliğini yitirmiş on binlerce çokgen oluşturulmuştur. OSM ve HGM poligonları böyle çokgenler tarafından temsil edilebilir. Sonrasında mesafe matrisleri hem çokgenler hem de dışbükey kabuklar için hesaplanmıştır.

Hedef, çokgenler arasındaki “toplam mesafe” ile dışbükey kabuklar arasındaki “toplam mesafe” arasında bir ilişkiyi tabir etmektir. Bunun için konuya bir de “iç mesafe” eklenmiştir, bu iç mesafe çokgenin kendisi ile dışbükey kabuğu arasındaki mesafeyi tanımlar. Her çokgen için iç mesafe bellidir, tekrar tekrar hesaplanmasına gerek yoktur. Dış mesafe yani dışbükey kabuklar arasındaki mesafeler hesaplanırsa, geriye gerçek mesafe, iç mesafe ve dış mesafe arasında bir bağıntı bulmak kalır. Bu bağıntının dışbükey lineer bir bağıntı olduğu varsayılmıştır ve regresyon yapılmıştır.

Bu işlem çokgen köşe sayıları değiştirilerek tekrarlanmıştır. Buradan sdeo:ConvexHullCoeff için ilk tahmin ortaya konmuştur. Kurum çalışanı ya da araştırmacı, ontoloji editöründe bu katsayıyı (sdeo:ConvexHullCoeff) düzenleyip ince ayar yapabilir. Şekil 31 ile gösterilmiştir.

Burada seçilen değerlerle dışbükey kabukla yapılan üçgenin (cep üçgen) ağırlık merkezine yaklaşılmıştır.



Şekil 31. sdeo:ConvexHullCoeff.

Rasterizasyon ve Eşleniklik

Düzenlemeler yapıldıktan sonra verilerin raster eşlenikleri sdeo:Rasterization'a bağlı olan sdeo:resolution parametresi kullanılarak derece biriminde oluşturulur. Trabzon civarında 0.0001 derece, doğu batı doğrultusunda 8,5 metreye, kuzey güney doğrultusunda 11,1 metreye karşılık gelmektedir (oran 1:1,3). Yuvarlayarak 1 metre 0.00001 dereceye karşılık geliyor denebilir. Rasterizasyon görsel olarak, vektör verideki poligonların içlerinin doldurulmasına denktir, burada Hormann ve Agathos (2001) algoritması kullanılmaktadır.

Raster görüntüde içi dolu sanal ara poligonlar vardır ve Raster görüntü uint16 formatındadır. HGM verisinin raster halinde, gri değer olarak sdeo:shortID kullanılır. HGM verisindeki çakışmalar bu aşamada da tespit edilir.

OSM eşleniği olup olmadığı, sdeo:shortID değerinin 32768'den (2^{15}) büyük olup olmadığına bakarak anlaşılabilir; henüz eşleniği olmayanların değerleri küçüktür.

Daha önceki işlemlerle tavan-arası OSM verisi, HGM verilerine yaklaştırılmıştı. Rasterizasyon sonrası da birbirlerine yakın olurlar, dolayısıyla eşlenik olanlar için SSIM ya da IW-SSIM gibi yapısal metrikler uygulanırsa, yüksek değerlerin çıkması beklenir. Eşleniklik, yapısal metrik değerlerinin yüksekliğine göre belirlenmektedir.

Aşağıda sdeo:Rasterization adlı ontoloji ögesi için mevcut RDF üçlülere gösterilmektedir.

```
sdeo:Rasterization rdf:type owl:NamedIndividual ,
    sdeo:RasterOperations ;
sdeo:coeffValue 0.9 ;
sdeo:resolution 0.0001 ;
sdeo:structmetric "IWSSIM" .
```

Eşlenik yapılar tespit edildikten sonra, sdeo:isAlike ile ilişkilendirilme yapılır ve eğer sdeo:shortID değeri küçük ise 32768 eklenir. Daha önce bir eşleştirmeden dolayı OSM poligonuna bir değer atanmamışsa, dolayısıyla orada değerler sıfır ya da bir ise, bu iki katına çıkarılmış sdeo:shortID değerleri OSM poligonuna da atılır. Birden fazla kere eşlenmiş OSM poligonunun gri değeri sadece bir tanesinin sdeo:shortID değerine eşittir.

Gösterim ve Denetleme

Raster görüntüler ayrıca tek bir görüntünün bantları olarak kaydedilir. Bu görüntü de ontolojide belirlenen bant kombinasyonlarına göre (sdeo:FinalBandCombination1) kullanıcıya gösterilir (Bantlar bilgisayara kaydedildiği (*.tif formatında) için kurum personeli ayrıca bunları inceleyebilir). Renkli bir görüntü oluşur. Bu görüntüdeki renk değerlerine göre kullanıcı HGM verisindeki eksik binaları ve diğer durumları görebilir.

Birden çok eşleniği olan HGM yapılarının değeri Bant 3'te sıfırlanır. Bant 5 için tavan-arası OSM verisine uygulanan işlemler güncel OSM verisine, tavan-arası OSM'nin parametrelerine göre uygulanır; bu işlemlerde sdeo:OSMID (örneğin "way\132928025") değişmez. Bant 6 için Bant 5'teki eşlenikli veriler karartılır. Kurum yetkilisi altıncı banta bakarak OSM'de güncellemeye karar verebilir.

Bant kombinasyonu "1,1,2" seçilirse, aynı OSM yapısına eşleştirilmiş HGM verileri görsel olarak, farklı renklerinden dolayı görülebilir. Konumsal olarak yakın olan HGM yapılarının sdeo:shortID değerleri o kadar da yakın olmayacaktır.

Altı bant vardır ve aynı anda üçer tanesi gösterilebildiği için 216 (6³) farklı kombinasyon vardır. Uydu ya da hava görüntüleri, başka kurumların haritaları, SAR görüntüleri kullanılarak zenginleştirme yapılabilir.

Bantlar

Bant 1 = HGM Raster

Bant 2 = OSM Raster

Bant 3 = Eşleniklik HGM Raster

Bant 4 = SDQO ve çakışmalarda kusurlu görülmüş HGM Raster

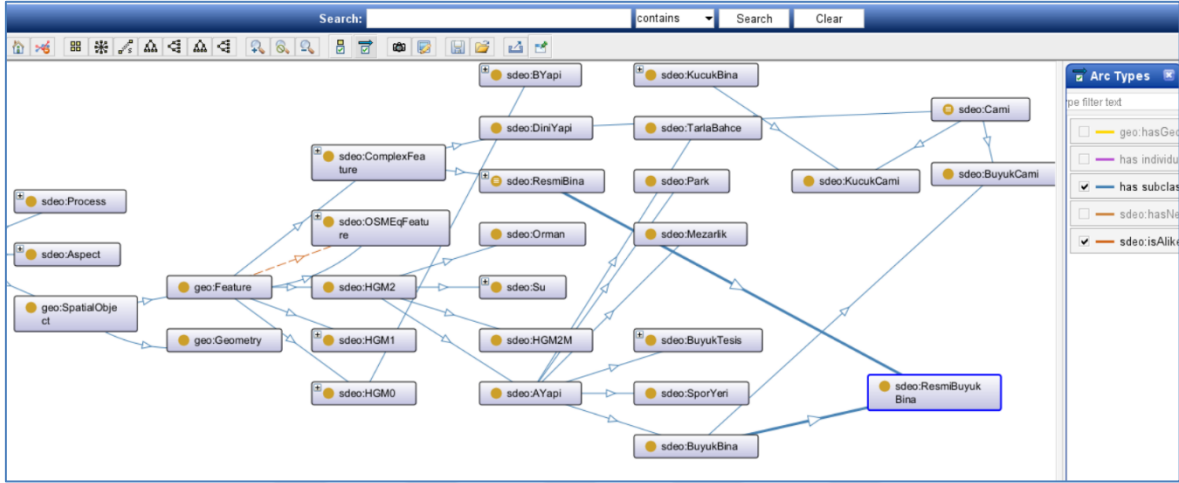
Bant 5 = Güncel OSM Raster

Bant 6 = Eşleniklik Güncel OSM Raster



3. BULGULAR VE TARTIŞMA

Tez çalışmasında üretilen ontoloji tabanlı veri kalitesi denetleme ve iyileştirme metodolojisinin uygulama çalışması Trabzon verisi üzerinde yapılmıştır. Çalışma SDEO ontolojisi güdümünde ilerlemiştir. SDEO sınıfları için Şekil 32 incelenebilir.



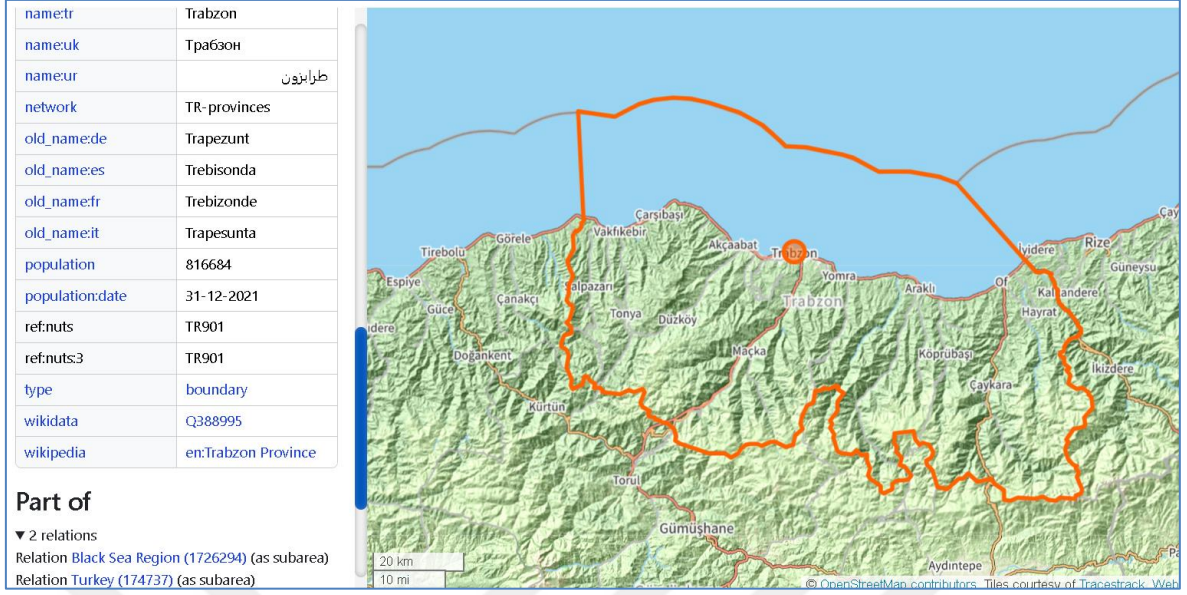
Şekil 32. sdeo:HGM2 sınıfı ve ilişkileri OntoGraf görüntüsü.

Trabzon bölgesi Şekil 33 ile gösterilmiştir. Ontoloji üzerinde tercih edilen çözünürlük değeri sdeo:resolution 0,0001 derecedir. Uygulama için bir de katı çözünürlük değeri vardır, bu değer ile tüm koordinatlar yuvarlanır ve grid yapısı oluşur.

3.1. Çalışmanın Alanlara Uygulanabilirliği Ve Hata Tespiti

Çalışma kutuplar ve 180. boylam yakınında olmayan kara-yoğun, yapılaşmış bölgelerde ontolojide uyarlamalar yapılarak gerçekleştirilebilir. Bölgenin boyutunun seçilen çözünürlüğe oranı önemlidir.

Seçilen çözünürlük değerinde (0,0001 derece), kuzey-güney-doğu-batı doğrultusundaki kare şeklinde binalar enlem boylam coğrafi koordinat sisteminde (WGS84), Trabzon'da 11,12'ye 8,55'lik gibi bir oranla (yani 1,3:1) dikdörtgen şeklinde haritalanırlar. Bu oranın bölge üzerinde sabit kalacağı varsayımının geçerli olabilmesi önemlidir. Benzer olarak, daire şeklinde binalar (ya da merkezî bir noktanın etrafı) da elips şeklini alacaktır. Elipslerin aynı basıklıkta kalması istenmektedir. Bölgeyi küçülterek ya da çözünürlüğü kötüleştirerek bu sağlanabilir.



Şekil 33. 223447 nolu OSM ilişkisi, Trabzon..

Trabzon, denize sınırı olan dereleri de olan ama insan yapısının çokça olduğu kara-yoğun bir bölgedir. OSM verilerine göre enlem boylam sınırları (sınırlayıcı dikdörtgen) aşağıdaki sorgu sonucuna göre (bb) şöyledir.

minlat="40.5135382" minlon="39.1233252" maxlat="41.3060000"
maxlon="40.5077656"

Sorgu:

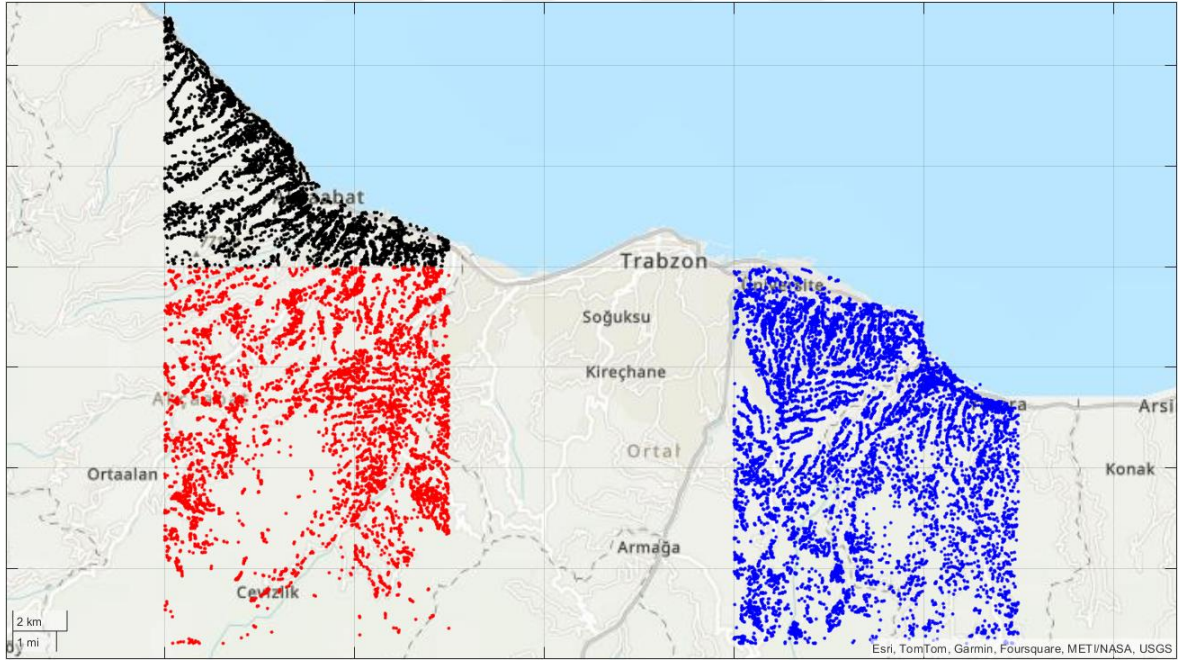
```
area[name="Trabzon"];
( (relation[admin_level=4](area);
  // >;
););
out count;
out bb qt;
```

Ayrıca, OSM verilerine göre building sayısı 60 binden fazladır, tamlık açısından iyi durumdadır. Trabzon HGM bölgelerinde büyük bina katmanlarında toplamda 2200'den fazla detay vardır. Küçük binalarla (hasarlı vs. çıkarıldıktan sonra da) birlikte, bu sayı 100 bini aşar.

OSM Trabzon alanında olan unsurların başka hangi alanlar içinde olduğuna bakılmıştır. Beklendiği gibi Türkiye, Anadolu ve Karadeniz Bölgesi alanları çıkmıştır. Bunların yanında bazı sıradağ alanları, Giresun ve Bayburt illeri çıkmıştır, dolayısıyla OSM'de hem Trabzon'da hem de bu diğer illerde olduğu belirtilmiş düğümler, çizgiler vardır. Sınır düğümleri için bu normaldir. Ortak unsurlar arandığında 307180424 OSM çizgi

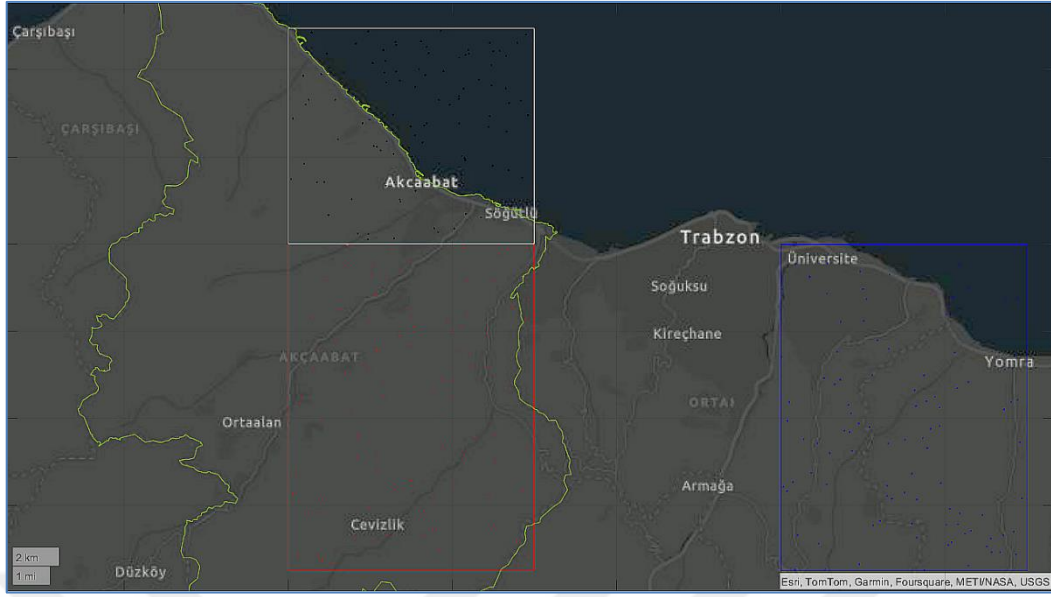
id'li Eynesil'e biraz sarkmış olan Beşikdüzü sahil şeridi (Mapbox kaynaklı) ve 189264645 OSM çizgi id'li Beşikdüzü'ne biraz sarkmış olan Eynesil sahil şeridi (Landsat kaynaklı) bulunmuştur. 2014 yılında 2649699309 OSM düğümü bu çizgilerin düğümü olmuştur. Bunun hata olup olmadığı yetkili kurum personeli ve araştırmacı tarafından kararlaştırılmalıdır. Bu, tez çalışmasında hata olarak görülmemektedir. Düğüm aynı zamanda sınır düğümüdür. Bu işlem Overpass sorguları ile yapılmıştır, ayrıca veri kökenlerinin (lineage) görülmesine bir örnektir.

HGM verisinde Trabzon 23 bölgeye ayrılmıştır. Bunlardan G43-b1 paftası Karadeniz Teknik Üniversitesi'ne, F43-d4 paftası Akçaabat merkezine, G43-a1 paftası da Akçaabat kırsalına denk gelmektedir. Küçük bina dağılımı Şekil 34 ile gösterilmiştir. İstatistikler için Akçaabat ve Trabzon OSM alanlarına da bakılmış, ama yöntemler bu üç bölgede çalıştırılmıştır.



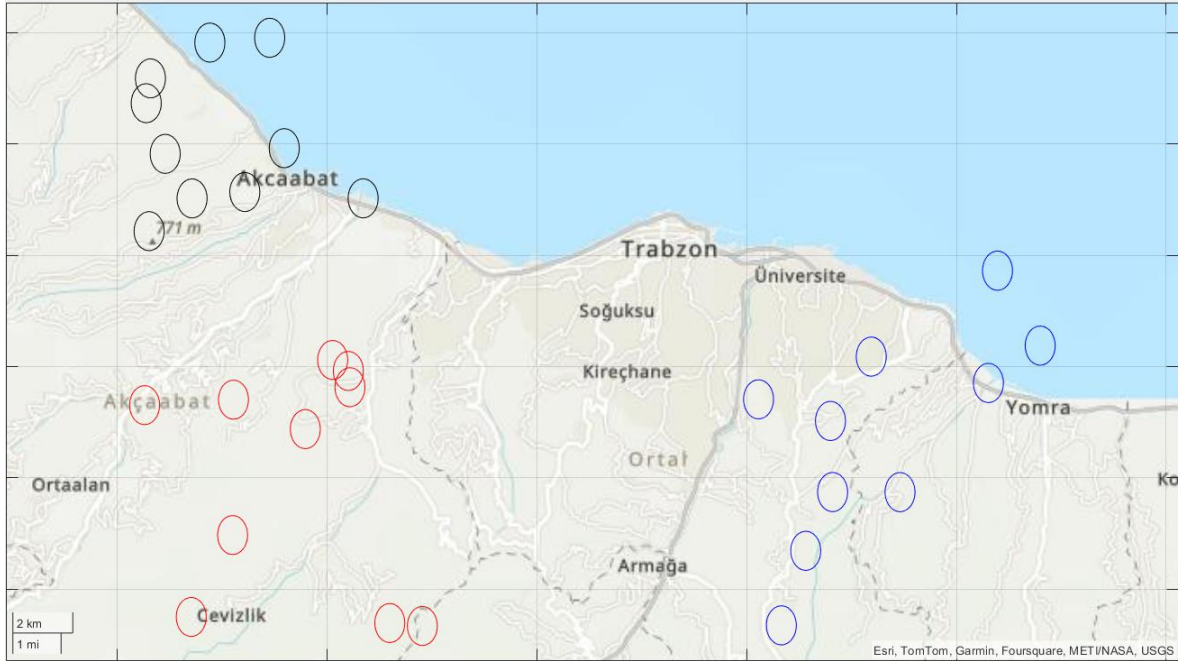
Şekil 34. Bölgeler ve HGM küçük bina dağılımları.

Şekil 36, iki bölgenin bulunduğu Akçaabat OSM alanını (yeşil sınırlı) diğer bölgelerle birlikte göstermektedir (kırmızı, beyaz ve mavi dikdörtgenler).



Şekil 35. OSM ilişkisi 1236156 (Akcaabat) ve HGM bölgeleri.

Enlem boylam oranlarındaki bozukluğu tespit etmek için bölgede rastgele olarak merkez noktaları seçilmiştir. Görüntüde ayırım yapılabilmesi için çap 1km seçilmiştir. Şekil 36, bu çap ve merkezli, yeryüzünde dairesel olan arazileri enlem boylam coğrafi koordinat sisteminde resmetmektedir.



Şekil 36. 1km çapındaki yeryüzünde dairesel alanlar.

Elipslerin ortalama basıklık değeri 0,638'dir, sıfırdan uzak ve hatta 1'e yakındır, dolayısıyla elipsler basıktır. Standart sapma ise sadece 0,0018'dir, dolayısıyla ortalamanın binde 2,8'i kadardır (varyasyon katsayısı). Bu sonuçlara göre hesaplamalar yapılırken, şekildeki elipslerin ortalamasını yansıtan elips için (0,638) enlem boylam oranı 1,3010 çıkmaktadır, bu sonuç daha önceki oranla uyumludur.

Aynı deneme, 40m çap için ve 100'er tane yapılırken aynı ortalama çıkmıştır ve standart sapma daha düşük (0,0016) çıkmıştır. Bu sonuçlara göre bölgeler yeterince küçüktür, hatta çözünürlükte bir miktar iyileştirme yapılabilir.

3.2. Veri Üst Bilgisi

Konum zaman verileri aşağıda gösterildiği gibi, Overpass QL betiğinin üst kısmında yer alır. Burada çıktı olarak JSON formatı seçilmiştir, CSV ve XML formatları da seçilebilir. Tavan-arası verileri alabilmek için sınırlayıcı dörtgenler kullanılmıştır. Bu dörtgenlerin koordinatları HGM verisinden hesaplanmıştır. Tavan-arası verisinin tarihi SDEO ile düzenlenir. Kurum personeli ya da araştırmacı OSM üzerinden incelemeler yaparak tarihe karar verir. Ayrıca görünen bir vandalizm olup olmadığını inceler.

İstatistikler için Trabzon ve Akçaabat OSM alanlarına bakılmıştır.

```
// server degistirmek icin
//{{data:overpass,server=https://overpass.kumi.systems/api/}}
//[out:json][timeout:180];
//ktu
[bbbox: 40.99, 39.77, 40.996, 39.778][date:"2023-10-02T00:00:00Z"];
//akcaabatbbox
//[bbbox: 40.918665, 39.477266, 41.041072, 39.661033][date:"2023-10-02T00:00:00Z"];
//akcaabat
//area[name="Akçaabat"];
//trabzon
//area[name="Trabzon"];
//{{geocodeArea:Trabzon}};
```

3.3. OSM İstatistikleri ve İşlemleri

Trabzon'da tüm OSM çizgilerinin % 1,12'sine, kapalı OSM çizgilerinin % 1,70'ine etiket konmamıştır. Bunlar yarı otomatik olarak yetkili verilerden etiketlendirilebilir. Kapalı OSM çizgilerinin % 90,40'ında bir, % 5,41'inde iki etiket vardır. Özellikle amenity etiketleri çoklu olarak yer alır.

3.3.1. Kaynaklar ve Uydu Görüntülerinin Kullanımı

Aşağıda gösterilen betik parçası ile, etiketlenmiş OSM çizgileri .ob değişkenine, kaynak incelemesi yapıldıktan sonra da kapalı OSM çizgileri .b değişkenine atanmıştır.

Maxar verisi .ob OSM çizgilerinden çıkarılmıştır. Bing verisi durmaktadır. Kaynak istatistikleri betiğin bu parçasının sonucunu girdi olarak kullanır.

```

way-> .ob; // area varsa way(area) -> .ob;
way.ob(if:count_tags()==0) -> .tagsiz;
(way.ob; - way.tagsiz;) -> .ob;

// Kaynaklar

//Maxar (DigitalGlobe)
(way.ob[source ~ "axar"]; way.src[source=digitalglobe]; ) -> .maxar;
(way.ob; - way.maxar;) -> .ob;

// Microsoft (Bing)
way.ob[source] -> .src;
(way.src[source=bing]; way.src[source=Bing]; way.src[source="Bing Aerial Imagery"]; ) ->
.bing;
(way.bing; way.src[source ~ 'icrosoft']; ) -> .bing;
// (way.ob; - way.bing;) -> .ob;

// ABD (Landsat, GPS, PGS(sahil seridi)
(way.src[source=landsat]; way.src[source=GPS]; way.src[source=PGS];) -> .abd;

// GOV.TR
way.src[source~'gov.tr'] -> .srcgovtr;

// Uygulamalar (Mapbox, Strava)
(way.src[source~"apbox"]; way.src["source"~"trava"]; ) -> .srcuyg;

// Diger
(way.src[source=estimated]; way.src[source=unreliable];) -> .srcziv;

// Coklu kaynak
way.src[source~";"] -> .karisik;
(way.karisik; way.srcziv;) -> .srcziv;

//Kapali OSM cizgileri
way.ob(if:is_closed()) -> .b;

////////////////////////////////////

```

Maxar (eski adıyla DigitalGlobe) dünyada önde gelen bir uydu görüntüsü sağlayıcısıdır; özellikle 2013 yılında rakibi GeoEye'ı bünyesine kattıktan sonra yüksek konumsal çözünürlükte lider bir duruma gelmiştir. Maxar OSM kurumu ile anlaşma yaparak

kendi görüntülerinin OSM üzerinde, OSM dışına çıkmadan, kullanımına olanak vermiştir. Bu tez çalışmasında lisans problemi nedeniyle Maxar kaynaklı veriler kullanımdan çıkarılmıştır. Türkiye’deki yetkili kurumlar Maxar’dan görüntü satın alırken OSM ile ilgili izinleri de talep edebilir.

2025’te görüntü sağlamaya başlamayı planlayan Albedo şirketi konumsal çözünürlüğü, Maxar’ın 31cm seviyesinden 10cm seviyesine yükseltecek, yapay zekâ algoritmalarını da kullanacaktır (SatImagingCorp, 2023).

Microsoft (örneğin Bing) kaynaklı verilerin ticari kullanımı yasaktır, Bing’te mevcut olan sokak adları gibi öznitelikler de Bing verilerinden aktarılmaz, StreetSide verileri kullanılabilir.

USGS, Landsat ve Sentinel-2 görüntülerini ücretsiz sunar ama kaynaklarda Landsat görüntülerine denk gelirse de Trabzon için Sentinel-2 görüntüleri kaynak olarak kullanılmamıştır. Landsat çözünürlüğü 30m’dir ama 1970’li yıllardan uydu görüntüleri için elimizdeki ender kaynaklardan biridir. Sentinel-2’nin konumsal çözünürlüğü üç kat daha iyidir; 10 metredir. Bu çözünürlük, Bing, Mapbox, Maxar ile karşılaştırıldığında düşüktür (ama bu servisler de Landsat görüntülerini ellerinde daha iyi verinin olmadığı bazı yerlerde sunarlar).

PGS verilerinde (Prototype Global Shoreline Data) de kalite problemi vardır. GPS kaynaklı veriler, GLONASS (Rusya), Galileo (Avrupa), BeiDou (Çin) gibi konumlama servislerini de kullanıyor olabilirler.

Trabzon OSM çizgilerinin % 9,0’ının kaynak verisi vardır. Aşağıda kaynaklı olan verilerin kaynağına göre yüzdeleri belirtilmiştir. Örneğin Maxar, kaynaklı verilerin % 52,2’sinin kaynağıdır, dolayısıyla tüm verilerin yaklaşık % 4,5’inin kaynağıdır. Ayrıca tüm veriler indirilip değerlendirilmeyecektir. İndirilecek olan verilerde, Microsoft Bing kaynak oranı düşmektedir.

• Maxar	52,2
• Microsoft	42,0
• Landsat	0,1
• GPS	3,7
• PGS	0,1
• Mapbox	1,6
• Strava	0,1
• gov tr	0,0

3.3.2. Kapalı Yol İncelemesi

Kapalı OSM çizgilerinin hepsi bina değildir. Kaynak incelemesinden sonra betikte, highway OSM çizgilerine bakılmıştır. Bunların % 0,57'si kapalıdır. Kapalı highway OSM çizgilerinin aşağıda görüldüğü gibi % 70,4'ü yaya yoludur (pedestrian, footway). Ayrıca 35 tane döner kavşak belirtilmiştir (junction roundabout).

• pedestrian	60.1
• footway	10.3
• trunk	5.8
• trunk_link	4.5
• secondary	4.9
• tertiary	3.6
• service	7.2
• residential	1.3
• track	0.4
• steps	0.4
• path	0.4
• unclassified	0.9

Betiğin kapalı yol inceleme kısmı aşağıda gösterilmiştir.

```
// Kapalı yollar
way.b[highway] -> .hw;
way.hw[junction=roundabout] -> .donerkavsak;
(way.hw[highway=pedestrian]; way.hw[highway=footway];) -> .yayayolu;
(way.hw; - way.yayayolu;) -> .hw;

(way.hw[highway=trunk]; way.hw[highway=trunk_link]; way.hw[highway=secondary];
way.hw[highway=tertiary]; way.hw[highway=service];) -> .yollar;
(way.hw; - way.yollar;) -> .hw;
//way.yayayolu -> .indir;
//way.yollar -> .indir;

/////
```

3.3.3. Tarlabahce

HGM'deki Tarlabahce katmanı için aşağıdaki betik kullanılarak OSM verisi bir araya getirilmiştir. Kapalı OSM çizgilerinin % 0,54'ü bu tip veridir. Akçaabat kırsalında bir tane

orchard ve bir tane farmland olmak üzere iki tane detay vardır. Buralarda bir hata olmasa da, analiz için sayı çok azdır. Diğer iki bölgede de durum aynıdır.

```
// Tarlabahce
(way.b[landuse]; -way.b[building];) -> .lu;
(
  way.lu[landuse=orchard];
  way.lu[landuse=farmyard];
  way.lu[landuse=farmland];
  way.lu[landuse=village_green];
  way.b[leisure=garden];
) -> .tarlabahce;
// way.tarlabahce -> .indir;
(way.lu; -way.tarlabahce;) -> .lu;
////////
```

3.3.4. Veri Seçme ve İndirme

Ana veri için indirilecekler aşağıdaki betik kısmı ile ayarlanmıştır. building OSM çizgileri yanında böyle etiketlenmemiş olabilen okullar, dini yapılar, hastaneler vs. eklenmiş, bir kısım veri de çıkarılmıştır. Sonra da bu OSM çizgilerinin düğümleri de indirilecek şekilde veriye eklenmiştir. Çıkış komutu “.indir out meta center qt;” Bu komutta meta terimi kullanıldığı için koordinatlar ve etiketler dahil çoğu veri indirilir. Ayrıca center terimi kullanıldığı için detayların ağırlık merkezleri de eklenir. qt veri ekleme sıralamasını değiştirir; qt ile daha hızlı bir sorgu ve indirme gerçekleştirebilir. bb eklenseydi, her düğümün ve çizginin sınırlayıcı dörtgenleri de çıkacaktı. Düğümler için ağırlık merkezleri kendileridir, ek veri oluşmaz.

```

// Indirilecek ayari
(
    way.b[amenity =university];
    way.b[amenity =school];
    way.b[amenity= kindergarten];
) -> .okul;
(
    way.b[amenity =place_of_worship];
    way.b[landuse=religious];
) -> .dini;
(
    way.b[amenity=hospital];
) -> .saglik;
(
    way.b[landuse=cemetery];
    way.b[landuse=grave_yard];
    way.b[amenity=grave_yard];
) -> .mezarlik;
(
    way.b[building];
    way.b[landuse=construction];
    way.okul;
    way.dini;
    way.saglik;
    //way.mezarlik;
    way.b[amenity=restaurant];

) -> .indir;

(
    way.b[amenity=parking];
    way.b[natural];
    way.b[leisure];
    way.b[landuse=industrial];
    way.b[landuse=forest];
    way.b[landuse=residential];
    way.b[highway=unclassified];
    way.b[amenity=animal_shelter];
    way.b[amenity=bicycle_parking];
    way.b[amenity=bicycle_rental];
    way.b[amenity=bus_station];
    way.b[amenity=fountain];
    way.b[amenity=marketplace];
    way.b[amenity=taxi];
    way.b[amenity=water_point];
) -> .cikar;
(way.indir; - way.cikar; ) -> .indir;

// Dugumleri ekleme
(way.indir; >;) -> .indir;

/////

```

3.4. Dini Yapı Karşılaştırması

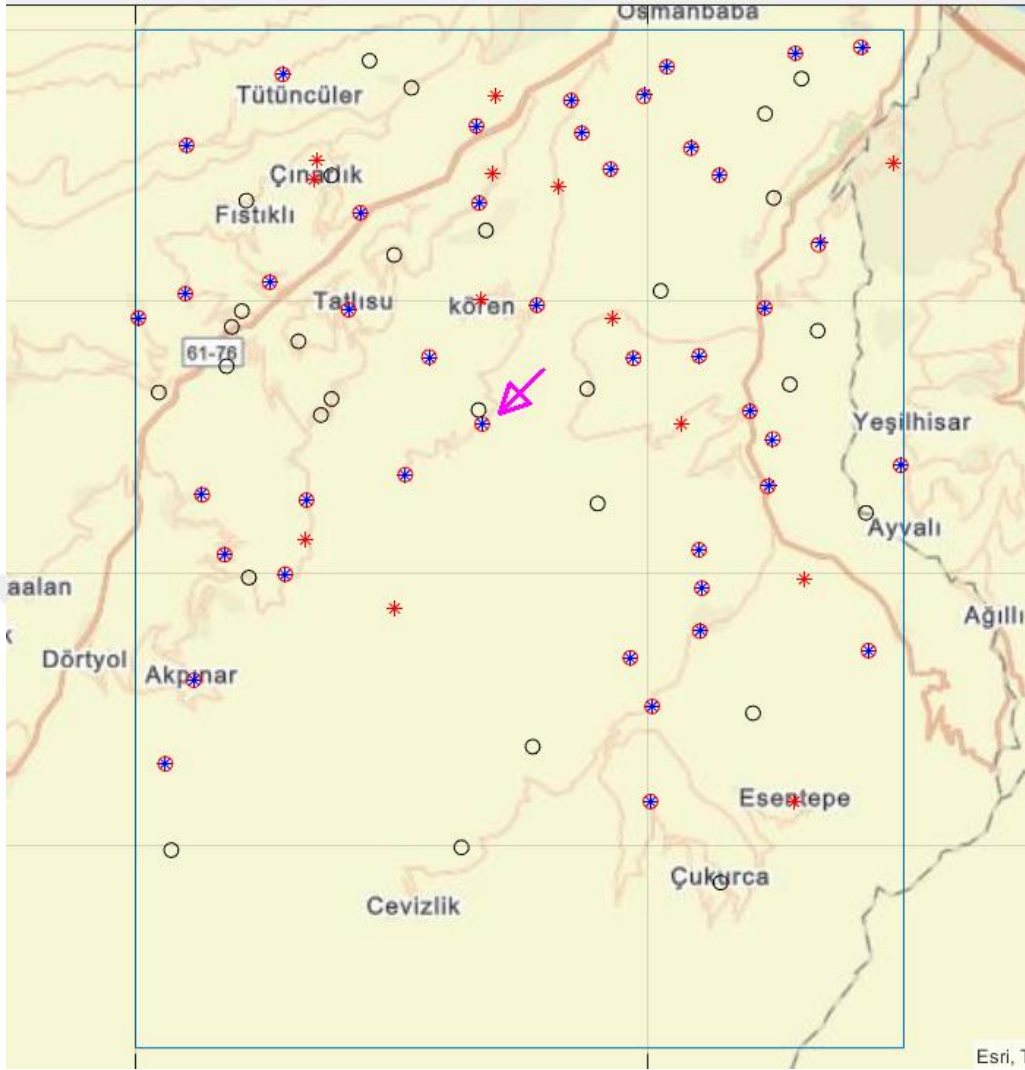
Camiler için yafonkka: 2, bikula:10 (ibadethane), kucukbina katmanındaki (topodetays: 65) ve bina için topodetaya:7, buyukbina katmanındaki (topodetays: 66) bina için topodetaya:4. Hem kucukbina, hem de buyukbina katmanlarında aynı alt sıra numarası denk gelmiştir.

Akçaabat kırsal bölgesinde toplam 7601 kucukbina, 50 buyukbina vardır. Buyukbina'ların sadece bir tanesinin adı kayıtlıdır. kucukbina'ların 24 tanesinin adı kayıtlıdır (% 0,3). Veri isim özneteliği açısından eksiktir.

129 kucukbina dini yapı cami olarak belirtilmiş, 68 tane dini yapı cami olarak belirtilmemiştir. Buyukbina katmanında dini yapı yoktur.

HGM'de camiler sıfır boyutlu iken OSM'de kapalı çizgidir (iki boyut). Tüm Trabzon güncel OSM verisinde bunun istisnaları sadece 3 tanedir (Uzun Sokak Emine Usta Mescidi, Özdil Merkez Camii ve Cumapazarı'nda bir cami). Bu istisnalar, OSM için kalite hatasıdır (mantıksal tutarlılık).

Şekil 37 ile HGM ve OSM camilerinin eşleşmesi harita üzerinde gösterilmiştir. Kırmızı çemberler eşleşen OSM camilerini, içlerinde gözüken mavi yıldızlar eşleşen HGM camilerini gösterir.



Şekil 37. Akçaabat camilerinin eşleşmesi.

HGM camilerinin ontolojide seçilen eşik değerinde (0,0005 derece yani yaklaşık 50 metre), % 77'si eşleşmiştir. Eşik değeri 0,0001'e (10 metre) düşürülürse, eşleşme oranı % 64'e düşmektedir.

Yaklaşık 50 metrede, eşleşmeyen % 23 HGM camisi incelendiğinde (kırmızı yıldızlar), eşleşmeyen OSM camileri (siyah çemberler) yakın değildir. Kırmızı yıldızlarda HGM camileri vardır, yakınlarında OSM camisi yani siyah çember yoktur. Ortadaki bir tane (okla gösterilen) siyah çemberin durumu Şekil 38 ile ayrıca incelenmiştir. Orada da hata yoktur.

50 metre civarında bir hata payı serbestisi ile hatalı bir eşleşme ve eşleşmeme yoktur. Aynı durum diğer bölgelerde de geçerlidir. Fakat eşik değeri düşürüldüğünde, HGM verisindeki detayı tek bir nokta ile temsil etme konusundaki seçim, tek başına hatalara sebep

olmuştur. Her HGM camisi ağırlık merkezleri ile ya da mesela güneybatı köşeleri ile temsil edilebilirdi.



Şekil 38. Yeşilyurt yakınında eşleşen cami eşdeğer çifti ve diğer OSM camisi.

MInPolygon yöntemi yine aynı veri setine uygulandığında, önce OSM camileri “dışarıya doğru” yapı merkezinden olan uzaklığa orantılı olarak katsayı kadar (örneğin 0,05) merkezden köşeler uzaklaştırılır (lineer kombinasyon). HGM noktasının bu genişletilmiş poligonun içinde olup olmadığına bakılır.

50 metre seviyesinde eşleşen camiler sdeo:isAlike ile ilişkilendirilmiştir. sdeo:KucukCami sınıfı sdeo:Anchor sınıfının alt sınıfı yapılmıştır (üç çapa çifti daha vardır). Seçilen yöntem (SAP) ve dönüşüme (tan-affin) göre eşleşen OSM camilerine yeni koordinat değerleri belirlenmiş ve ontolojide düzeltmeler yapılmıştır. Dönüşüm tekrar hesaplanırsa etkin bir şekilde değişmez ($A \cdot A'$ tersi var)). Dönüşüm için homeomorfizm seçilirse ama fonksiyonlar monoton olacak şekilde seçilirse benzer sonuç çıkacaktır. Böyle bir durumda çapa noktalarının sıralaması da önemlidir.

Direkt afin dönüşümde dönüşüm matrisi birim matrise yakın çıkmaktadır. Tan-afin dönüşümü, dönüşümün kendisi reel düzlemde gerçekleşsin diye yapılmıştır. Tanjant fonksiyonu sadece sürekli artan analitik bir fonksiyon değildir, türevinin değeri en az birdir. Bundan dolayı, aradaki farkları büyütür. Tanjant uygulandıktan sonra dönüşüm matrisi birim matrizen farklılaşır. Uygulamada çıkan tan-afin dönüşüm matrisi aşağıdadır.

0.9995	-0.0013	0.0009
-0.0003	0.9994	0,0005
0	0	1

3.5. Sonuç Metrikleri

Yöntem uygulandığında Matthews korelasyonu % 81,8 gibi yüksek oran çıkmıştır.

Diğer metrikler:

- Kesinlik (Precision): %87,35
- Duyarlılık (Recall/Sensitiviy) %89,55
- Doğruluk (Accuracy): %91,51
- F1 Skoru: %88,44
- Özgüllük (Specificity): %92,63

$$Kesinlik (Precision) = \frac{1}{1 + FP_-}$$

$$Duyarlılık (Recall, Sensitivity) = \frac{1}{1 + FN_-}$$

$$F1 = 2 \frac{Kesinlik + Duyarlılık}{Kesinlik * Duyarlılık}$$

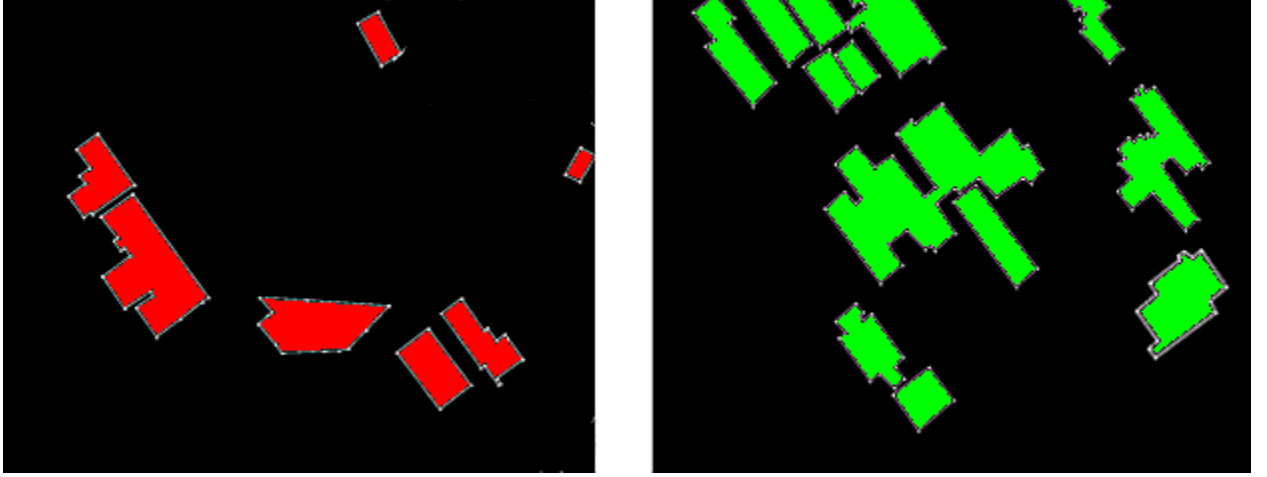
$$Doğruluk (Accuracy) = \frac{1 + TN_-}{1 + TN_- + FP_- + FN_-}$$

$$Özgüllük (Specificity) = \frac{TN_-}{(TN_- + FP_-)}$$

$$MCC = \frac{TN_- - FP_- \cdot FN_-}{\sqrt{(1 + FP_-) \cdot (1 + FN_-)(TN_- + FP_-)(TN_- + FN_-)}}$$

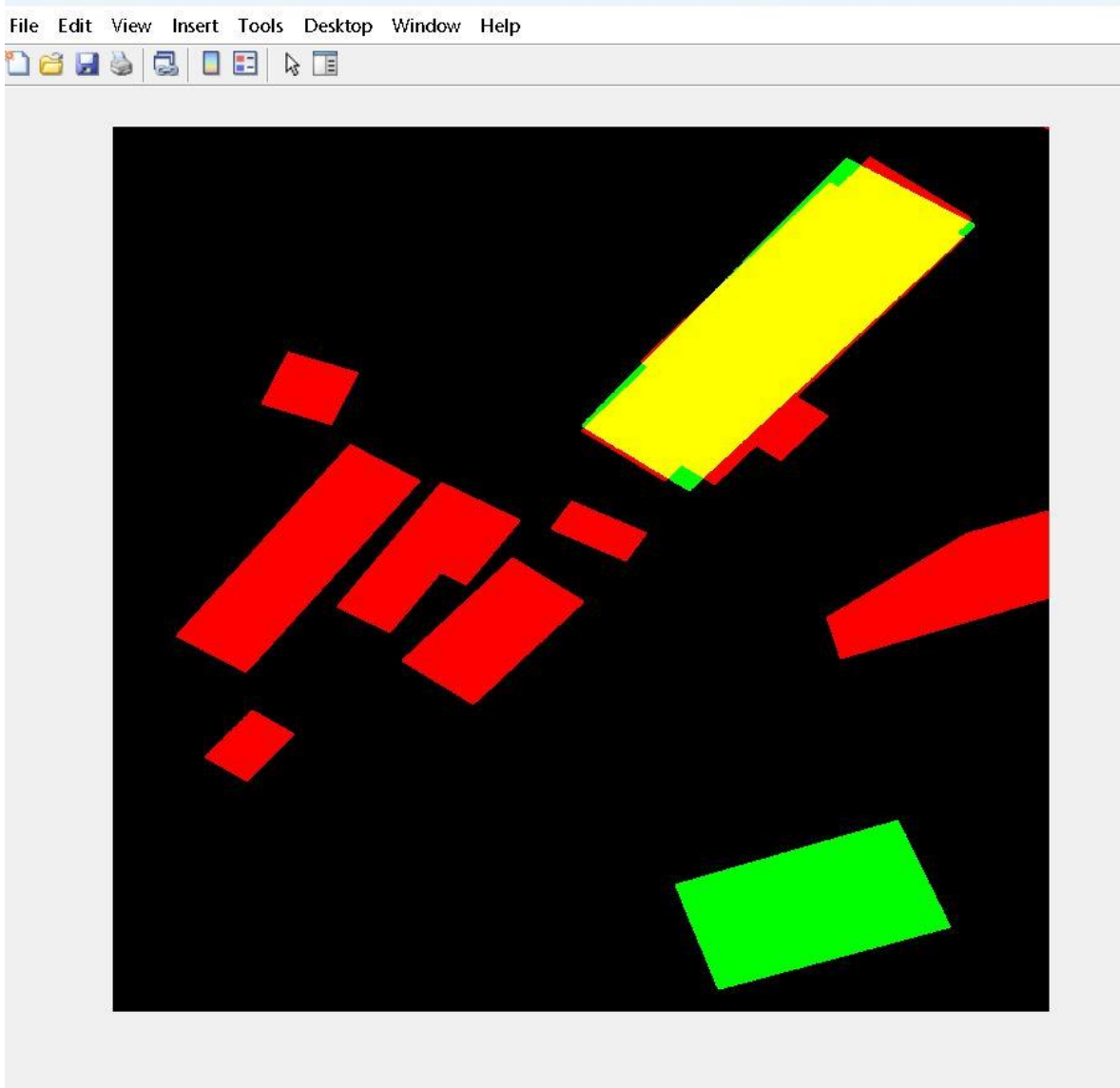
Burada $FP_- = FP/TP$, $FN_- = FN / TP$, $TN_- = TN/FP$.

HGM verilerinin yoğunluğunun görece azlığı TN/TP oranını arttırmıştır. HGM iki boyutlu verilerinin daha sık olduğu KTÜ bölgesinde de durum böyledir. HGM sıklığı Şekil 39 ile resmedilmiştir. Burada verilerin güncelliği önem göstermektedir. Kenar iyileştirme yapılmıştır.



Şekil 39. HGM’de olmayan binalar (sol taraf). HGM binaları (sağ taraf)

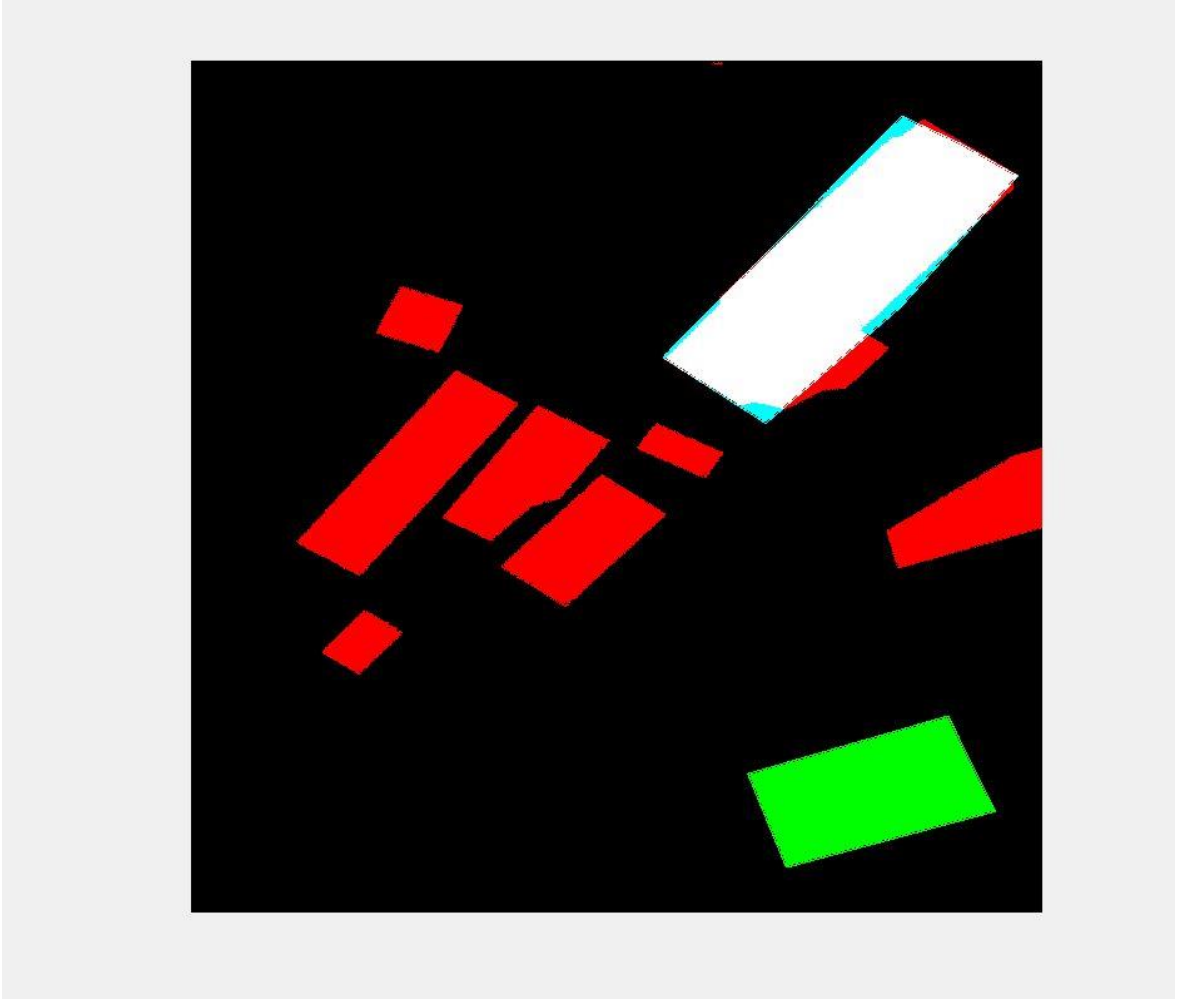
Orijinal poligonlarda Akçaabat merkezi bölgesinin bir kısmı Şekil 40 ile gösterilmiştir. Burada da OSM verisi daha fazladır. Yeşil bina sadece HGM’de vardır. Kırmızı binalar sadece OSM’de vardır. Sarı (kırmızı+yeşil) renk çakışmayı gösterir. Bir binada yüksek oranda çakışma olmuştur. Burada sadece iki bant kullanılmıştır. Bunun için önceden kaydedilmiş görüntüler değerlendirilmiştir. Uygulama, üç kanallı kullanıma yöneliktir. Kanallar aynı bandın olabilir.



Şekil 40. Akçaabat merkezi bölgesinden bir parça

Aynı bölge SAP poligonları ve 2,1,3 bant-kanal kombinasyonu kullanılarak Şekil 41 ile gösterilmiştir. Birinci yani kırmızı kanaldan ikinci bant yani OSM bandı geçirilmiştir. İkinci kanalda birinci bant olan HGM bandı yansıtılmıştır. Üçüncü bant ve burada üçüncü kanal ise eşlenikli HGM detaylarını belirtir. Beyaz renkli yerler TP, cyan renkli yerler FP hata tiplerini yansıtır. Bu görüntüde sarı renkli yerler bulunmamaktadır, sarı renkli yerler FN hata tipine karşılık gelir. Az sayıda pikselde magentalık önemli ölçüde vardır.

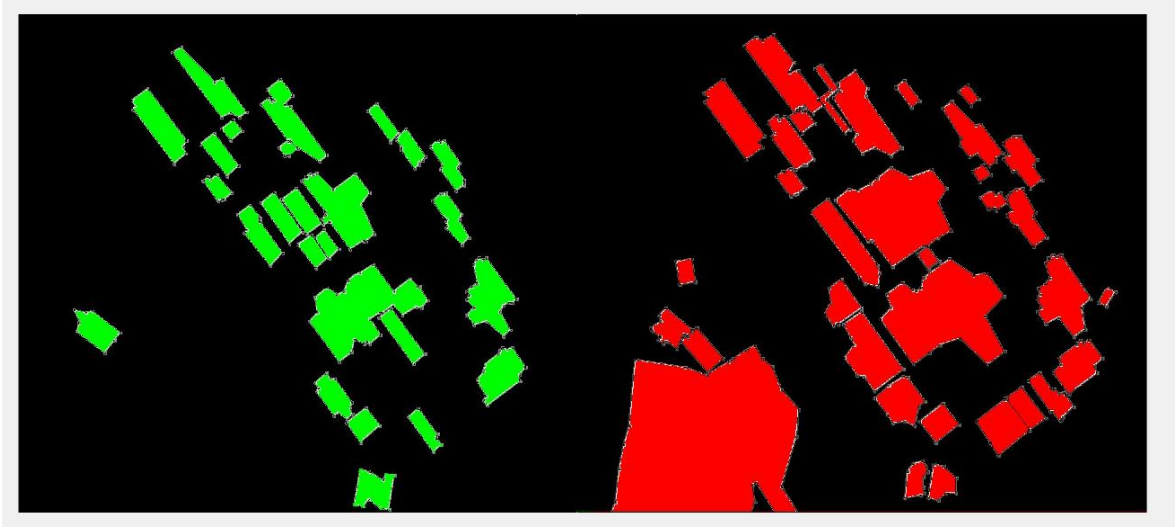
Eşlenene kadar OSM yapılarının gri değerleri bire eşittir. Bu gösterim için bantlar mantıksal tipe çevrilmiştir. Bu durumda sırf bu kısım için kesinlik değeri % 93 civarındadır ve Trabzon'un diğer yerlerinden iyidir. Tam duyarlılık gerçekleşmiştir.



Şekil 41. Akçaabat şehir merkezinden bir parça (SAP).

KTÜ Makine Mühendisliği binası OSM’de tek kapalı çizgi ile temsil edilirken HGM’de beş poligon olarak kaydedilmiştir. Bu da FN hatasını arttırmaktadır.

Rasterizasyon sonrası SAP poligonları KTÜ bölgesi için Şekil 42 ile gösterilmiştir. Örneğin inşaat halinde olan alanlar HGM verisinde yansıtılmamıştır.



Şekil 42. KTÜ mantıksal SAP raster, HGM(yeşil) ve OSM (kırmızı).

3.6. Olası Vandalizm ve İlgili Hatalar

Amenity ile etiketlenmiş düğümlerin çizgiler ile bağlantısı incelenirken 7288577554 numaralı bir güncellemenin sonuçları tespit edilmiştir. Aynı değişiklik setindeki (Changeset: 147964174) diğer işlemler incelendiğinde, bir vandalizm şüphesi oluşmuştur. Yerel olarak yapılan bu değişikliklerin yanında dünya çapında yapılan değişiklikler de vardır. Bunlar kolayca tespit edilir ve hızlıca geri alınır. Çok sayıda düğümün ve diğer objelerin değiştiği setler birçok kullanıcının takibi ile hızlıca belirlenir.

4. SONUÇLAR VE ÖNERİLER

Bu çalışmanın ortaya koyduğu metodoloji, yetkili konumsal verilerin kalitesini, kitle kaynak verilerle entegrasyonunu sağlayıp iyileştirme hedefinde önemli bir adım atmıştır. Tez çalışmasında Trabzon ilinde, Akçaabat kırsalı, Akçaabat şehir merkezi ve KTÜ olmak üzere seçilen üç ana bölgede, HGM'den alınan yetkili veriler ile OSM üzerinden toplanan kitle kaynaklı verilerin etkili bir şekilde entegre edilmesini sağlamıştır. Tez çalışmasının temel hedefi, kitle kaynak verilerinde özellikle bina verilerindeki değişimi belirleyip yetkili verilerde güncellenecek bina detaylarını belirlemektir. Bu noktadan hareketle diğer alan detaylar için yöntem genişletilebilir. Önerilen yöntem, dünya üzerinde farklı bölgelerde de uygulanabilecek durumdadır.

Çalışmada MATLAB yazılımı kullanılmıştır. Bu yazılımla hem işlemler gerçekleştirilmiş hem de kurum personellerinin kullanımına yönelik GUI geliştirilmiştir. MATLAB aynı betik içinde Python ve Java kodlarını direkt olarak, bir dosya yüklemeyen çalıştırabilir, fakat verimlilik için modüller ve *.jar dosyalarının kullanımı önerilir. C/C++ kodları derlenip MATLAB yürütülebilir mex dosyaları ile betiklerde kullanılabilir. Çeşitli algoritmalar, OWLAPI gibi Java sınıfları için mex ve *.jar dosyalarına başvurulmuştur. MATLAB, matris işlemlerinde, dolayısıyla görüntü işlemede iyidir. Gelecekteki çalışmalarda yapay zekâ algoritmaları için Python modülleri kullanılabilir.

Çalışmada sunulan metodoloji açık dünya kavramını ve seçeneklerde esnekliği dikkate almaktadır. Ontoloji tasarımında kullanıcılara uzmanlığını aşan yük bindirilmemesi göz önüne alınmıştır. Ontoloji editörü ile yapılan değişiklikler, GUI ile tercih edilen seçeneklerle birleşmesiyle istenilen sonuçlara yaklaştıır. Literatürde bahsedilen yöntemlere benzeyebilen yöntemler de dâhil olmak üzere ondan fazla yöntem, parametrelerinin oluşturduğu ekstra bir seçenek havuzuyla birlikte ontolojide sunulmaktadır. SAP yöntemi iki boyutlu yapılar için tez çalışmasında tasarlanmış özgün bir yöntemdir.

Sonuç olarak bu araştırma, özgün bir ontoloji olan SDEO ontolojisinin geliştirilmesiyle, konumsal veri yönetimi ve analizinde yenilikçi bir paradigma sunmaktadır. Araştırmanın bulguları, kitle kaynaklı ve yetkili verilerin entegrasyonu ve analizinin, veri kalitesinin artırılmasında önemli bir rol oynayabileceğini göstermektedir. Özellikle, bu çalışmanın metodolojisi ve ontoloji tabanlı yaklaşımı, ilgili alanlarda önemli gelişmelere yol açabilir ve coğrafi bilgi biliminin geleceğine katkıda bulunabilir.

4.1. Metodolojik İnovasyonlar ve Uygulamalar

Bu araştırmanın merkezinde, özgün olarak geliştirilen SDEO ontolojisi yer almaktadır. SDEO ontolojisi, veri kalitesinin denetimi ve iyileştirilmesi süreçlerini yönlendirerek, kitle kaynaklı ve yetkili veriler arasındaki entegrasyonu kolaylaştırmıştır. Önerilen ontoloji tabanlı yaklaşım, veri yönetimi süreçlerinde önemli iyileştirmeler sağlamış ve semantik olarak zenginleştirilmiş verilerin işlenmesine yeni olanaklar sunmuştur. SDQO ontolojisi çalışmada ön denetim için kullanılmıştır.

Kitle kaynaklı veri ve yetkili veriler arasındaki entegrasyon, ontolojiler sayesinde daha verimli bir şekilde gerçekleştirilmiştir. Ontoloji kullanımı, verilere semantik bir bağlam kazandırır. Ontolojiler, farklı veri setlerinden gelen bilgilerin birleştirilmesinde ve karşılıklı ilişkilerin kurulmasında önemli bir rol oynar. Örneğin basit bir hiyerarşik işlemle, detayların nerede nasıl kullanılacağı değiştirilebilir.

Rasterizasyon aşamasına geçildikten sonra gri değerlerine, azaltılmış bitli kimlik verilerini atamak, ilişkilendirme işlemlerini ve denetimini kolaylaştırmıştır. Ontolojilerde ilişkilendirmeler kullanılarak, örneğin OSM güncellemeleri, veriye sonraki süreçlerde yansıtılabilir.

SAP yöntemi inovatif bir yöntemdir. Ayrıntılı bilgi 4.4 bölümünde yer almaktadır. Bu yöntem sonunda test edilen bölgelerle ilgili çok sayıda bant oluşturulur. Kurum elemanı ve araştırmacının tercihi ile hangi kanallardan hangi bantların gösterileceğine ontoloji düzenlenerek karar verilir. Bunun sonucunda çok sayıda farklı kombinasyonda bilgi yansıtılmasının olanağı doğmuştur.

4.2. Karşılaşılan Zorluklar Ve Çözüm Önerileri

OSM ve HGM verilerinin entegrasyonu sırasında karşılaşılan zorluklar arasında OSM verilerinin güncellik farkı ve vandalizm riski gibi dezavantajlar bulunmaktadır. Ancak, bu zorlukların üstesinden gelindiğinde değerli veri setleri elde edilebilir. Tavan-arası verilerini dikkatli bir şekilde seçip kullanmanın sebepleri arasında bu sorunu aşmak vardır.

Dış kaynaklardan elde edilen verilerin kullanımı, lisans kısıtlamalarını dikkate almayı gerektirir. Özellikle, Maxar kaynaklı OSM verilerinin dikkatli bir şekilde kullanılması ve 2023 yılındaki Maxar satın alımı sonrası değişen kullanım koşullarına uyum sağlanması önemlidir. OSM'nin açık kaynak olması, her verinin OSM dışında özgürce kullanılabileceğini göstermemektedir. OSM şirketlerle anlaşmalar yaparak veri kaynaklarını

çeşitlendirmeyi başarmıştır ama bu anlaşmalarda açık dünya fikrine uyum her zaman sağlanmamıştır.

Kurumlar Maxar şirketinden görüntü satın alıyorsa, satın alma sürecinde OSM kullanımı ile ilgili olarak danışabilir. Uygulama bölgelerini kapsayan Trabzon'daki kaynağı belirtilmiş OSM verilerinin (Mapillary etiketliler hariç) yarısından çoğu Maxar kaynaklıdır.

Microsoft Bing kaynaklı veriler görece az sayıdadır, OSM Wikideki şablonlar göz önüne alınmış ve bu veriler elenmemiştir. Ticari bir kullanım zaten mevcut aşamada öngörülmemektedir.

Türkiye OSM tamlığı açısından dünyada kötü seviyede olan ülkeler arasındadır. Buna rağmen verilerin tamlığı Trabzon için yerindedir.

HGM verilerinde poligon tipinde olması beklenebilecek olan veriler, küçük diye nokta tipinde tutulmaktadır. Bu duruma bağlı sorunları aşmak için nesne özellikleri ve SWRL kuralları kullanılmıştır.

MATLAB ortamında ESRI Shapefile dosyalarının kullanımında yaşanan zorluklar, tablo ve tablo fonksiyonlarının kullanımı ve çeşitli düzenlenmeler ile aşılmıştır.

4.3. Gelecek Araştırmalar İçin Öneriler

Gelecek çalışmalar, bu araştırmanın özgün metodolojisini daha geniş veri setleri ve coğrafi alanlar üzerinde uygulayarak, sonuçların genelleştirilebilirliğini değerlendirebilir. Farklı kitle kaynaklı verilerin kalitesinin sistematik olarak değerlendirilmesi, yetkili verilerde kitle kaynak veriler kullanılarak güncellenen detayların belirlenmesini daha etkili hale getirebilir. Makine öğrenimi teknolojilerinin entegrasyonu, veri analizi ve yönetim süreçlerindeki verimliliği ve doğruluğu artırabilir. Ayrıca, özellikle çizgi tipindeki yapılar için olan yöntemler daha da geliştirilebilir. Uzunluk ve açıların devreye girmesi sağlanabilir.

Araştırma, kitle kaynaklı veri ve yetkili verilerin entegrasyonu ve analizi için sağlam bir metodolojik çerçeve sunmuştur. Gelecek çalışmalar, bu çerçeveyi genişleterek, entegrasyonun çeşitli alanlardaki potansiyelini daha da artırabilir.

4.4. SAP Yöntemi

Metodoloji çerçevesinde yöntemler sunulmuştur ve bu yöntemler parametrelerle zenginleştirilmiştir. Örneğin dönüklük için homeomorfizm gibi seçenekler sunulmuştur.

SAP yöntemi özellikle iki açıdan özgün bir yöntemdir. Ara poligonlar ve bunların köşeleri yerine Steiner tipi köşeler eklenerek sanal ara yapılar oluşturulur. İlişkilendirmelerde karşılaştırılacak olanlar, orijinal poligonlar değil, karşılaştırmanın veri kaynaklarındaki doğa farklılığına uyumlu olması istenen bir şekilde türetilmiş olan SAP poligonlarıdır. Diğer bir yandan, poligon yakınlığı/benzerliği kavramı incelenmiştir.

Makine öğrenimini canlandıran konulardan biri memeli hayvan ve insan görüşündeki araştırma sonuçlarının bilgisayar biliminde değerlendirilip onlara uymaya çalışan algoritmalar üretmek olmuştur. İnsan görüşünün katmanlılığı derin öğrenme kavramını ortaya çıkarmıştır. İnsan, vektörel olarak görmez ve konumsal çözünürlük düşüktür. Bir wabi-sabi etkisiyle bunlar, örneğin bir uydu görüntüsünde detay çıkarımını çok hızlı yapabilmesine yol açmıştır. Rasterizasyon, SAP yönteminin önemli bir diğer noktasıdır. Karşılaştırmalar, rastere dönüşümden sonra yapısal metrikler ile yapılır. Literatürdeki tampon yöntemleri, bu yönde bir çaba olarak görülebilir.

Tüm bu işlemlerin sonunda, kurumda var olabilecek OSM hakkında bilgili bir personel düşünülerek, OSM’de yapılabilecek işlemlerin listesi de oluşturulmaktadır. OSM verisi, uygun bulunursa, HGM öznitelikleri ile zenginleştirilebilir.

5. KAYNAKLAR

- Albakri, M. M. (2016). Semantic Similarity Assessment of Volunteered Geographic Information. *Journal of Engineering*, 22(1), 215-229. doi:10.31026/j.eng.2016.01.14
- Alomari, A. H., Al-Omari, B. H. & Al-Hamdan, A. B. (2020). Validating trip travel time provided by smartphone navigation applications in Jordan. *Jordan Journal of Civil Engineering*, 14(4).
- Anderson, J., Sarkar, D. & Palen, L. (2019). Corporate Editors in the Evolving Landscape of OpenStreetMap. *ISPRS International Journal of Geo-Information*, 8(5), 232. doi:10.3390/ijgi8050232
- Antoniou, G. & Van Harmelen, F. (2008). *A Semantic Web Primer* (2nd. ed.). MIT Press.
- Antoniou, V. (2011). *User Generated Spatial Content: An Analysis of the Phenomenon and its Challenges for Mapping Agencies* Doctoral thesis. <http://discovery.ucl.ac.uk/1318053/>
- Antoniou, V. & Skopeliti, A. (2015). Measures and indicators of vgi quality: An overview. *ISPRS Annals of the Photogrammetry, Remote Sensing and Spatial Information Sciences*, 2(3W5), 345-351. doi:10.5194/isprsannals-II-3-W5-345-2015
- Antoniou, V. (2017). Assessing VGI Data Quality. *Mapping and the Citizen Sensor* (pp. 137-163). Ubiquity Press. doi:10.5334/bbf.g
- Antunes, F., Fonte, C., Brovelli, M., Minghini, M., Molinari, M. & Mooney, P. (2015a). Assessing OSM Road Positional Quality with Authoritative Data. VIII National Conference on Cartography and Geodesy (VIII CNCG), PRT.
- Arcbazar. (2023). ArcBazar. Erişim adresi: <https://tr.arcbazar.com/about>.
- Arsanjani, J. J., Barron, C., Bakillah, M. & Helbich, M. (2013). Assessing the quality of OpenStreetMap contributors together with their contributions. Proceedings of the AGILE,.
- Arsanjani, J. J. & Bakillah, M. (2015). Understanding the potential relationship between the socio-economic variables and contributions to OpenStreetMap. *International Journal of Digital Earth*, 8(11), 861-876. doi:10.1080/17538947.2014.951081
- Arsanjani, J. J., Jamal, Mooney, P., Zipf, A. & Schauss, A. (2015). Quality Assessment of the Contributed Land Use Information from OpenStreetMap Versus Authoritative Datasets. *Lecture Notes in Geoinformation and Cartography* (Vol. 0, pp. 37-58). doi:10.1007/978-3-319-14280-7_3
- Ather, A. (2009). *A Quality Analysis of OpenStreetMap Data: M. Eng. Dissertation ME Thesis, University College London, London, UK.*

<ftp://ftp.cits.rncan.gc.ca/pub/cartonat/Reference/VGI/Dissertation-OpenStreepMap-Quality-Aather-2009.pdf>

- Avbelj, J., Muller, R. & Bamler, R. (2015). A Metric for Polygon Comparison and Building Extraction Evaluation. *IEEE Geoscience and Remote Sensing Letters*, 12(1), 170-174. doi:10.1109/LGRS.2014.2330695
- Ballatore, A. & Arsanjani, J. J. (2019). Placing Wikimapia: an exploratory analysis. *International Journal of Geographical Information Science*, 33(8), 1633-1650.
- Ballatore, A. & Zipf, A. (2015). A Conceptual Quality Framework for Volunteered Geographic Information. *Lecture Notes in Computer Science (including subseries Lecture Notes in Artificial Intelligence and Lecture Notes in Bioinformatics)* (Vol. 9368, pp. 89-107). Springer. doi:10.1007/978-3-319-23374-1_5
- Bärlocher, M. (2013). OpenSeaMap-The free nautical chart. *Hydro Int*, 17.
- Barron, C., Neis, P. & Zipf, A. (2014). A Comprehensive Framework for Intrinsic OpenStreetMap Quality Analysis. *Transactions in GIS*, 18(6), 877-895. doi:10.1111/tgis.12073
- Barth, A. (2013). New York City and OpenStreetMap collaborating through open data. *Mapbox*. <https://www.mapbox.com/blog/nyc-andopenstreetmap-cooperating-through-open-data>. Erişim tarihi: 19 Şubat 2024. Erişim adresi: <https://blog.mapbox.com/new-york-city-and-openstreetmap-collaborating-through-open-data-9441ebdff24>.
- 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. doi:10.2478/pcr-2020-0010
- Batini, C. & Scannapieca, M. (2006). Methodologies for Data Quality Measurement and Improvement. *Data Quality* (pp. 161-200). Springer Berlin Heidelberg. doi:10.1007/3-540-33173-5_7
- Bayazit, M. (2009). Poly Decomp. Erişim tarihi: 23 Mart 2024. Erişim adresi: <https://mpen.ca/406/bayazit>.
- Berners-Lee, T., Fielding, R. & Masinter, L. (2005). *Uniform resource identifier (URI): Generic syntax*.
- Berners-Lee, T. & Hendler, J. (2001). Publishing on the Semantic Web. *Nature*, 410(6832), 1023-1024.
- Bezděk, A. & Sebera, J. (2013). Matlab script for 3D visualizing geodata on a rotating globe. *Computers & geosciences*, 56, 127-130.

- BKG. (2023). Das BKG feierte Geburtstag. Eriřim adresi: <https://www.bkg.bund.de/DE/Das-BKG/Veranstaltungen/Jubilaem70/jubilaem70.html>.
- Bordogna, G., Carrara, P., Criscuolo, L., Pepe, M. & Rampini, A. (2016). On predicting and improving the quality of Volunteer Geographic Information projects. *International Journal of Digital Earth*, 9(2), 134-155. doi:10.1080/17538947.2014.976774
- Borkowska, S., Bielecka, E. & Pokonieczny, K. (2023). OpenStreetMap – building data completeness visualization in terms of “Fitness for purpose”. *Advances in Geodesy and Geoinformation*, article-no. doi:10.24425/agg.2022.141922
- Brassel, K., Bucher, F., Stephan, E.-M. & Vckovski, A. (1995). Completeness. *Elements of Spatial Data Quality* (pp. 81-108). Elsevier. doi:10.1016/B978-0-08-042432-3.50012-4
- Brovelli, M. A., Minghini, M.A. & Molinari, M. E. (2016). An Automated GRASS-Based Procedure to Assess the Geometrical Accuracy of the OpenStreetMap Paris Road Network. *The International Archives of the Photogrammetry, Remote Sensing and Spatial Information Sciences*, XLI-B7, 919-925. doi:10.5194/isprs-archives-XLI-B7-919-2016
- Brovelli, M. A., Minghini, M. A., Molinari, M. E. & Mooney, P. (2015). A FOSS4G-based procedure to compare OpenStreetMap and authoritative road network datasets. *Geomatics Workbooks* n° 12 – “FOSS4G Europe Como 2015”.
- Brovelli, M. A., Minghini, M. A., Molinari, M. E. & Mooney, P. (2017). Towards an Automated Comparison of OpenStreetMap with Authoritative Road Datasets. *Transactions in GIS*, 21(2), 191-206. doi:10.1111/tgis.12182
- Brovelli, M. A., Minghini, M. A., Molinari, M. E. & Zamboni, G. (2016). Positional Accuracy Assessment of the OpenStreetMap Buildings Layer through Automatic Homologous Pairs Detection: The Method and a Case Study. *ISPRS - International Archives of the Photogrammetry, Remote Sensing and Spatial Information Sciences*, XLI-B2, 615-620. doi:10.5194/isprsarchives-XLI-B2-615-2016
- Brovelli, M. A. & Zamboni, G. (2018). A New Method for the Assessment of Spatial Accuracy and Completeness of OpenStreetMap Building Footprints. *ISPRS International Journal of Geo-Information*, 7(8), 289. doi:10.3390/ijgi7080289
- Canada, G. (1997). Standards and Specifications of the National Topographic Data Base, Geomatics Canada, Canada.
- Canavosio-Zuzelski, R., Agouris, P. & Doucette, P. (2013). A Photogrammetric Approach for Assessing Positional Accuracy of OpenStreetMap© Roads. *ISPRS International Journal of Geo-Information*, 2(2), 276-301. doi:10.3390/ijgi2020276

- CEN. (1991). CEN/TC 287 - Geographic Information. Erişim adresi: <https://standards.iteh.ai/catalog/tc/cen/9db592e4-9c2e-4874-8788-b854976afd16/cen-tc-287>.
- Chausson, C. (2017). BAN: a collaborative national database of geo-referenced addresses in France. Erişim adresi: <https://joinup.ec.europa.eu/collection/open-government/document/ban-collaborative-national-database-geo-referenced-addresses-france>.
- Chow, T. E., Chien, J. & Meitzen, K. (2023). Validating the Quality of Volunteered Geographic Information (VGI) for Flood Modeling of Hurricane Harvey in Houston, Texas. *Hydrology*, 10(5), 113. doi:10.3390/hydrology10050113
- Ciepluch, B., Mooney, P., Jacob, R., Zheng, J. & Winstanely, A. (2011). Assessing the quality of open spatial data for mobile location-based services research and applications. *Archiwum Fotogrametrii, Kartografii i Teledetekcji*, 22, 105-116.
- Cipeluch, B., Jacob, R., Winstanley, A. & Mooney, P. (2010). Comparison of the accuracy of OpenStreetMap for Ireland with Google Maps and Bing Maps. Accuracy 2010 - Proceedings of the 9th International Symposium on Spatial Accuracy Assessment in Natural Resources and Environmental Sciences, University of Leicester.
- Clarke, D. G. & Clark, D. M. (1995). Lineage. *Elements of spatial data quality* (pp. 13-30). Elsevier.
- Cömert, Ç., Akinci, H., Ulutaş, D., Kara, G. & Yılmaz, C. (2008). Implementing Spatial Data Infrastructures with Semantic Web Services. INSPIRE Conference,.
- Cömert, Ç. & Akinci, H. (2004). Ulusal Konumsal Veri Altyapısı ve E-Dönüşüm Türkiye. Yapı ve Kentte Bilişim 04 KENTTe-Donuşum, e-belediyecilik, e-muhendislik ve e-kentlilik 3.Ulusal Kongresi,.
- Debattista, J., Auer, Sö. & Lange, C. (2016). Luzzu—A Methodology and Framework for Linked Data Quality Assessment. *Journal of Data and Information Quality*, 8(1), 1-32. doi:10.1145/2992786
- Debattista, J., Lange, C. & Auer, S. (2014). daQ, an Ontology for Dataset Quality Information. LDOW.
- Debattista, J., Lange, C. & Auer, S. (2015). Luzzu Quality Metric Language--A DSL for Linked Data Quality Assessment. *ArXiv preprint arXiv:1504.07758*.
- Devillers, R., Bédard, Y., Jeansoulin, R. & Moulin, B. (2007). Towards spatial data quality information analysis tools for experts assessing the fitness for use of spatial data. *International Journal of Geographical Information Science*, 21(3), 261-282. doi:10.1080/13658810600911879

- Devillers, R., Bédard, Y. & Jeansoulin, R. (2005). Multidimensional management of geospatial data quality information for its dynamic use within GIS. *Photogrammetric Engineering & Remote Sensing*, 71(2), 205-215.
- Dorn, H., Törnros, T. & Zipf, A. (2015). Quality Evaluation of VGI Using Authoritative Data—A Comparison with Land Use Data in Southern Germany. *ISPRS International Journal of Geo-Information*, 4(3), 1657-1671. doi:10.3390/ijgi4031657
- Drummond, J. (1995). Positional accuracy. *Elements of Spatial Data Quality* (pp. 31-58). Elsevier. doi:10.1016/B978-0-08-042432-3.50010-0
- 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. doi:10.1111/tgis.12210
- Edelsbrunner, H. (2011). Alpha shapes-a survey. *Tessellations in the sciences: Virtues, techniques and applications of geometric tilings*.
- El-Ashmawy, K. L. A. (2016). Testing the positional accuracy of OpenStreetMap data for mapping applications. *Geodesy and Cartography*, 42(1), 25-30. doi:10.3846/20296991.2016.1160493
- Elias, E. N., Amorim, F. R., Silva, L. S., Schmidt, M. A., Camboim, S. P. & Fernandes, V. O. (2021). A proposal for a QGIS plugin for spatio-temporal analysis of OSM data quality: the case study for the city of Salvador, Brazil. Proceedings of the Academic Track State of the Map. doi:10.11588/heidok.00030584
- Erhardt, S. (2016). Status der OpenTopoMap.
- Esmaili, R., Naseri, F. & Esmaili, A. (2013). Quality Assessment of Volunteered Geographic Information. *American Journal of Geographic Information System*, 2(2), 19-26.
- ESRI. (1988). *ESRI Shapefile Technical Description*. <https://www.esri.com/content/dam/esrisites/sitecore-archive/Files/Pdfs/library/whitepapers/pdfs/shapefile.pdf>.
- Estellés-Arolas, E. & González-Ladrón-de-Guevara, F. (2012). Towards an integrated crowdsourcing definition. *Journal of Information Science*, 38(2), 189-200. doi:10.1177/0165551512437638
- 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. doi:10.1080/13658816.2013.867495
- FGDC. (2000). Content standard for digital geospatial metadata workbook. *Version2. 0*.
- Fonte, C. C., Antoniou, V., Bastin, L., Estima, J., Arsanjani, J. J., Bayas, J.-C. L. & ... Vatseva, R. (2017). Assessing VGI data quality. *Mapping and the citizen sensor*, 137-163.

- Forghani, M. & Delavar, M. (2014). A Quality Study of the OpenStreetMap Dataset for Tehran. *ISPRS International Journal of Geo-Information*, 3(2), 750-763. doi:10.3390/ijgi3020750
- Fürber, C. (2015). Architecture of the Semantic Data Quality Management Framework (SDQM). *Data Quality Management with Semantic Technologies* (pp. 87-109). Springer.
- Fürber, C. & Hepp, M. (2010). Using SPARQL and SPIN for Data Quality Management on the Semantic Web. *Lecture Notes in Business Information Processing* (Vol. 47 LNBIP, pp. 35-46). Springer. doi:10.1007/978-3-642-12814-1_4
- Fürber, C. & Hepp, M. (2011). SWIQA - A Semantic Web information quality assessment framework. *19th European Conference on Information Systems, ECIS 2011*.
- Geisler, S., Quix, C., Weber, S. & Jarke, M. (2016). Ontology-Based Data Quality Management for Data Streams. *Journal of Data and Information Quality*, 7(4), 1-34. doi:10.1145/2968332
- Girres, J. & Touya, G. (2010). Quality Assessment of the French OpenStreetMap Dataset. *Transactions in GIS*, 14(4), 435-459. doi:10.1111/j.1467-9671.2010.01203.x
- Goodchild, M. F. (2011). Citizens as Sensors: The World of Volunteered Geography. *The Map Reader* (Vol. 69, pp. 370-378). Wiley. doi:10.1002/9780470979587.ch48
- Goodchild, M. F. & Hunter, G. J. (1997). A simple positional accuracy measure for linear features. *International journal of geographical information science*, 11(3), 299-306.
- Goodchild, M. F. & Li, L. (2012a). Assuring the quality of volunteered geographic information. *Spatial statistics*, 1, 110-120.
- Goodchild, M. F. & Li, L. (2012b). Assuring the quality of volunteered geographic information. *Spatial Statistics*, 1, 110-120. doi:10.1016/j.spasta.2012.03.002
- Google. (2017). Google Map Maker has closed. Erişim tarihi: 24 Mart 2024. Erişim adresi: <https://support.google.com/mapmaker/answer/7195127?hl=en>.
- Gould, P. (1999). *Becoming a geographer*. Syracuse University Press.
- GSI. (2023). Maps and Geospatial Information. Erişim adresi: https://www.gsi.go.jp/ENGLISH/page_e30031.html.
- Guha, R. & McCool, R. (2003). TAP: a Semantic Web platform. *Computer Networks*, 42(5), 557-577.
- Guptill, S. C. & Morrison, J. L. (2013). *Elements of spatial data quality*. Elsevier.

- Haklay, M. (2010). How Good is Volunteered Geographical Information? A Comparative Study of OpenStreetMap and Ordnance Survey Datasets. *Environment and Planning B: Planning and Design*, 37(4), 682-703. doi:10.1068/b35097
- Haklay, M., Basiouka, S., Antoniou, V. & Ather, A. (2010). How Many Volunteers Does it Take to Map an Area Well? The Validity of Linus' Law to Volunteered Geographic Information. *The Cartographic Journal*, 47(4), 315-322. doi:10.1179/000870410X12911304958827
- Haklay, M. & Weber, P. (2008). Openstreetmap: User-generated street maps. *IEEE Pervasive computing*, 7(4), 12-18.
- Harvard. (2015). Facebook – Crowd-sourced translation. Erişim adresi: <https://d3.harvard.edu/platform-digit/submission/facebook-crowd-sourced-translation/>.
- Hashemi, P. & Abbaspour, R. (2015). Assessment of Logical Consistency in OpenStreetMap Based on the Spatial Similarity Concept. *Lecture Notes in Geoinformation and Cartography* (Vol. 0, pp. 19-36). Springer. doi:10.1007/978-3-319-14280-7_2
- Helbich, M., Amelunxen, C. & Neis, P. (2012). Comparative Spatial Analysis of Positional Accuracy of OpenStreetMap and Proprietary Geodata. Proceedings of GI_Forum 2012: Geovisualization, Society and Learning,. http://koenigstuhl.geog.uni-heidelberg.de/publications/2010/Helbich/Helbich_et_al_AGILE2011.pdf.
- Herzog, T. N., Scheuren, F. J. & Winkler, W. E. (2007). *Data Quality and Record Linkage Techniques*. *Data Quality and Record Linkage Techniques* Springer New York: New York, NY. doi:10.1007/0-387-69505-2
- HGM. (2019). *Türkiye Topoğrafik Veritabanı (TOPOVT) Verisi*. <https://www.harita.gov.tr/urun/turkiye-topografik-veritabani-topovt-verisi/5>.
- HGM. (2020a). *TOPOVT Detay Grupları ve Açıklamaları*. https://www.harita.gov.tr/vektor/TOPOVT_Detay_Tanimlari_2020.pdf.
- HGM. (2020b). *TOPOVT Veri Sözlüğü*. https://www.harita.gov.tr/vektor/TOPOVT_Veri_Sozlugu_2020.pdf.
- Homburg, T. (2020). Connecting Semantic Situation Descriptions with Data Quality Evaluations—Towards a Framework of Automatic Thematic Map Evaluation. *Information*, 11(11), 532. doi:10.3390/info11110532
- Homburg, T. & Boochs, F. (2019). Situation-Dependent Data Quality Analysis for Geospatial Data Using Semantic Technologies. *Lecture Notes in Business Information Processing* (Vol. 339, pp. 566-578). Springer International Publishing. doi:10.1007/978-3-030-04849-5_49
- 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. *Lecture Notes in Business Information Processing* (Vol. 373 LNBIP, pp. 543-555). Springer. doi:10.1007/978-3-030-36691-9_46
- Hormann, K. & Agathos, A. (2001). The point in polygon problem for arbitrary polygons. *Computational geometry*, 20(3), 131-144.
- Horridge, M. & Bechhofer, S. (2011). The OWL API: A Java API for OWL Ontologies. *Semantic web*, 2(1), 11-21.
- Idehen, K. U. (2017). Semantic Web Layer Cake Tweak, Explained. *Medium.com*. Erişim tarihi: 26 Mart 2024. Erişim adresi: <https://medium.com/openlink-software-blog/semantic-web-layer-cake-tweak-explained-6ba5c6ac3fab>.
- IGN. (1997). Bulletin d'information de l'IGN – Qualité d'une base de données géographique: concepts et terminologie., N. 67.
- IGN. (2023). Identity card. Erişim adresi: <https://www.ign.fr/institut/identity-card>.
- INSPIRE. (2014). *Data Specification on Hydrography - Technical Guidelines*. European Commission Joint Research Centre (Vol. D2.8.1.8).
- INSPIRE. (2023). INSPIRE. Erişim tarihi: 21 Mart 2024. Erişim adresi: <https://inspire-geoportal.ec.europa.eu/srv/eng/catalog.search#/home>.
- ISO. (2001). *ISO 19113:2002, Geographic Information - Quality Principles. The International Organization for Standardisation*. <https://www.iso.org/standard/26018.html>.
- ISO. (2013). *ISO 19157:2013: Geographic Information - Data Quality. International Organization for Standardization (ISO) International Organization for Standardization Geneva, Switzerland*. <https://www.iso.org/standard/32575.html>.
- ISO. (2023). *ISO 19157: 2023 Geographic Information-Data quality. International Organization for Standardization Geneva, Switzerland*. <https://www.iso.org/standard/78900.html>.
- Jackson, M. (2014). A flexible framework for assessing the quality of crowdsourced data. Multiway multidimensional analysis View project ROSELT View project. AGILE Digital Editions, AGILE Digital Editions. <https://www.researchgate.net/publication/265187958>.
- Jackson, S., Mullen, W., Agouris, P., Crooks, A., Croitoru, A. & Stefanidis, A. (2013). Assessing Completeness and Spatial Error of Features in Volunteered Geographic Information. *ISPRS International Journal of Geo-Information*, 2(2), 507-530. doi:10.3390/ijgi2020507
- Jacobs, K. (2018). Quality Assessment of Volunteered Geographic Information: An Investigation into the Ottawa-Gatineau OpenStreetMap Database. . <https://curve.carleton.ca/system/files/etd/fb66a114-871d-4cac-bfb1->

092a65a28ccc/etd_pdf/8c1c30f0d837bec377efda28b59225aa/jacobs-qualityassessmentofvolunteeredgeographicinformation.pdf.

- Juhász, L. & Hochmair, H. H. (2016). User contribution patterns and completeness evaluation of Mapillary, a crowdsourced street level photo service. *Transactions in GIS*, 20(6), 925-947.
- Juhász, L., Novack, T., Hochmair, H. H. & Qiao, S. (2020). Cartographic vandalism in the era of location-based games—The case of OpenStreetMap and Pokémon GO. *ISPRS International Journal of Geo-Information*, 9(4), 197.
- Juran, J. (1967). *Management of quality control*. JM Juran.
- Kainz, W. (1995). Logical consistency. *Elements of spatial data quality*, 202, 109-137.
- Kalantari, M. & La, V. (2015). Assessing OpenStreetMap as an Open Property Map. *OpenStreetMap in GIScience: Experiences, Research, and Applications* (pp. 255-272). Springer. doi:10.1007/978-3-319-14280-7_13
- Kartverket. (2021). Terms of use. Erişim adresi: <https://www.kartverket.no/en/api-and-data/terms-of-use>.
- Kastanakis, B. (2016). *Mapbox Cookbook*. Packt Publishing Ltd.
- Keil, M. & Snoeyink, J. (2002). On the time bound for convex decomposition of simple polygons. *International Journal of Computational Geometry & Applications*, 12(03), 181-192.
- Keßler, C. & De Groot, R. T. A. (2013). Trust as a proxy measure for the quality of volunteered geographic information in the case of OpenStreetMap. *Geographic information science at the heart of Europe* (pp. 21-37). Springer.
- Keßler, C., Trame, J. & Kauppinen, T. (2011). Tracking editing processes in volunteered geographic information: The case of OpenStreetMap. Identifying objects, processes and events in spatio-temporally distributed data (IOPE), workshop at conference on spatial information theory,.
- Kounadi, O. (2009). *Assessing the quality of OpenStreetMap data* *Geographical information science*, University College of. ftp://ftp.cits.nrcan.gc.ca/pub/cartonat/Reference/VGI/Rania_OSM_dissertation.pdf
- Küçük, K. & Anbaroğlu, B. (2020). Spatial Accuracy Assessment of Buildings in Openstreetmap. *International Journal of Environment and Geoinformatics*, 7(2), 205-212. doi:10.30897/ijegeo.710508
- Külcü, Ö. (2020). Verification of information and evaluation of approaches of information professionals in accessing accurate information. *Navigating Fake News, Alternative Facts, and Misinformation in a Post-Truth World* (pp. 162-183). IGI Global.

- Laor, T. & Galily, Y. (2022). In WAZE we trust? GPS-based navigation application users' behavior and patterns of dependency. *PLOS ONE*, 17(11), e0276449. doi:10.1371/journal.pone.0276449
- Lien, J.-M. & Amato, N. M. (2004). Approximate convex decomposition of polygons. Proceedings of the twentieth annual symposium on Computational geometry,.
- LinkedGeoData. (2023). Imprint. Erişim tarihi: 16 Mart 2024. Erişim adresi: <http://linkedgeodata.org/docs/imprint.html>.
- Ludwig, I., Voss, A. & Krause-Traudes, M. (2011). A Comparison of the Street Networks of Navteq and OSM in Germany. *Advancing geoinformation science for a changing world*, 65-84.
- Mahabir, R., Stefanidis, A., Croitoru, A., Crooks, A. & Agouris, P. (2017). Authoritative and Volunteered Geographical Information in a Developing Country: A Comparative Case Study of Road Datasets in Nairobi, Kenya. *ISPRS International Journal of Geo-Information*, 6(1), 24. doi:10.3390/ijgi6010024
- MapGive. (2014). MapGive. Erişim adresi: <https://mapgive.state.gov/>.
- Mapillary. (2020). Mapillary Joins Facebook on the Journey of Improving Maps Everywhere. Erişim tarihi: 24 Mart 2024. Erişim adresi: <https://blog.mapillary.com/news/2020/06/18/Mapillary-joins-Facebook.html>.
- Mashhadi, A., Quattrone, G. & Capra, L. (2015). The impact of society on volunteered geographic information: The case of OpenStreetMap. *OpenStreetMap in GIScience: Experiences, research, and applications* (pp. 125-141). Springer.
- McEwen, R. B., Witmer, R. E. & Ramey, B. S. (1983). *USGS Digital Cartographic Data Standards*. US Department of the Interior, Geological Survey.
- Mobasheri, A. (2017). A Rule-Based Spatial Reasoning Approach for OpenStreetMap Data Quality Enrichment; Case Study of Routing and Navigation. *Sensors*, 17(11), 2498. doi:10.3390/s17112498
- Moellering, H. (1987). *A draft proposed standard for digital cartographic data*. National Committee for Digital Cartographic Data Standards Columbus, OH, USA.
- Mooney, P., Corcoran, P. & Winstanley, A. C. (2010). Towards quality metrics for OpenStreetMap. Proceedings of the 18th SIGSPATIAL International Conference on Advances in Geographic Information Systems, ACM: New York, NY, USA. doi:10.1145/1869790.1869875
- Moradi, M., Roche, S. & Mostafavi, M. A. (2023a). A Novel Feature Matching Method for Matching OpenStreetMap Buildings with Those of Reference Dataset. *Lecture Notes in Computer Science (including subseries Lecture Notes in Artificial Intelligence and Lecture Notes in Bioinformatics)* (Vol. 13912 LNCS, pp. 139-152). doi:10.1007/978-3-031-34612-5_10

- Moradi, M., Roche, S. & Mostafavi, M. A. (2023b). A Novel Feature Matching Method for Matching OpenStreetMap Buildings with Those of Reference Dataset. International Symposium on Web and Wireless Geographical Information Systems, Springer.
- Mostafavi, M.-A., Edwards, G. & Jeansoulin, R. (2004). An ontology-based method for quality assessment of spatial data bases. Third International Symposium on Spatial Data Quality,. <http://hal.inria.fr/inria-00000447>.
- Nalci, İ. (2011). *Türkiye için ulusal mekânsal veri altyapısı önerisi*.
- Nash, E., Nikkilä, R., Wiebenson, J., Walter, K. & Bill, R. (2011). Interchange of geospatial rules - towards georules interchange format (GeoRIF)?. *Gis.Science*, 24(3), 82-94.
- Nash, E., Wiebenson, J., Nikkilä, R., Vatsanidou, A., Fountas, S. & Bill, R. (2011). Towards automated compliance checking based on a formal representation of agricultural production standards. *Computers and Electronics in Agriculture*, 78(1), 28-37. doi:10.1016/j.compag.2011.05.009
- Neis, P., Zielstra, D. & Zipf, A. (2011). The street network evolution of crowdsourced maps: OpenStreetMap in Germany 2007–2011. *Future Internet*, 4(1), 1-21.
- Neis, P., Zielstra, D. & Zipf, A. (2013). Comparison of Volunteered Geographic Information Data Contributions and Community Development for Selected World Regions. *Future Internet*, 5(2), 282-300. doi:10.3390/fi5020282
- NLS. (2023). About the National Land Survey of Finland. Erişim adresi: <https://www.maanmittauslaitos.fi/en/about-nls/organisation/about-national-land-survey-finland>.
- OpenGeoEdu. (2024). OSM data model. Erişim adresi: <https://learn.opengeoedu.de/en/opendata/vorlesung/freiwillig-erhobene-daten/openstreetmap/datenmodell>.
- Ordnance Survey. (2023). OS History. Erişim adresi: <https://www.ordnancesurvey.co.uk/about/history>.
- O'Reilly, T. & Battelle, J. (2009). *Web squared: Web 2.0 five years on*. “O'Reilly Media, Inc.”
- OSM. (2022). OpenStreetMap stats. Erişim adresi: https://web.archive.org/web/20221118152110/https://planet.openstreetmap.org/statistics/data_stats.html.
- OSM Wiki. (2018). Organizations. Erişim adresi: <https://wiki.openstreetmap.org/wiki/Category:Organizations>.
- OSM Wiki. (2019). Prototype Global Shoreline. Erişim adresi: https://wiki.openstreetmap.org/wiki/Prototype_Global_Shoreline.

- OSM Wiki. (2022). Key:mapillary. Erişim tarihi: 24 Mart 2024. Erişim adresi: <https://wiki.openstreetmap.org/wiki/Key:mapillary>.
- OSM Wiki. (2023a). No:Kartverket import. Erişim adresi: https://wiki.openstreetmap.org/wiki/No:Kartverket_import.
- OSM Wiki. (2023b). GPS.
- OSM Wiki. (2024). NYC Buildings Addresses. Erişim adresi: https://wiki.openstreetmap.org/wiki/Import/Catalogue/NYC_Buildings_Addresses.
- Paiva, C. dos A. & Camboim, S. P. (2022). Inference of positional accuracy of collaborative data from intrinsic parameters. *Transactions in GIS*, 26(4), 1898-1913. doi:10.1111/tgis.12913
- Papadopoulou, C.-A. & Giaoutzi, M. (2014). Crowdsourcing as a Tool for Knowledge Acquisition in Spatial Planning. *Future Internet*, 6(1), 109-125. doi:10.3390/fi6010109
- Pfund, J. (2008). Mapping with the masses: Google Map Maker. AGU Fall Meeting Abstracts,.
- Qiu, H., Ayara, A. & Glimm, B. (2021). Ontology-Based Map Data Quality Assurance. *Lecture Notes in Computer Science (including subseries Lecture Notes in Artificial Intelligence and Lecture Notes in Bioinformatics)* (Vol. 12731 LNCS, pp. 73-89). Springer International Publishing. doi:10.1007/978-3-030-77385-4_5
- Raines, G. L. (1997). Elements of spatial data quality. *Computers & Geosciences*, 23(1), 135. doi:10.1016/S0098-3004(97)87525-5
- RCMRD. (2023). About Us. Erişim adresi: <https://rcmr.org/en/about-us>.
- Resmi Gazete. (2005). *Büyük Ölçekli Harita ve Harita Bilgileri Üretim Yönetmeliği*. (Vol. no. 25876).
- Salgé, F. (1995). Semantic accuracy. *Elements of spatial data quality* (Vol. 1, pp. 139-151). Oxford, UK: Elsevier.
- Salvucci, G. & Salvati, L. (2021). Official Statistics, Building Censuses, and OpenStreetMap Completeness in Italy. *ISPRS International Journal of Geo-Information*, 11(1), 29. doi:10.3390/ijgi11010029
- Sarretta, A. & Minghini, M. (2021). Towards the Integration of Authoritative and OpenStreetMap Geospatial Datasets in Support of the European Strategy for Data. *The International Archives of the Photogrammetry, Remote Sensing and Spatial Information Sciences*, XLVI-4/W2-2021(4/W2-2021), 159-166. doi:10.5194/isprs-archives-XLVI-4-W2-2021-159-2021

- SatImagingCorp. (2023). Albedo Satellite Constellation (10cm). Erişim adresi: <https://www.satimagingcorp.com/satellite-sensors/albedo-10cm/>.
- Sehra, S. S., Singh, J. & Rai, H. S. (2014). A Systematic Study of OpenStreetMap Data Quality Assessment. 2014 11th International Conference on Information Technology: New Generations, IEEE. doi:10.1109/ITNG.2014.115
- Senaratne, H., Mobasher, A., Ali, A. L., Capineri, C. & Haklay, M. (2017). A review of volunteered geographic information quality assessment methods. *International Journal of Geographical Information Science*, 31(1), 139-167. doi:10.1080/13658816.2016.1189556
- Servigne, S., Lesage, N., & Libourel, T. (2006). Quality Components, Standards, and Metadata. *Fundamentals of Spatial Data Quality* (pp. 179-210). Wiley. doi:10.1002/9780470612156.ch10
- Silva, L. S. L., & Camboim, S. P. (2021). Authoritative Cartography in Brazil and Collaborative Mapping Platforms: Challenges and Proposals for Data Integration. *Boletim de Ciências Geodésicas*, 27(spe), 1-19. doi:10.1590/s1982-21702021000100003
- Sehra, S. S., Singh, J. & Rai, H. S. (2013). Assessment of OpenStreetMap Data - A Review. *International Journal of Computer Applications*, 76(16), 17-20. doi:10.5120/13331-0888
- SoI. (2023). List and Pricing of Digital Products of Survey of India. Erişim adresi: <https://onlinemaps.surveyofindia.gov.in/PricingPolicy.aspx>.
- Stadler, C., Lehmann, J., Höffner, K. & Auer, S. (2012). Linked geodata: A core for a web of spatial open data. *Semantic Web*, 3(4), 333-354.
- Stark, H.-J. (2011). Quality Assessment of Volunteered Geographic Information Using Open Web Map Services Within OpenAddresses. Geoinformatics Forum Salzburg, Wichmann. <http://hdl.handle.net/11654/12725>.
- Steiner, J. (1971). *Jacob Steiner's gesammelte werke*. American Mathematical Soc.
- Sui, D., Elwood, S. & Goodchild, M. (2013). *Crowdsourcing Geographic Knowledge. Crowdsourcing Geographic Knowledge: Volunteered Geographic Information (VGI) in Theory and Practice* (D. Sui, S. Elwood ve M. Goodchild, Ed.). Springer Netherlands: Dordrecht. doi:10.1007/978-94-007-4587-2
- Sui, D. Z. (2008). The wikification of GIS and its consequences: Or Angelina Jolie's new tattoo and the future of GIS. *Computers, Environment and Urban Systems*, 32(1), 1-5. doi:10.1016/j.compenvurbsys.2007.12.001
- Touya, G. (2012). What Is the Level of Detail of OpenStreetMap ? Workshop on role of volunteered geographic information in advancing science: quality and credibility, At Columbus, OH, USA,.

https://www.researchgate.net/publication/281967547_What_Is_the_Level_of_Detail_of_OpenStreetMap.

TriMet. (2009). TriMet Planner. Erişim adresi: <https://trimet.org/home/planner>.

Turner, A. (2006). *Introduction to Neogeography*. “O’Reilly Media, Inc.”

Ullah, T., Lautenbach, S., Herfort, B., Reinmuth, M. & Schorlemmer, D. (2023). Assessing Completeness of OpenStreetMap Building Footprints Using MapSwipe. *ISPRS International Journal of Geo-Information*, 12(4), 143. doi:10.3390/ijgi12040143

Uluğtekin, N. & Bildirici, İ. Ö. (1997). Coğrafi bilgi sistemi ve harita. Kartografya. *Harita Kurultayı, Ankara*, 85.

USGS. (2023). Landsat Data. Erişim adresi: <https://landsat.gsfc.nasa.gov/data/#:~:text=Since%202008%2C%20when%20the%20USGS,products%2C%20tools%2C%20and%20applications>.

Van Exel, M., Dias, E. & Frujtier, S. (2010). The impact of crowdsourcing on spatial data quality indicators. Proceedings of the GIScience 2010 Doctoral Colloquium, Zurich, Switzerland,.

Vikiveri. (2022). Giriş. Erişim adresi: <https://www.wikidata.org/wiki/Wikidata:Introduction/tr>.

Vinet, L. & Zhedanov, A. (2011). A ‘missing’ family of classical orthogonal polynomials. *Journal of Physics A: Mathematical and Theoretical*, 44(8), 085201. doi:10.1088/1751-8113/44/8/085201

W3C. (2004). SWRL: A Semantic Web Rule Language Combining OWL and RuleML. Erişim adresi: <http://www.w3.org/Submission/SWRL/>.

W3C. (2012). Web Ontology Language (OWL). Erişim tarihi: 19 Mart 2024. Erişim adresi: Web Ontology Language (OWL).

W3C. (2014). RDF Schema 1.1. Erişim tarihi: 24 Mart 2024. Erişim adresi: <https://www.w3.org/TR/rdf-schema/>.

Wand, Y. & Wang, R. Y. (1996). Anchoring data quality dimensions in ontological foundations. *Communications of the ACM*, 39(11), 86-95. doi:10.1145/240455.240479

Wang, F., Mäs, S., Reinhardt, W. & Kandawasvika, A. (2005). Ontology Based Quality Assurance for Mobile Data Acquisition. Proceedings of the 19th International Conference on Informatics for Environmental Protection: Networking Environmental Information,.

Waze. (2024). Waze. Erişim adresi: <https://www.waze.com>.

Wikimapia. (2012). About Wikimapia. Erişim adresi: <https://wikimapia.org/about/>.

- Wikimedia. (2024). Wikipedia. Erişim adresi: <https://www.wikipedia.org/>.
- Wuebbles, D. J. (2012). Celebrating the “blue marble”. *Eos, Transactions American Geophysical Union*, 93(49), 509-510.
- Xiao-li, W. & De, Z. (2010). Selecting optimal threshold value of Douglas-Peucker algorithm based on curve fit. 2010 First International Conference on Networking and Distributed Computing, IEEE.
- Xu, X. & Cai, H. (2020). Semantic approach to compliance checking of underground utilities. *Automation in Construction*, 109, 103006. doi:10.1016/j.autcon.2019.103006
- Xu, Y., Chen, Z., Xie, Z. & Wu, L. (2017). Quality assessment of building footprint data using a deep autoencoder network. *International Journal of Geographical Information Science*, 31(10), 1929-1951. doi:10.1080/13658816.2017.1341632
- Yamashita, J., Seto, T., Iwasaki, N. & Nishimura, Y. (2023). Quality assessment of volunteered geographic information for outdoor activities: an analysis of OpenStreetMap data for names of peaks in Japan. *Geo-spatial Information Science*, 26(3), 333-345. doi:10.1080/10095020.2022.2085188
- Yilmaz, C. (2018). *A framework for ontology-based spatial data quality assessment, design and development*. Karadeniz Teknik Üniversitesi, Fen Bilimleri Enstitüsü, Doktora Tezi, Trabzon.
- Zaveri, A., Rula, A., Maurino, A., Pietrobon, R., Lehmann, J. & Auer, S. (2015). Quality assessment for Linked Data: A Survey. *Semantic Web*, 7(1), 63-93. doi:10.3233/SW-150175
- Zhang, H. & Malczewski, J. (2019). Quality Evaluation of Volunteered Geographic Information. *Crowdsourcing* (pp. 1173-1201). IGI Global. doi:10.4018/978-1-5225-8362-2.ch058
- Zhang, W.-B., Leung, Y. & Ma, J.-H. (2019). Analysis of positional uncertainty of road networks in volunteered geographic information with a statistically defined buffer-zone method. *International Journal of Geographical Information Science*, 33(9), 1807-1828. doi:10.1080/13658816.2019.1606430
- Zheng, S. & Zheng, J. (2014). Assessing the completeness and positional accuracy of OpenStreetMap in China. *Thematic cartography for the society* (pp. 171-189). Springer.
- Zhou, Q., Zhang, Y., Chang, K. & Brovelli, M. (2022). Assessing OSM building completeness for almost 13,000 cities globally. *International Journal of Digital Earth*, 15(1), 2400-2421. doi:10.1080/17538947.2022.2159550

- Zhou, Qi. (2017). Rethinking the Buffering Approach for Assessing OpenStreetMap Positional Accuracy. *Lecture Notes in Geoinformation and Cartography* (pp. 435-448). Springer Berlin Heidelberg. doi:10.1007/978-3-319-57336-6_30
- Zhou, Qi. (2018). Exploring the relationship between density and completeness of urban building data in OpenStreetMap for quality estimation. *International Journal of Geographical Information Science*, 32(2), 257-281.
- Zielstra, D. & Hochmair, H. H. (2011). Comparative study of pedestrian accessibility to transit stations using free and proprietary network data. *Transportation Research Record*, 2217(1), 145-152.
- Zielstra, D. & Hochmair, H. H. (2013). Positional accuracy analysis of Flickr and Panoramio images for selected world regions. *Journal of Spatial Science*, 58(2), 251-273. doi:10.1080/14498596.2013.801331
- Zielstra, D. & Zipf, A. (2010). A comparative study of proprietary geodata and volunteered geographic information for Germany. .
http://agile2010.dsi.uminho.pt/pen/shortpapers_pdf/142_doc.pdf.

ÖZGEÇMİŞ

Deniz YILDIRIM, ortaöğretim ve lise eğitimini İstanbul Lisesi'nde tamamlamıştır. 1997 yılında, Bilkent Üniversitesi Matematik Bölümü'ne burslu olarak kabul edilmiş ve lisans eğitimine başlamıştır. 2001 yılında, bölümünden üçüncülük derecesiyle mezun olmuştur. Aynı yıl, Purdue Üniversitesi (ABD) Matematik Bölümü'nde doktora programına başlamış ve 2001-2008 yılları arasında bu kurumda lisansüstü araştırma görevlisi olarak çalışmıştır. Bu süre zarfında, çeşitli seviyelerde Calculus ve Cebir dersleri vermiştir. 2008 ve 2009 yıllarında, yurt dışı eğitim danışmanlığı alanında danışmanlık görevlerinde bulunmuştur. 2009 yılında, Karadeniz Teknik Üniversitesi (KTÜ) Fen Bilimleri Enstitüsü'nde Harita Mühendisliği Uzaktan Algılama Anabilim Dalı'nda yüksek lisans programına başlamış ve aynı yıl KTÜ Mühendislik Fakültesi'nde Araştırma Görevlisi olarak göreve başlamıştır. Harita Mühendisliği anabilim dalından yüksek lisans derecesini almıştır. İngilizce ve Almanca dillerine hakimiyeti, çok iyi seviyededir.