

**KARADENİZ TEKNİK ÜNİVERSİTESİ
FEN BİLİMLERİ ENSTİTÜSÜ**

HARİTA MÜHENDİSLİĞİ ANABİLİM DALI

**COVID-19 VERİLERİNİN BAĞLANTILI VERİ OLARAK
YAYINLANMASI VE SORGULANMASI**

YÜKSEK LİSANS TEZİ

Hatice Dilan KARABULUT

**EYLÜL 2022
TRABZON**



KARADENİZ TEKNİK ÜNİVERSİTESİ
FEN BİLİMLERİ ENSTİTÜSÜ

HARİTA MÜHENDİSLİĞİ ANABİLİM DALI

**COVID-19 VERİLERİNİN BAĞLANTILI VERİ OLARAK
YAYINLANMASI VE SORGULANMASI**

Hatice Dilan KARABULUT

**Karadeniz Teknik Üniversitesi Fen Bilimleri Enstitüsünce
"HARİTA YÜKSEK MÜHENDİSİ"
Unvanı Verilmesi İçin Kabul Edilen Tezdir.**

Tezin Enstitüye Verildiği Tarih : 11 / 08 / 2022

Tezin Savunma Tarihi : 14 / 09 / 2022

Tez Danışmanı : Dr. Öğr. Üyesi Gülten KARA

Trabzon 2022

ÖNSÖZ

Karadeniz Teknik Üniversitesi Harita Mühendisliği Anabilim Dalında hazırlanan “COVID-19 Verilerinin Bağlantılı Veri Olarak Yayınlanması ve Sorgulanması” konulu tezin ortaya çıkış sürecinde birçok kişinin önemli katkıları olmuştur. Öncelikle bilgi ve tecrübesini benimle paylaşan ve çok değerli zamanımı benden esirgemeyen tez danışmanım Dr. Öğretim Üyesi Gülten KARA’ya çok teşekkür ederim. Değerli katkı ve eleştirileri için jüri üyelerine teşekkürlerimi sunarım. Trabzon İl Sağlık Müdürlüğü’ne ve Bulaşıcı Hastalıklar Birimi’ndeki Ferhat KÖSE’ye katkılarından dolayı teşekkür ederim. Benden hiçbir konuda desteğini esirgemeyen ve bugünlere gelmemde sonsuz emekleri olan aileme teşekkürü bir borç bilirim.

Hatice Dilan KARABULUT
Trabzon, 2022

TEZ ETİK BEYANNAMESİ

Yüksek Lisans Tezi olarak sunduğum “COVID-19 Verilerinin Bağlantılı Veri Olarak Yayınlanması ve Sorgulanması” başlıklı bu çalışmayı baştan sona kadar danışmanım Dr. Öğretim Üyesi Gülten KARA’ın sorumluluğunda tamamladığımı, verileri kendim topladığımı, analizleri ilgili laboratuvarlarda yaptığımı, 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. 14/09/2022

Hatice Dilan KARABULUT

İÇİNDEKİLER

| | <u>Sayfa No</u> |
|---|-----------------|
| ÖNSÖZ..... | III |
| TEZ ETİK BEYANNAMESİ..... | IV |
| İÇİNDEKİLER..... | V |
| ÖZET | IX |
| SUMMARY | X |
| ŞEKİLLER DİZİNİ | XI |
| TABLolar DİZİNİ..... | XIV |
| KISALTMALAR DİZİNİ | XV |
| 1. GENEL BİLGİLER | 1 |
| 1.1. Giriş..... | 1 |
| 1.2. Literatür Araştırması | 2 |
| 1.2.1. COVID-19 Açık Veri Kaynakları | 2 |
| 1.2.2. COVID-19 Verilerinin Semantik Tanımlarının Oluşturulması | 4 |
| 1.3. Tezin Amacı ve Kapsamı | 8 |
| 1.4. Semantik Web | 9 |
| 1.5. Semantik Web Bileşenleri..... | 10 |
| 1.5.1. URI/IRI | 11 |
| 1.5.2. XML..... | 14 |
| 1.5.3. RDF..... | 15 |
| 1.5.4. RDFS..... | 20 |
| 1.5.5. RDF Graph..... | 21 |
| 1.5.6. Ontoloji | 22 |
| 1.5.7. Ontoloji Dilleri..... | 24 |
| 1.5.7.1. RDF/RDFS..... | 24 |
| 1.5.7.2. SKOS | 24 |
| 1.5.7.3. DAML+OIL..... | 25 |
| 1.5.7.4. OWL | 25 |
| 1.5.8. SPARQL-GeoSPARQL..... | 27 |
| 1.5.9. Unifying Logic..... | 31 |

| | | |
|----------|--|----|
| 1.5.10. | Proof..... | 31 |
| 1.6. | Bağlantılı Veri..... | 31 |
| 1.7. | Konumsal Bağlantılı Veri | 36 |
| 1.8. | COVID-19 Bağlantılı Veri Uygulamaları..... | 36 |
| 1.9. | Bağlantılı Veri Teknolojileri..... | 42 |
| 1.9.1. | RDF Veri Modeline Dönüşüm Araçları..... | 42 |
| 1.9.1.1. | Geometry2RDF..... | 43 |
| 1.9.1.2. | Shp2GeoSPARQL | 43 |
| 1.9.1.3. | GeomRDF..... | 43 |
| 1.9.1.4. | TripleGeo | 43 |
| 1.9.1.5. | GeoTriples..... | 44 |
| 1.9.1.6. | OpenRefine RDF Extension | 45 |
| 1.9.1.7. | Karma – A Data Integration Tool | 45 |
| 1.9.2. | SPARQL Endpoint..... | 46 |
| 1.9.2.1. | OpenLink Virtuoso | 46 |
| 1.9.2.2. | Fuseki..... | 47 |
| 1.9.2.3. | D2R Server..... | 47 |
| 1.9.2.4. | 4store SPARQL Server | 47 |
| 1.9.2.5. | PublishMyData | 47 |
| 1.9.3. | Eşleştirme Araçları..... | 48 |
| 1.9.3.1. | Silk-The Linked Data Integration Framework..... | 48 |
| 1.9.3.2. | LIMES (Link Discovery Framework For Metric Spaces) | 48 |
| 1.9.3.3. | The xCurator Project..... | 49 |
| 1.9.4. | Bağlantılı Verinin Yayınlanması | 49 |
| 1.9.4.1. | Pubby | 49 |
| 1.9.4.2. | CKAN | 50 |
| 1.9.4.3. | DataHub | 50 |
| 1.9.4.4. | LOD Cloud..... | 50 |
| 1.9.5. | Konumsal Üçlü Depoları | 50 |
| 1.9.5.1. | Parliament | 51 |
| 1.9.5.2. | Strabon | 51 |
| 1.9.5.3. | Fuseki..... | 52 |
| 1.9.6. | Bağlantılı Verinin Doğrulanması | 52 |

| | | |
|----------|---|----|
| 1.9.6.1. | RDFSShape-RDF Playground | 52 |
| 1.9.6.2. | Vapour..... | 52 |
| 1.9.6.3. | W3 RDF Validator | 52 |
| 2. | YAPILAN ÇALIŞMALAR | 53 |
| 2.1. | TURCOVID19 Verilerinin Elde Edilmesi..... | 54 |
| 2.2. | TURCOVID19 Verilerinin Semantik Tanımları İçin Ontolojilerin Değerlendirilmesi..... | 55 |
| 2.2.1. | CODO | 55 |
| 2.2.2. | ROC Ontology | 57 |
| 2.2.3. | WHOnto..... | 60 |
| 2.3. | TURCOVID19 Verileri İçin Kullanılacak Ontolojilerin Belirlenmesi..... | 62 |
| 2.4. | TURCOVID19 Ontolojisinin Geliştirilmesi | 63 |
| 2.4.1. | Amacın Tanımı (Definition of Purpose)..... | 64 |
| 2.4.2. | Yetkinlik Sorularının Tanımı (Derivation of Competency Questions)..... | 64 |
| 2.4.3. | Kavramların Çıkarılması (Term Extraction)..... | 65 |
| 2.4.4. | Analiz (Analysis) | 65 |
| 2.4.5. | Bilgi Sentezi (Knowledge Synthesis) | 67 |
| 2.4.6. | Yeniden Kullanım ve Standardizasyon (Reuse and standardization) | 67 |
| 2.4.7. | Temsili Modelin Tasarımı (Design of Representational Model)..... | 68 |
| 2.4.8. | Ontoloji geliştirme (Ontology development)..... | 68 |
| 2.4.8.1. | Sınıflar (Classes)..... | 69 |
| 2.4.8.2. | Özellikler (Object-Data Properties) | 69 |
| 2.4.9. | Değerlendirme (Evaluation)..... | 78 |
| 2.5. | COVID-19 Verilerinin Semantik Tanımlarının Oluşturulması | 79 |
| 2.6. | RDF Linklerinin Kurulması | 82 |
| 2.6.1. | Workspace..... | 83 |
| 2.6.2. | Linkage Rule Editör | 85 |
| 2.6.3. | Generating Links..... | 87 |
| 2.7. | Bağlantılı Verinin Yayınlanması | 88 |
| 3. | BULGULAR VE İRDELEMELER..... | 89 |
| 3.1. | COVID-19 Verilerinin Sorgulanması ve Karşılaştırılması..... | 90 |
| 3.2. | Triple Store/RDF Store ile İlişkisel Veritabanlarının Karşılaştırılması..... | 96 |
| 3.3. | TURCOVID19 Verileri İçin Belirlenen Ontolojilerin İrdelenmesi | 98 |

| | | |
|------|---|-----|
| 3.4. | Mevcut COVID-19 Baęlantılı Verilerin İrdelenmesi | 100 |
| 3.5. | Farklı Veri Setleri ile Oluřturulan Linklerin İrdelenmesi | 101 |
| 3.6. | TURCOVID19 Baęlantılı Verilerinin Sorgulanması..... | 109 |
| 4. | SONUÇLAR VE ÖNERİLER | 117 |
| 5. | KAYNAKLAR | 120 |

ÖZGEÇMİŐ



Yüksek Lisans

ÖZET

COVID-19 VERİLERİNİN BAĞLANTILI VERİ OLARAK YAYINLANMASI VE
SORGULANMASI

Hatice Dilan KARABULUT

Karadeniz Teknik Üniversitesi
Fen Bilimleri Enstitüsü
Harita Mühendisliği Anabilim Dalı
Danışman: Dr. Öğretim Üyesi Gülten KARA
2022, 129 Sayfa

Günümüzde COVID-19 pandemisi ile başa çıkmaya yardımcı olmak ve gelecekteki olası krizlere hazırlanmak için mevcut durumun değerlendirilmesi ve analiz gerçekleştirilmesine yönelik araçlar önem kazanmaktadır. Bu bağlamda birçok kuruluş, COVID-19 salgının izlenmesine olanak tanıyan virüsün etkisine ilişkin küresel verileri toplamak ve yayınlamak için sistemler ve projeler geliştirmiştir. Uluslararası düzeyde COVID-19 verileri web üzerinde belirli bir kapsamda açık veri olarak yayınlanmıştır. Ulusal düzeyde ise Sağlık Bakanlığı, günlük olarak tablolar halinde COVID-19 verilerini paylaşmıştır. Ayrıca TURCOVID19 web sayfasında, Sağlık Bakanlığı tarafından yayınlanan COVID-19 verileri tablolar halinde düzenlenmiştir ve açık veri olarak paylaşılmıştır. Sayfada bu tablolar istatistiksel grafikler ve bilgiler herkesin kullanıma açıktır. Böyle büyük bir veri topluluğunda yer alan verilerin Semantik Web Teknolojileri kullanılarak semantik tanımlarının oluşturulması çok önemli ve gereklidir. Açık veri olarak sunulan COVID-19 verileri kullanılarak mevcut verilerden yeni bilgi ve verilerin çıkarılması için semantik tanımların oluşturulması gerekir. COVID-19 verilerinin salgının tam olarak anlaşılması, tüm verilerin değerlendirilerek sonuçlarına hâkim olmayı gerektirir. Özellikle şema ve veri düzeyinde semantik tanımların ve eşleştirmelerin yapılması yeni verilerin çıkarılması için en önemli gereksinimdir. Semantik tanımların oluşturulması için “belirli bir alandaki kavramların kavramsallaştırılması” olarak tanımlanan ontolojiler geliştirilir. COVID-19 verilerinin semantik tanımlarının oluşturulması için mevcut ontolojiler değerlendirilerek TURCOVID19 ontolojisi geliştirilmiştir. Geliştirilen TURCOVID19 Ontolojisi ile TURCOVID19 verileri Karma 2.2 yazılımı ile ilişkilendirilmiştir. TURCOVID19 açık verilerinin semantik tanımları RDF olarak yayınlanmıştır. Silk Link Discovery Framework ile mevcut farklı Web veri kaynakları içindeki varlıklar arasında bağlantılar oluşturulmuştur. Bu RDF bağlantıları, LOD bulutundaki orijinal RDF COVID-19 veri kümesiyle birlikte yayınlanabilir. Fuseki üçlü deposu, veri depolama ve SPARQL endpoint için kullanılmıştır. SPARQL query arabirimi kullanılarak RDF COVID-19 veri kümesinin sorgulanması ve semantik tanımlar ile mevcut sınıflar ve ilişkiler kullanılarak yeni bilgilerin çıkarılması sağlanmıştır.

Anahtar Kelimeler: COVID-19, Ontoloji, Semantik Web Teknolojileri, RDF, TURCOVID19.

Master Thesis

SUMMARY

PUBLISHING AND QUERYING COVID-19 DATA AS LINKED DATA

Hatice Dilan KARABULUT

Karadeniz Technical University
The Graduate School of Natural and Applied Sciences
Geomatics Engineer Graduate Program
Supervisor: Assistant Prof. Gülten KARA
2022, 129 Pages

Today, tools for evaluating the current situation and performing analysis are gaining importance in order to help cope with the COVID-19 pandemic and prepare for possible future crises. In this context, many organizations have developed systems and projects to collect and publish global data on the impact of the virus, allowing monitoring of the COVID-19 pandemic. At the international level, COVID-19 data has been published on the web as open data to a certain extent. At the national level, the Ministry of Health has shared COVID-19 data in tables on a daily basis. In addition, COVID-19 data published by the Ministry of Health are organized into tables and shared as open data on the TURCOVID19 web page. These tables are available to everyone as statistical graphics and information on the page. It is significant and necessary to create semantic definitions of the data in such a large data collection using Semantic Web Technologies. Using COVID-19 data presented as open data, semantic definitions need to be created in order to extract new information and data from existing data. A full understanding of the COVID-19 data of the epidemic requires evaluating all data and mastering the results. Especially, making semantic definitions and mappings is the most important requirement for inferring new data at the schema and data level. Ontologies which are defined as “the conceptualization of concepts in a particular field” are developed for the creation of semantic definitions. The TURCOVID19 ontology was developed by evaluating existing ontologies to create semantic definitions of COVID-19 data. Developed TURCOVID19 Ontology and TURCOVID19 data were associated with Karma 2.2 software. Semantic descriptions of TURCOVID19 open data are published as RDF. Links are created between Silk Link Discovery Framework and entities in different existing Web data sources. These RDF links can be published with the original RDF COVID-19 dataset in the LOD cloud. Fuseki triple store is used for data storage and SPARQL endpoint. Using the SPARQL query interface, the RDF COVID-19 dataset was queried and new information was extracted using semantic definitions and existing classes and relationships.

Key Words: COVID-19, Ontology, Semantic Web Technologies, RDF, TURCOVID19.

ŞEKİLLER DİZİNİ

Sayfa No

| | | |
|-----------|---|----|
| Şekil 1. | Semantik Web yapısı | 11 |
| Şekil 2. | URI, URL ve URN ilişkisi | 12 |
| Şekil 3. | Bir kaynağa ait URI ve tanımlayıcı dokümanı | 13 |
| Şekil 4. | 303 URI Çözümü..... | 13 |
| Şekil 5. | İçerik uzlaştırma ve hash URI çözümü..... | 14 |
| Şekil 6. | Basit XML kod örneği..... | 15 |
| Şekil 7. | RDF üçlüsü | 16 |
| Şekil 8. | RDF çizgesi örneği | 17 |
| Şekil 9. | RDF/XML serilizasyonunda RDF üçlüsü | 18 |
| Şekil 10. | N-Triples serilizasyonunda RDF üçlüsü..... | 18 |
| Şekil 11. | Turtle serilizasyonunda RDF üçlüsü | 18 |
| Şekil 12. | JSON-LD serilizasyonunda RDF üçlüsü..... | 19 |
| Şekil 13. | RDFa içeren bir HTML belgesi örneği | 20 |
| Şekil 14. | RDFS örneği..... | 21 |
| Şekil 15. | RDF çizgesi örneği RDF çizgesine ait RDF/XML serilizasyonu | 22 |
| Şekil 16. | SPARQL örnek sorgusu | 29 |
| Şekil 17. | Üçlü ifadenin tablosal sunumu | 34 |
| Şekil 18. | RDF verilerinin grafik sunumu | 35 |
| Şekil 19. | LOCD Sisteminin çok katmanlı mimarisi | 37 |
| Şekil 20. | Covid-on-the-Web overview | 38 |
| Şekil 21. | OpenKG-COVID19'a genel bakış | 40 |
| Şekil 22. | Wikidata'daki COVID-19 ile ilgili bilgilerin veri modeli | 41 |
| Şekil 23. | Türkiye İBBS-1 bölgeleri | 55 |
| Şekil 24. | CODO sınıfları ve arasındaki ilişkiler | 56 |
| Şekil 25. | Alınan yanıtı sahip istatistiklerin değerleri örneği..... | 58 |
| Şekil 26. | OXCGRY yanıtını temsil eden veri özelliklerinin hiyerarşisi | 58 |
| Şekil 27. | ROC Ontolojisi sınıfları ve arasındaki ilişkiler | 59 |
| Şekil 28. | WHOnto Ontolojisi | 61 |
| Şekil 29. | WHOnto veri özellikleri hiyerarşisi | 61 |

| | |
|---|----|
| Şekil 30. CODO Ontolojisi geliştirme sürecinin adımları | 63 |
| Şekil 31. İBBS Bölgeleri'ne ait sınıflar ve hiyerarşik gösterimi | 67 |
| Şekil 32. TURCOVID19 Ontolojisi ontograph gösterimi | 68 |
| Şekil 33. TURCOVID19 Ontolojisi'ne ait sınıflar | 69 |
| Şekil 34. Ontoloji sınıfları ve aralarındaki ilişkileri temsil eden nesne özellikleri..... | 70 |
| Şekil 35. hasAgeGroup ilişkisinin gösterimi | 70 |
| Şekil 36. Contains ilişkisinin gösterimi | 71 |
| Şekil 37. Veri özelliklerinin domain-range ve annation alanlarındaki ilişkilerin gösterimi | 72 |
| Şekil 38. İllere ait yapılan toplam doz aşısı sayısına göre 100 bin kişideki haftalık vaka sayısı sorgusu | 79 |
| Şekil 39. Karma ara yüzünde covid19_gunlukveriler.csv dosyasının görüntülenmesi | 79 |
| Şekil 40. Karma ara yüzünde semantik tiplerin tanımlanması | 80 |
| Şekil 41. Karma ara yüzünde semantik ilişkilerin tamamlanması..... | 81 |
| Şekil 42. Karma ara yüzünde RDF dosyasının oluşturulması | 81 |
| Şekil 43. RDF dosyası | 82 |
| Şekil 44. LOD cloud üzerinden yayınlanan mevcut COVID-19 veri kaynakları | 83 |
| Şekil 45. Prefixlerin tanımı..... | 84 |
| Şekil 46. SPARQL Endpoint ile veri kaynağının tanımlanması..... | 84 |
| Şekil 47. RDF dump ile veri kaynağının tanımlanması..... | 85 |
| Şekil 48. Linking Task gösterimi..... | 85 |
| Şekil 49. Path operatörünün gösterimi..... | 86 |
| Şekil 50. Levenshtein Distance karşılaştırma operatöründe bağlantı tanımı..... | 87 |
| Şekil 51. Eşleştirmelerin belirlenmesi | 87 |
| Şekil 52. Dosya yükleme | 88 |
| Şekil 53. SPARQL sorgu dili ile sorgulama..... | 89 |
| Şekil 54. RDF üçlüsü sorgulama sonucu | 89 |
| Şekil 55. Ardahan iline ait günlük 1. ve 2. doz sayılarını gösteren SPARQL sorgusu | 91 |
| Şekil 56. Ardahan iline ait günlük 1. ve 2. doz sayılarını gösteren Excel sorgu sonucu..... | 91 |
| Şekil 57. Ege Bölgesi'ne ait vaka kayıtları sorgu örneği..... | 93 |
| Şekil 58. Excel sorgu sonucu | 93 |
| Şekil 59. SPARQL sorgusu sonucu | 94 |
| Şekil 60. Bölgelere göre yeni hasta sayısı SPARQL sorgusu sonucu | 95 |

| | |
|--|-----|
| Şekil 61. Excel’de bölgelere göre yeni hasta sayısı sorgu sonucu..... | 96 |
| Şekil 62. Yaş ve cinsiyet gruplarına ait ölüm oranları sorgu sonucu | 98 |
| Şekil 63. Doğu Karadeniz Bölgesi’ne ait istatistikler ve kapsadığı alt bölge ve illeri gösteren sorgu sonucu | 100 |
| Şekil 64. Haftalık vaka sayıları verilerindeki il isimleri ile NUTS bölge birimleri isimleri arasındaki bağlantı sonuçları | 102 |
| Şekil 65. owl:sameAs ilişkisi ile farklı veri setlerindeki bilgilerin gösterimi | 103 |
| Şekil 66. Silk ile belirlenen eşleştirme sonuçları..... | 104 |
| Şekil 67. Link editör ile bağlantıların tanımlanması..... | 104 |
| Şekil 68. Dbpedia Pandemic Sınıfı ile TURCOVID19 sınıflarının eşleştirilmesi..... | 105 |
| Şekil 69. Dbpedia Pandemic Sınıfı ile TURCOVID19 sınıflarının eşleştirilmesi..... | 105 |
| Şekil 70. Wikidata COVID-19 (Q84263196) sınıfı ile TURCOVID19 sınıflarının eşleştirilmesi..... | 106 |
| Şekil 71. Wikidata COVID-19 (Q84263196) sınıfı..... | 106 |
| Şekil 72. Dbpedia ve TURCOVID19 ontojilerindeki data propertylerin eşleştirilmesi..... | 107 |
| Şekil 73. Covid-on-the-web ve TURCOVID19 ontojisindeki datapropertylerin eşleştirilmesi | 107 |
| Şekil 74. Karma ile URI referansının yapılması..... | 108 |
| Şekil 75. URI referansının özne (subject) durumunda gösterimi | 108 |
| Şekil 76. NumberOfProvincialLevelCases.ttl dosyasının gösterimi | 109 |
| Şekil 77. Türkiye ve dünya geneli COVID-19 verileri SPARQL sorgusu..... | 110 |
| Şekil 78. WHO Coronavirus (COVID-19) Dashboard sayfası..... | 111 |
| Şekil 79. COVID-19 bağlantılı verilerinin sorgulanması | 112 |
| Şekil 80. Centres for Disease Control and Prevention Web Sitesi | 112 |
| Şekil 81. Pandemi ile ilgili webde bulunan belgeler | 113 |
| Şekil 82. Trabzon iline ait Bağlantılı Verilerin sorgulanması | 114 |
| Şekil 83. Trabzon iline ait genel bilgiler..... | 114 |
| Şekil 84. Trabzon iline ait nüfus ve konum bilgileri | 115 |
| Şekil 85. İllere ait bağlantılı verilerin sorgulanması..... | 116 |
| Şekil 86. İllere ait geometri bilgileri..... | 116 |

TABLULAR DİZİNİ

| | <u>Sayfa No</u> |
|---|------------------------|
| Tablo 1. RDF, RDFS, OWL dillerinin karşılaştırılması | 27 |
| Tablo 2. RDF veri modeline dönüşüm araçlarının karşılaştırılması..... | 46 |
| Tablo 3. TURCOVID19 verilerinin CODO'daki Karşılığı | 57 |
| Tablo 4. TURCOVID19 verilerinin ROC Ontolojisi'ndeki karşılığı | 60 |
| Tablo 5. TURCOVID19 verilerinin WHOnto'daki karşılığı | 62 |
| Tablo 6. TURCOVID19 Ontolojisi'nin farklı düzeydeki sınıfları | 65 |
| Tablo 7. Sınıflara ait veri özelliklerinin gösterilmesi | 72 |

KISALTMALAR DİZİNİ

| | |
|-----------|---|
| ACTA | Argumentative Clinical Trial Analysis |
| API | Application Programming Interface |
| CBS | Coğrafi Bilgi Sistemi |
| CIDO | Coronavirus Infectious Disease Ontology |
| CKAN | Comprehensive Knowledge Archive Network |
| CODO | The COviD-19 Ontology |
| CORD-19 | COVID-19 Open Research Dataset |
| CSS | Cascading Style Sheets |
| CSSE | Johns Hopkins Coronavirus Resource Center |
| CSV | Comma Separated Values |
| DAML | DARPA Agent Markup Language |
| DO | Human Disease Ontology |
| DSÖ | Dünya Sağlık Örgütü |
| ECDC | European Centre for Disease Prevention and Control |
| ETL | Extract- Transform- Load |
| GeoSPARQL | A Geographic Query Language for RDF Data |
| GML | Geography Markup Language |
| HTML | Hyper Text Markup Language |
| http | Hyper Text Transfer Protocol |
| https | Secure Hypertext Transfer Protocol |
| IDO | The Infectious Disease Ontology |
| ILOSTAT | The Leading Source of Labour Statistic |
| IRI | Internationalized Resource Identifier |
| İBBS | Türkiye İstatistiki Bölge Birimleri Sınıflandırması |
| JSON | JavaScript Object Notation |
| KML | Keyhole Markup Language |
| LOCD | Linked Open COVID-19 Data |
| LOD | Linked Open Data |
| OGC | Open Geospatial Consortium |

| | |
|------------|--|
| OIL | Ontology Inference Layer |
| OS | Ordnance Survey |
| OWL | Web Ontology Language |
| OxCGRT | Oxford COVID-19 Government Response Tracker |
| PDF | Portable Document Format |
| RDF | Resource Description Framework |
| RDFa | Resource Description Framework in Attributes |
| RDFS | Resource Description Framework Schema |
| RIF | Rule Interchange Format |
| RML | RDF Mapping Language |
| ROC | An Ontology For Country Responses Towards COVID-19 |
| SGML | Standard Generalized Markup Language |
| SHACL | Shapes Constraint Language |
| ShEx | Shape Expressions |
| SKOS | Simple Knowledge Organization System |
| SPARQL | SPARQL Protocol and RDF Query Language |
| SQL | Structured Query Language |
| stRDF | spatial/temporal RDF |
| stSPARQL | SPARQL for stRDF |
| TSV | Tab Separated Value |
| TURCOVID19 | Türkiye COVID-19 Pandemi İzlem Ekranı |
| UCS | Universal Character Set |
| URI | Uniform Resource Identifier |
| URL | Uniform Resource Locator |
| URN | Uniform Resource Name |
| VTYS | Veri Tabanı Yönetim Sistemi |
| W3C | World Wide Web Consortium |
| WGS | World Geodetic System |
| WHO | World Health Organization |
| WKT | Well-known Text |
| XHTML | Extensible Hypertext Markup Language |
| XML | Extensible Markup Language |
| YAMO | Yet Another Methodology for Ontology |

1. GENEL BİLGİLER

1.1. Giriş

Bağlantılı veri prensipleri ile açık ve tekrar kullanılabilir şekilde kamu kuruluşları ve kişiler verilerini yayınlamaktadır. Geleneksel Web’de bulunun bilgilerin çoğunluğu metinsel kaynaklıdır. Metinsel kaynaklı olan bilgilerin anlamlandırılması, bilgisayarlar tarafından anlaşılması ve aralarında ilişki kurulmasında sınırlılıklar ortaya çıkmaktadır. Bu nedenle, aralarında ilişki kurulamayan bilgilerden, otomatik yeni bilgi çıkarsanması da imkânsız hale gelmektedir. Web’in günümüzdeki hacmi ve büyüme hızı dikkate alındığında artan bilgi ve içeriklerin sadece insanlar tarafından değil, bilgisayarlar tarafından da anlaşılmaya başlanması zorunluluk haline gelmiştir. Arama motorları incelendiğinde Semantik Web ile Geleneksel Web arasındaki en önemli fark ortaya çıkmaktadır. Kullanıcı, anlamsal arama motoruna konu ile bazı anahtar kelime girmek yerine doğrudan sisteme sorusunu yöneltebilmektedir. Semantik Web ile sahip olunan bu yeteneklerin ardından Bağlantılı Veri (Linked Data) kavramı ortaya çıkmıştır. Linked Data sayesinde her bilgi belirli bir anlama sahip olacak şekilde modellenir, daha sonra da modellenmiş bu bilgiler birbirleriyle ilişkilendirilir. Böylece, bilgilerin birbirine bağlanmasıyla akıllı bir veri tabanı tasarımı oluşmaktadır.

Sağlık hizmetlerinde daha iyi planlama ve karar verme için değerli bilgiler edinmelerini desteklemek amacıyla Web’de çeşitli COVID-19 veri kümeleri yayınlanmaktadır. Web’de farklı kuruluşlar tarafından üretilen COVID-19 veri kümeleri tek tip veri gösterimine dikkat edilmeden farklı adlandırma kurallarına ve yapılarla sahip tablosal biçimlerde (CSV, Excel, PDF vb.) temsil edilmekte ve yayınlanmaktadır. Ancak, veri kümeleri, veri keşfini ve yeniden kullanımını zorlaştıran heterojen formatlarda üretilir. COVID-19 veri kümelerinin bu çeşitliliği paydaşlar tarafından etkin planlama ve karar verme için veri yönetimi, paylaşım ve entegrasyon konusunda ciddi zaman alıcı ve hataya açık sorunlar ortaya çıkarmaktadır.

COVID-19’a ait sağlık uygulamalarında önemli bilgileri ortaya çıkarmak ve bunları etkili bilgiye dönüştürüp yeniden kullanımını sağlamak için daha esnek veri temsili, analizi, sorgulama ve görselleştirme yöntemlerine ihtiyaç vardır. Semantik Web Teknolojileri (SWT) yardımıyla COVID-19 pandemisine ait bilgilerin tek tip yapı temsili

için RDF formatına dönüştürülerek Linked Data olarak yayınlanması araştırmacılara sağlık sektöründe kaliteli bilgi ve gelişmiş yöntemler üretme fırsatları sağlayacaktır.

1.2. Literatür Araştırması

Tez çalışması kapsamında ülkemizdeki COVID-19 verilerinin semantik tanımlarının oluşturulması ve bu verilerin sorgulanması hedeflenmiştir. Bu bağlamda öncelikle ilgili literatürdeki proje, girişim ve akademik çalışmalar incelenmiştir. Uluslararası ve ulusal düzeyde COVID-19 verilerinin açık veri olarak yayınlandığı platformlar incelenmiştir. Uluslararası düzeyde Johns Hopkins Coronavirus Resource Center, European Centre for Disease Prevention and Control, Oxford COVID-19 Government Response Tracker, World Health Organization (WHO) Coronavirus (COVID-19) Dashboard Data Explorer, The Leading Source of Labour Statistic veri platformları tarafından, ulusal düzeyde Sağlık Bakanlığı tarafından yayınlanan COVID-19 verileri ve proje yöneticisi Abdullah UÇAR ve çalışma takımı tarafından oluşturulan TURCOVID19 veri platformu incelenmiştir.

1.2.1. COVID-19 Açık Veri Kaynakları

Mevcut durumda COVID-19 verilerini açık veri olarak sunan platformlar değerlendirilmiştir. Uluslararası ve ulusal düzeyde Covid-19 verilerinin açık veri olarak paylaşıldığı platformlar aşağıda verilmiştir.

European Centre for Disease Prevention and Control (ECDC -Avrupa Hastalık Önleme ve Kontrol Merkezi): ECDC, Avrupa'nın bulaşıcı hastalıklara karşı savunmasını güçlendirmeyi amaçlayan bir Avrupa Birliği ajansıdır ve enfeksiyon oranları hakkında veri sağlamaktadır. Temel işlevleri gözetim, salgın istihbarat, müdahale, bilimsel tavsiye, mikrobiyoloji, hazırlık, halk sağlığı eğitimi, uluslararası ilişkiler, sağlık iletişimi ve bilimsel dergi Eurosurveillance vb. geniş bir faaliyet yelpazesini kapsamaktadır. Ayrıca ECDC, COVID-19 hastalığının ülkeler üzerindeki etkinliğini izler ve hakkında bilgi sağlamaktadır (URL-1).

The Leading Source of Labour Statistic (ILOSTAT-İşgücü İstatistiklerinin Önde Gelen Kaynağı): COVID-19 salgını, birçok pratik ve kavramsal zorluğun yanı sıra işgücü piyasaları ve çalışma dünyası üzerindeki etkiyi tanımlamak için gelişmiş verilere yönelik

büyük bir talep yarattı (URL-2). ILOSTAT, kullanıcılara işgücüne ilgili çok çeşitli konularda kapsamlı uluslararası veriler sağlar. Veri serisinde işgücüne katılım, istihdam (ekonomik faaliyet, meslek, ekonomik sınıf ve istihdamdaki statüye göre yapılan ayrımlar dâhil), işsizlik ve işgücü verimliliği tahminlerini içerir (URL-3).

Oxford COVID-19 Government Response Tracker (OxCGRT-Oxford Covid-19 Hükümet Müdahale İzleyicisi): OxCGRT, hangi hükümetlerin hangi önlemleri ne zaman aldığına dair sistematik bilgiler toplar. Bu, karar vericilerin ve vatandaşların hükümet tepkilerini tutarlı bir şekilde anlamalarına yardımcı olabilir ve pandemi ile mücadele çabalarına yardımcı olabilir. OxCGRT, hükümetlerin aldığı birkaç farklı ortak politika yanıtı hakkında sistematik olarak bilgi toplar, bu politikaları hükümet eyleminin kapsamını yansıtacak bir ölçekte kaydeder ve bu puanları bir dizi politika endeksinde toplar (Hale vd., 2021).

Johns Hopkins Coronavirus Resource Center (CSSE-Johns Hopkins Üniversitesi Sistem Bilimi ve Mühendisliği Merkezi tarafından COVID-19 Veri Deposu) : CSSE, tarafından COVID-19 verilerinin gerçek zamanlı olarak görselleştirmek ve izlemek için çevrimiçi bir etkileşimli pano geliştirilmiştir. Açık olarak paylaşılan gösterge tablosu, etkilenen tüm ülkeler için doğrulanmış COVID-19 vakalarının, ölümlerin ve iyileşmelerin yerini ve sayısını göstermektedir. Toplanan ve görüntülenen tüm veriler, başlangıçta Google E-Tablolar aracılığıyla ve şimdide bir GitHub deposu aracılığıyla ücretsiz olarak sunulur (Dong vd., 2020).

World Health Organization (WHO) Coronavirus (COVID-19) Dashboard Data Explorer (Dünya Sağlık Örgütü (DSÖ) COVID-19 Gösterge Tablosu Veri Gezini): WHO coronavirus disease (COVID-19) explorer, sık sık güncellenen veri görselleştirmeleri sağlamayı amaçlar. Dünya Sağlık Örgütü (DSÖ), 31 Aralık 2019'dan 21 Mart 2020'ye kadar DSÖ, resmi sağlık bakanlıklarının web sitelerini ve sosyal medya hesaplarını izleyerek tamamlanan Uluslararası Sağlık Tüzüğü (UST, 2005) kapsamında resmi iletişimler yoluyla onaylanmış COVID-19 vaka ve ölüm sayılarını toplamıştır. 22 Mart 2020'den bu yana, küresel veriler DSÖ bölgesine özgü gösterge panoları ve DSÖ genel merkezine günlük olarak bildirilen toplu sayım verileri aracılığıyla derlenmektedir (URL-4).

T.C. Sağlık Bakanlığı COVID-19 Bilgilendirme Platformu: Sağlık Bakanlığı Türkiye'deki COVID-19 pandemisinin günlük verilerini turkuaz tablo ile sunmaktadır. İlgili web sayfasında hem ilgili güne ait veriler sunulmakta, hem de "Genel Koronavirüs

Tablosu” sekmesinde geçmişe yönelik rakamlar tablo haline sunulmaktadır. Türkiye geneli pandemi sürecinde ilk vaka 11 Mart’ta, ilk ölüm ise 17 Mart’ta açıklanmıştır (URL-5).

Türkiye COVID-19 Pandemi İzlem Ekranı (TURCOVID19): TURCOVID19, yeni koronavirüs pandemisiyle ilgili Türkiye’deki epidemiyolojik verilerin monitörize edildiği, kamusal kullanıma açık, ücretsiz, üyelik gerektirmeyen, kâr amacı gütmeyen interaktif bir platformdur. TURCOVID19 markası, Türk Patent ve Marka Kurumu (TÜRKPATENT) tarafından 2020-127426 marka numarası ile tescillenmiştir.

Halk sağlığı yaklaşımı 4 aşamadan oluşmaktadır:

1. Veri toplama ve problemi tanımlama
2. Risk faktörlerini, koruyucu faktörleri ve sebepleri tespit etme
3. Müdahale geliştirme, uygulama ve etkinliklerini ölçme
4. Etkili müdahaleleri tüm topluma uygulama ve politika geliştirme

Bu bağlamda TURCOVID19, halk sağlığı yaklaşımının birinci basamağındaki veri toplama kısmına katkı sunmayı hedeflemektedir. Böylece karar vericilerin sorunu moniterize etmesini kolaylaştırmayı, bilim insanlarının çalışmalarına veri desteği sunarak hızlandırmayı, kamunun konuya dair algısını geliştirmeyi amaçlamaktadır (Ucar vd., 2020).

Ek olarak dünya çapında Kaggle, OurWorldinData, WHORegion veri kümeleri CSV formatta yayınlanmaktadır. Günlük olarak güncellenip ve ücretsiz kullanılmaktadırlar. Açık veri kümeleri farklı ülkelere ait COVID-19 vaka bilgilerini kapsamlı olarak ele almamaktadır. Ulusal bazda yayın yapan TURCOVID19 platformu kapsamlı ek bilgi kümeleri sağlamaktadır. T.C. Sağlık Bakanlığı tarafından yayınlanan veriler ve yapılan açıklamalar dikkate alınarak veri kümelerindeki ayrıntı dereceleri dünya çapında yayın yapan diğer veri kümelerinden farklı olduğu tespit edilmektedir.

1.2.2. COVID-19 Verilerinin Semantik Tanımlarının Oluşturulması

COVID-19 verilerinin semantik tanımlarının oluşturulması için ilgili literatürde birçok akademik çalışma yapılmıştır. İlgili literatürde geliştirilen ontolojiler Human Disease Ontology, The Infectious Disease Ontology in the age of COVID-19, The Coronavirus Infectious Disease Ontology, An Ontology for Collection and Analysis of Covid-19 Data, An Ontology For Country Responses Towards COVID-19, An Ontology For Coronavirus Infectious Disease, COVID-19 Surveillance Ontology,

DRUGS4COVID19, COVID-19 Vocabulary, WHO COVID-19 Rapid Version CRF Semantic Data Model, Linked COVID-19 Data Ontology ve KG Hub To Produce A Knowledge Graph For COVID-19 And SARS-COV-2 ontolojileri geliştirilmiştir. Bu ontolojilerin kapsamı, amacı ve uygulama alanları aşağıda detaylı olarak verilmiştir.

Human Disease Ontology (DO-İnsan Hastalık Ontolojisi): İnsan hastalıklarıyla ilişkili biyomedikal verilerin entegrasyonu için açık kaynaklı bir ontoloji sağlamaktır. DO, biçimsel olarak ontolojik anlamda, anlamsal olarak hesaplanabilir bir yapıya sahiptir. Web ontoloji dilinde çok editörlü bir modele taşınan binlerce insan hastalığını sınıflandırır ve hastalık kavramı aracılığıyla farklı veri kümelerinin entegrasyonunu sağlar. Yakın zamanda COVID-19 kavramlarını içeren DO (diseases ontology) tarafından genişletilmiştir (URL-6).

The Infectious Disease Ontology in the age of COVID-19 (IDO-COVID-19 Çağında Bulaşıcı Hastalık Ontolojisi): IDO, yeni hastalıklar hakkındaki verilerin mevcut hastalık ontolojileri tarafından temsil edilen verilerle birden çok boyutta kolayca karşılaştırılmasına izin verecek şekilde yeni patojene özgü ontolojiler oluşturmak için basit bir reçete sağlamaktadır. Ayrıca IDO stratejisi, hekimlerin, araştırmacıların, halk sağlığı kuruluşlarının mevcut ve gelecekteki halk sağlığı krizlerine hızlı ve verimli bir şekilde yanıt vermelerini sağlayan güçlü bir veri entegrasyonu ve paylaşımı yöntemi sağlayarak ontoloji koordinasyonunu desteklemektedir. COVID-19 verilerinin ve daha genel olarak viral bulaşıcı hastalık verilerinin entegrasyonuna ve analizine yardımcı olmak için yakın zamanda üç yeni IDO uzantısı olan IDO Virus (VIDO), the Coronavirus Infectious Disease Ontology (CIDO) ve CIDO'nun COVID-19'a odaklanan bir uzantısı IDO-COVID-19 ontolojileri geliştirilmiştir (Babcock vd., 2021).

The Coronavirus Infectious Disease Ontology (CIDO-Coronavirüs Bulaşıcı Hastalık Ontolojisi): Mevcut COVID-19 pandemisi ve önceki SARS/MERS salgınları, küresel halk sağlığında bir dizi büyük krize neden olmuştur. Etkili, güvenli aşılardan ve ilaçlar geliştirmek adına COVID-19 ve ilişkili hastalık mekanizmalarını daha iyi anlamak için büyük ve katlanarak artan miktarda heterojen koronavirüs verilerini entegre etmektedir. Ayrıca COVID-19 ve diğer hastalıkların semptomları, ilaçlar, klinik araştırmalar vb. karşılaştırmalı analizini temsil etmektedir. Ontolojiler, standart bilgi ve veri temsili, entegrasyon, paylaşım ve analizde önemli bir rol oynamak için ortaya çıkmıştır. Bir Açık Biyomedikal Ontoloji (OBO) kitaplığı ontolojisi olarak CIDO, açık bir kaynaktır ve mevcut diğer OBO ontolojileriyle birlikte çalışabilir (He vd., 2021).

The COviD-19 Ontology for cases and patient information (CODO- Vakalar ve hasta bilgileri için COVID-19 Ontolojisi): CODO, COVID-19 pandemisi ile ilgili verilerin toplanması, analizi ve COVID-19 verilerini Web'de Grafik olarak yayınlamak için bir model sağlar. Ontoloji, heterojen veri kaynaklarından gelen verilerin entegrasyonunu kolaylaştıran standartlara dayalı bir açık kaynak modeldir. Veri kümeleri, literatür, hizmetler vb. gibi farklı COVID-19 veri kaynaklarını analiz ederek tasarlanmıştır. Önde gelen diğer sözlüklerdeki kavramları yeniden kullanarak ve W3C standartları RDF, OWL, SPARQL ve SWRL'yi kullanarak sözlükler için en iyi uygulamaları takip eder. Modelin ana odak noktaları şunlardır;

- COVID-19 vakaları: Coğrafi konumdaki (ilçe, eyalet ve ülke) günlük bazda COVID-19 vakaları için veriler (ör. Aktif, iyileşmiş, ölmüş, göç etmiş), mevcut kaynaklar ve gereksinimler,
- COVID-19 hasta verileri: Uyruk, semptom, şüpheli COVID-19 vaka düzeyi, COVID-19 tedavi tesisi, COVID-19 klinik tesisi, hastanın seyahat geçmişi, hastalar arasındaki kişisel ilişkiler, şüpheli bulaşma nedeni, hasta test sonuçlarının izlenmesi vb. ayrıntıların izlenmesini sağlamaktır (Dutta ve DeBellis, 2020).

An Ontology For Country Responses Towards COVID-19 (ROC- Ülkelerin COVID-19'a Karşı Müdahaleleri İçin Bir Ontoloji): Ülkelerin COVID-19'a verdiği yanıtlara yönelik ROC ontolojisi, COVID-19 pandemisine ilişkin verilerin toplanması, bağlanması ve paylaşılması için bir model sağlar. Kavramların temsili ve sözlüklerin oluşturulması için semantik standardizasyonu (W3C standartları RDF, OWL, SPARQL) takip eder. ROC, ülke önlemlerine odaklanır ve heterojen veri kaynaklarından gelen verilerin entegrasyonunu sağlar. Önerilen ontoloji, ülkelerde hükümetin COVID-19'a karşı alınan önlemler karşısında verilen yanıtların etkinliğini ve yan etkilerini incelemek ve değerlendirmek için istatistiksel analizi kolaylaştırmayı amaçlamaktadır. Ontoloji, OxCGRT tarafından halka açık bilgilerden toplanan verileri içerir. Bu veriler, ECDC tarafından çoğu ülke için sağlanan bilgilerin yanı sıra COVID-19 hakkında veri toplamak için kullanılan çeşitli veri havuzlarından derlenmiştir (Qundus vd., 2021).

An Ontology For Coronavirus Infectious Disease (WHOnto- Coronavirüs Enfeksiyon Hastalığı İçin Bir Ontoloji): Dünya Sağlık Örgütü (World Health Organization - WHO)'ya göre Coronavirüs ile enfekte vakalar için koronavirüs bulaşıcı hastalık alanında açık kaynaklı ontolojidir. WHOnto, aktif vakalar, ülke genelinde toplam ölüm sayısı

hakkında bilgi sağlar. WHOnto, etiyojisi, bulaşması, patogenezi, teşhisi, önlenmesi ve tedavisi dâhil olmak üzere çeşitli koronavirüs bulaşıcı hastalıklarının standartlaştırılmış insan ve bilgisayar tarafından yorumlanabilir açıklamalarını ve temsilini sağlamak için geliştirilmiştir (URL-7).

COVID-19 Surveillance Ontology (COVID-19 Gözetim Ontolojisi) : COVID-19 Surveillance Ontology, birinci basamakta sürveyansı desteklemek için tasarlanmış bir uygulama ontolojisidir. Ontoloji, birden fazla farklı kayıt sistemlerinden gelen verileri kullanarak COVID-19 vakalarının ve ilgili solunum koşullarının izlenmesini kolaylaştırır (URL-8).

DRUGS4COVID19: DRUGS4COVID19, ilaçların COVID-19 ile ilgili ilişkilerini tanımlar. İlaç, etki, hastalık, semptomlar, bozukluk, kimyasal madde vb. ontolojinin bazı temel sınıflarıdır (URL-9).

COVID-19 Vocabulary (COVID-19 COVID-19 Kelime Bilgisi): COVID-19 veri kümelerinin RDF'de tanımlanmasını sağlayan sınıflardan oluşan bir ontolojidir. Bu ontolojinin sınıflarından bazıları Veri Kümesi, Johns Hopkins Üniversitesi Veri Kümesi vb.'dir (URL-10).

WHO COVID-19 Rapid Version CRF Semantic Data Model (COVIDCRFRAPID-DSÖ'nün COVID-19 Vaka Kayıt Formunun RAPID Versiyonu için Semantik Bir Veri Modeli): Ontoloji, Dünya Sağlık Örgütü'nün 23 Mart 2020 tarihli COVID-19 RAPID vaka kayıt formu için semantik bir veri modelidir. Bu model, formun soru ve cevaplarına semantik referanslar sağlar (URL-11).

KG Hub To Produce A Knowledge Graph For COVID-19 And SARS-COV-2 (Kg-COVID-19- KG Merkezi, COVID-19 ve SARS-COV-2 İçin Bir Bilgi Grafiği): Eyleme geçirilebilir bilgi üretmek amacıyla COVID-19 ve SARS-COV-2 ile ilgili veri kümelerini indirmek ve dönüştürmek için KG Hub'ın tasarım modellerini izleyen ve daha sonra makine öğrenimi vb. kullanımlar için bir bilgi grafiği yayan çerçevedir (URL-12).

Linked COVID-19 Data Ontology: Johns Hopkins Üniversitesi tarafından oluşturulan veri deposu temel alınmıştır. Bu veri deposundan gelen veriler Avrupa Hastalık Önleme ve Kontrol Merkezi tarafından json dosyası oluşturularak açık veri olarak yayınlanır (URL-13).

Tez çalışması kapsamında The COviD-19 Ontology for cases and patient information (CODO) kullanılmıştır. CODO geliştirilmeden önce COVID-19 pandemisi ile ilgili olarak

geliştirilmiş çeşitli ontolojiler olmuştur. Mevcut ontolojiyi yeniden kullanılıp kullanılmayacağına dair anketler yapılmıştır. İlgili dokuz ontoloji bulunmuştur.

Hindistan ve dünyadan vakalara ait veriler üzerine anlamsal bir katman sağlamak amacı ile ilgili ontolojilerin hiçbiri ihtiyaç olanla aynı alanda olmadığı tespit edilmiştir (Qundus vd., 2021).

1.3. Tezin Amacı ve Kapsamı

COVID-19, bulaşıcı bir hastalık olarak 2019 yılının sonu ve 2020 yılının başından itibaren hızlı bir şekilde ülkelerin gündemine yerleşmiştir. Birçok kuruluş, salgının izlenmesine olanak tanıyan virüsün etkisine ilişkin küresel verileri toplamak ve yayınlamak için sistemler ve projeler kurmuştur. Gün geçtikçe girdi miktarı artmakta ve aşırı bilgi oluşmaktadır. Bilginin bilgisayarlar tarafından da anlaşılabilir hale getirilmesi ve işlenebilmesi için Web kaynaklarında bulunan COVID-19 veri setlerinin anlamsallığının makinelere anlaşılabilir şekilde ve somut olarak belirlenmesini sağlayan ontolojiler geliştirilmiştir. Bu ontolojiler, COVID-19 vakaları ile ilgili kavramsallaştırmaları ve örnek verileri içermektedir. Öncelikleri soru-cevap veya gösterge tablolarını izleme gibi yazılım uygulamalarına yöneliktir. Ulusal bazlı olan, TURCOVID19 platformundan elde edilen veri kümeleri mevcut kaynaklardan ayrıntı dereceleri bakımından farklılık göstermektedir. Bu noktada “COVID-19 Verilerinin Bağlantılı Veri Olarak Yayınlaması ve Sorgulanması” konulu tezin amacı; TURCOVID19 veri kümelerindeki tüm veri öğelerini anlamsal olarak modellenmesinin gerçekleştirilmesi için TURCOVID19 ontolojisi geliştirmek ve verilerin semantik tanımlarını oluşturmaktır. Açık veri olarak yayınlanan veri kümeleri tablosal biçimlerde temsil edilmektedir. Veriler geliştirilen ontoloji ile ilişkilendirilip yapılandırılmış RDF formatına dönüştürülerek farklılıkları irdelenecektir. Ayrıca ulusal bazlı ve güncel COVID-19 bilgilerinin sürdürülebilirliğini ve analiz edilen pandeminin çeşitli yönlerini kapsamaları için farklı veri kaynakları ile bağlantılar kurarak yayılması amaçlanmıştır.

Bağlantılı Veriler ile ilgili tezde genel bilgiler ve araçlar açıklanmıştır. Uygulama, TURCOVID19 platformundan temin edilen açık veriler ile yapılmıştır.

1.4. Semantik Web

“Kısaca bilgiye erişilebilir küresel ağ evrenidir” olarak World Wide Web (WWW veya Web) kavramı ortaya çıkmıştır. (Berners-Lee, 1996). Tim Berners-Lee (1996) tarafından “insanların ve bilgisayarların iletişime geçebileceği, bilgi paylaşabilecekleri bir ortam oluşturmak” amacıyla web kavramı ortaya çıkmıştır. Öneri raporu olarak Avrupa Nükleer Araştırma Merkezine (European Organization for Nuclear Research - CERN) 1979 yılında Berners-Lee webi sunmuştur. Web 1990 yılında geliştirilmiştir. Webde temel kavram olarak kabul edilen hiper metin (hyper text), belgelerin bağlantılar aracılığıyla birbirine bağlanmasını sağlamaktadır ve kullanıcıların belgeler arasındaki geçişini kolaylaştırmaktadır (Berners-Lee, 1979). Farklı sistemlerin, web tarayıcısı, HTML sunucusu ve web sunucusu gibi evrensel veri paylaşımının sağlanması ile ilk kez bir arada kullanılmıştır. İnternet her geçen gün artan farklı formatta ve büyük miktarda bilginin bulunduğu bir ortam haline gelmiştir. Web sayfalarının sadece insanların anladığı formata sahip olması ve bilgisayarların bu yapıya sahip olmaması nedeni ile kullanıcılar web sayfalarındaki aramalarda istediği bilgilere hızlı bir şekilde ulaşamamakta ve zaman açısından kayıp oluşmaktadır. Sadece insanların anlayabileceği yapıya sahip olan web sayfaları bilgisayarlar tarafından etkin olarak kullanılamamaktadır. Çoğu zaman kullanıcılar aradığı doğru ve güvenilir bilgilere ulaşamamakta; arama girişimleri emek ve zaman kaybı ile sonuçlanmaktadır (Karalar ve Özdemir, 2013). Anahtar kelime ile verinin düzenlenmesi ve kullanılmasında çözüm için yaklaşımlar geliştirilmiş ama yeterli olmamıştır (Hitzler vd., 2009). Mevcut sorunlara yaklaşım olarak geleneksel webde, yalnızca insanlar için anlamlı olan ve bulunan içeriklerin makineler tarafından anlamlı hale getirilmesi ve sorunlara çözüm olarak geliştirilen bu yaklaşıma Semantik web adı verilmiştir. Semantik Web ile geleneksel web arasındaki en önemli fark, arama motorları sistemlerinde fark edilmektedir. Bir anlamsal arama motoru kullanan kullanıcı, kendisini ilgilendiren konuda bazı anahtar kelimeler girmek yerine sorusunu sisteme doğrudan yöneltebilecektir (Ege, 2011). Tim Berners Lee tarafından Semantik Web kavramı 2001 yılında günümüz webinin bir uzantısı olarak tanıtılmıştır. Semantik Web ile bilgisayarlar ve insanların webi anlamasına vurgu yapılmıştır (Karalar ve Özdemir, 2013).

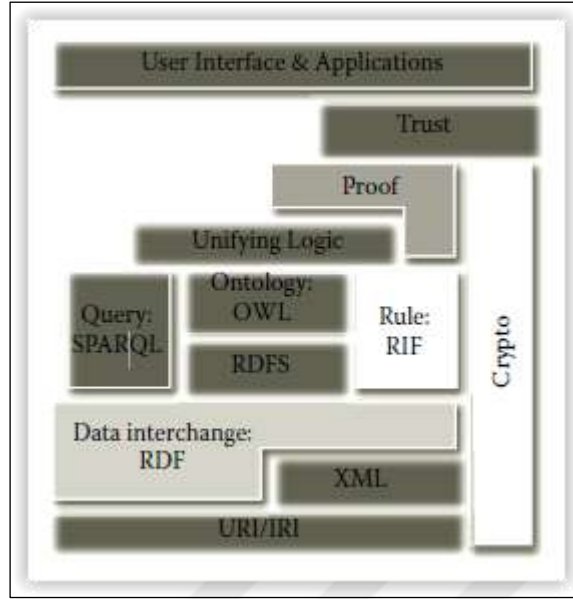
Semantik Web, web içerikleri ilgili yazılımlar tarafından anlaşılabilir, yorumlanabilir ve kullanılabilir bir biçimde ifade edilmektedir. Semantik web, ilgili yazılımların veriyi kolayca bulmasını, paylaşmasını ve bilgiyi birleştirmesini sağlama amacıyla geliştirilmiştir

(URL-14). WWW, yalnızca insanlar tarafından okunabilen belgelerden oluşmaktadır. Semantik Web'in önemi XML (eXtensible Markup Language) ve XML tabanlı geliştirilmiş olan RDF (Resource Description Framework) teknolojileri ile veri ve veri hakkında çıkarılmaların sağlanmasıyla ortaya çıkmaktadır. Ayrıca, webdeki tüm platformların izinler çerçevesinde veriye ulaşımı sağlanmaktadır. RDF üçlüsü olarak kodlanmış ifadelerde özne ve yüklem URI'ler ile temsil edilen kaynaklardır. Kaynaklara ya da bilgiye ait elemanlar nesne ile ifade edilir. XML etiketleri kullanılarak üçlüler oluşturulur (W3C-RDF, 2009). Ancak, bilgi miktarının artması ile RDF dili yetersiz kalmıştır ve RDF gibi XML tabanlı OWL dili geliştirilmiştir. OWL dili ontoloji geliştirmek için W3C tarafından bir standart olarak kabul edilmiştir.

Kavramlar arasındaki ilişkileri temsilen ontolojiler, semantik web'in temel bileşenlerinden biridir. Ayrıca, RDF verilerinin SQL benzeri sorgulanmasına izin veren SPARQL sorgulama dili vardır (Demirli ve Kütük, 2010). Geleneksel web'de bilgilerin büyük bir kısmının metinsel kaynaklı olması, bu bilgilerin anlamlandırılıp, bilgisayarlar tarafından anlaşılmasını ve aralarında ilişki kurmasını engellemektedir. Aralarında ilişki kurulamayan bilgilerden, otomatik yeni bilgi çıkarsaması imkânsız hale gelmektedir (Ege, 2011). Semantik web ile webde yer alan tüm bilgi kaynakları arasında bağlantıların ve böylece bilgisayarların hızlı ve kolay bir şekilde çıkarsama yapabilmesi amaçlanmıştır ve geliştirileye devam etmektedir.

1.5. Semantik Web Bileşenleri

World Wide Web Consortium (W3C) tarafından Semantik Web'e ait teknoloji ve standartlar geliştirilmiştir ve bu standartlar Şekil 1'de gösterilmiştir.

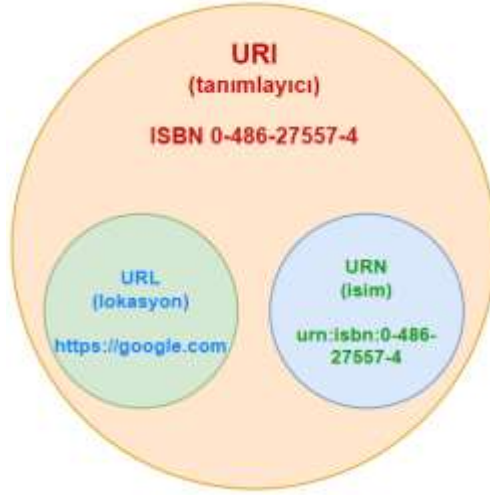


Şekil 1. Semantik web yapısı (Hart ve Dolbear, 2013).

URI/IRI, XML, RDF, RDFS, OWL, RIF, SPARQL, Logic, Proof, Trust gibi bileşenler Semantik Web'in yapısında yer almaktadır.

1.5.1. URI/IRI

URL (Uniform Resource Locator) internette bulunan bir kaynağı temsil etmektedir. “Tekdüze kaynak tanımlayıcısı” olan Uniform Resource Identifier (URI), işaret ettiği URL'nin kaynak konumundan sonra gelen ayırıcı adresini belirtir. URL kapsamında tutulan alt tanımlardır. Belirli bir formata sahip olan URI, Tim Berners-Lee tarafından oluşturulmuştur. URI, URL (konumlayıcılar) ve URN (adlar) olarak ikiye ayrılmaktadır. URI aynı anda hem bir isim hem de konumlandırıcı olabilmektedir. URI hem konum hem isim olarak kullanılmaktadır. Böylece, URL ve URN kavramları ile çakışma sorunu ortaya çıkmaktadır. URI örneği olarak www.example.com/book verilebilir. Burada konuma erişilebilir veya erişilemeyebilir. URL olarak <http://www.example.com/book>, isim ve konumdur. Bu URL örneği erişilebilirdir. URN olarak Book Seller, sadece isimdir. Bir URN genelde “urn:” prefix'i ile başlar. URI'nin uluslararası bir varyantı, tanımlayıcıda Unicode karakterlerin kullanımına izin veren ve bunun için URI'ye eşlemenin tanımlandığı Internationalized Resource Identifier (IRI) dir (URL-15).

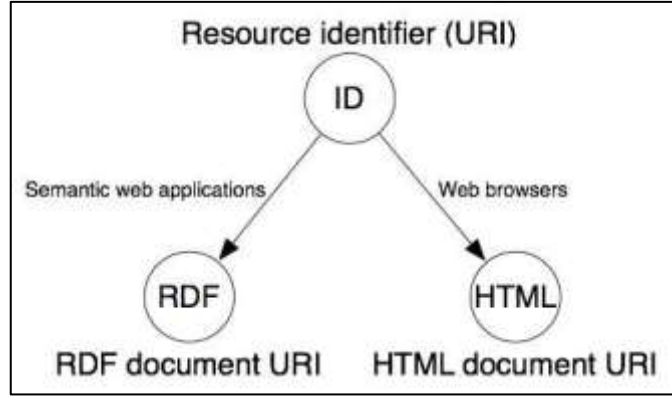


Şekil 2. URI, URL ve URN ilişkisi (URL-16).

RDF ile webde bulunan belge ve kaynakları bilgisayarlar tarafından işlenebilir olarak tanımlanmaktadır. Bu tanımlamaların web ortamında yayınlanması Semantik Web oluşturulmuştur. URI'ler ile RDF ve Semantik Web arasında bağlantı sağlanmaktadır (Sauermann vd., 2008). Geleneksel webdeki dokümanlarda URL kullanılmıştır. Her URL bir URI olabilir fakat bir URI yalnızca kaynakları adlandıran bir isim olarak kullanılmaktadır (Berners-Lee vd., 1998).

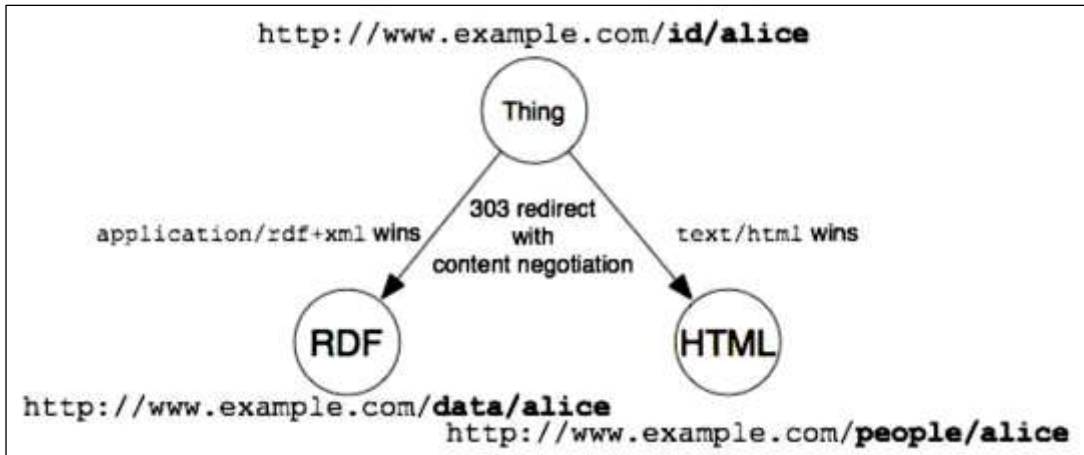
Web dökümanlarının görüntülenmesi, farklı format ve dillerde sunulması HTTP protokolleri ile sağlanmaktadır (Fielding vd., 1999). Semantik Web'de gerçek nesnelere veya soyut şeyler web belgeleri olarak tanımlanmaktadır. HTTP'nin Semantik Web'deki önemi URI'lerin yeniden referanslandırmasını sağlamaktır (Kara, 2014). HTTP kullanılarak URI tanımlanmalıdır. İnsanlar tarafından HTML, RDF verisibilgisayarlar tarafından elde edilmelidir.

Tim Berners-Lee tarafından semantik web ile anlamsal genişleme olan URI'ler için İyi URI kavramı ortaya atılmıştır. İyi URI, web dökümanı veya kaynağın URI ile temsil edilmesi ve hakkındaki bilgilere web yoluyla ulaşılmasını açıklamaktadır. Bir URI sadece gerçek dünya nesnesi ya da Web dokümanı olarak kullanılır. İkisini aynı anda temsil edemez (Ayers ve Völkel, 2008).



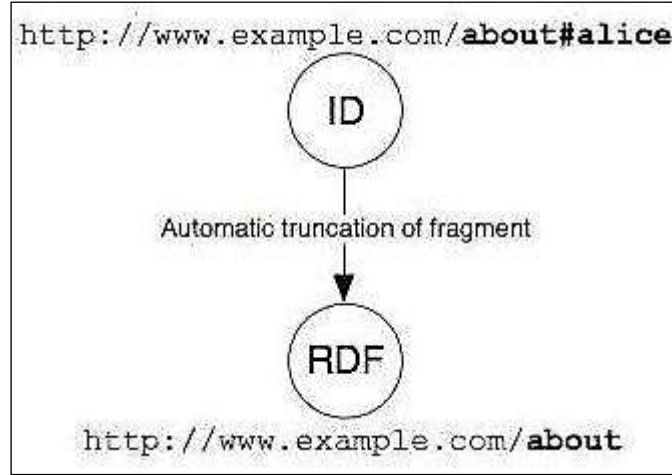
Şekil 3. Bir kaynağa ait URI ve tanımlayıcı dokümanı (Ayers ve Hölkel, 2008).

Ancak, URI web dökümanı ve konumunu aynı anda bir dünya nesnesini belirtmiyorsa kaynağı ifade eden URI'ye HTTP kullanarak erişebilirlik problemi ortaya çıkmaktadır. Kaynak ya da belge olarak URI'nin yalnızca birini tanımlayabilme sorunu, iki farklı yaklaşım ile çözülmektedir. 303 URI yaklaşımı ile nesneyi doğrudan ağ üzerinden göndermek yerine, istemciye bir HTTP yanıt kodu (response code) olarak '303 see other' kodu ile nesneyi tanımlayan Web dokümanının URI'si gönderilir.



Şekil 4. 303 URI çözümü (Ayers ve Hölkel, 2008).

Diğer bir yaklaşım URI'nin hash sembolü (#) ayrıldığı Hash URI'lerdir. URI'nin tamamından yanıt almak yerine ayrılan bu kısım ile sunucuya istek yapar (Health ve Bizer, 2011). URI'den ayrılan bu kısma bölüm kimliği (fragment identifier) denilmektedir (Berners-Lee vd., 2005).



Şekil 5. İçerik uzlaştırma ve hash URI çözümü (Ayers ve Hölkel, 2008)

Uygulamalarda kullanılacak URI yaklaşımı bazı değişkenlere bağlıdır. Daha çok küçük projeler ve veri setleri için Hash URI'ler kullanılmaktadır. Büyük veya büyüyecek veri setleri için 303 URI'ler kullanılmaktadır. En geniş kapsamlı Bağlı Veri uygulaması olan DBpedia, 303 URI yaklaşımını kullanmıştır. Hash URI'ler RDF belgeleri için Dünya Jeodezik Sistemi (World Geodetic System - WGS84), SKOS gibi sözlüklerdeki terimleri oluşturan çalışmalarda, binlerce terimi tek bir seferde çağırmak için kullanılmaktadır (Health ve Bizer, 2011).

1.5.2. XML

XML (Extensible Markup Language), genişletilebilir işaretleme dili anlamına gelmektedir. İnsanların ve makinelerin kolayca okunabilecek dökümanlar oluşturmayı sağlamaktadır. W3C tarafından tanımlanmış, veri saklama ve farklı sistemler ve uygulamalar arasında veri alışverişinin tek bir format yapılmasını sağlayan bir standarttır (URL-17). XML bir programlama dili değildir. Verileri görüntülemek değil, depolamak ve aktarmak amacıyla XML verileri tasarlanmıştır. SGML'in basitleştirilmiş bir alt kümesidir. SGML, standart bir meta dil olarak 1986 yılında ISO tarafından geliştirilmiştir. Kullanımı, karmaşık bir dil olması nedeni azalmıştır. Bu nedenle, yerine 1989 yılında HTML geliştirilmiştir fakat genişletilebilir bir dil değildir. İhtiyaçları karşılayamamıştır ve böylece XML yeni bir dil olarak tasarlanmıştır. HTML ile kullanılmasını sağlayan SGML'nin daha basit bir biçimi oluşturulmuştur (URL-18). XML, veriyi tanımlamaktadır. Verinin

görselleştirilmesi HTML ile sağlanmaktadır. HTML gibi etiketler (tags) XML’de bulunmaktadır. Önceden belirtilmiş etiketler yerine kullanıcıların kendilerine ait etiketler oluşturmasına imkân vererek HTML’den farklıdır. XML’in genişletilebilir bir dil olması kullanıcıların kendi etiketlerini tanımlayabilmesi ile açıklanmaktadır (Bikakis vd., 2013).

XML’in ana katkısı, Web dokümanları için ortak ve iletilebilir bir sözdizimi sağlamasıdır. XML sözdizimi yazılması ve okunması basittir. XML, webde bulunan verilerin anlamlarını tanımlamaktadır ve yeni nesil webin temelini oluşturmaktadır. Bilgisayarlar tarafından web verilerine anlamlar eklenmesini sağlayan XML, anlamsal ağın temelidir (Geroimenko ve Chen, 2006).

```

1 <?xml version="1.0" encoding="iso-8859-9"?>
2 <kitaplar>
3   <kitap1>
4     <ad>Kar</ad>
5     <yazar>Orhan Pamuk</yazar>
6   </kitap1>
7   <kitap2>
8     <ad>İnce Memed</ad>
9     <yazar>Yaşar Kemal</yazar>
10  </kitap2>
11  <kitap3>
12    <ad>Çalıkuşu</ad>
13    <yazar>Reşat Nuri Güntekin</yazar>
14  </kitap3>
15 </kitaplar>

```

Şekil 6. Basit XML kod örneği (URL-19)

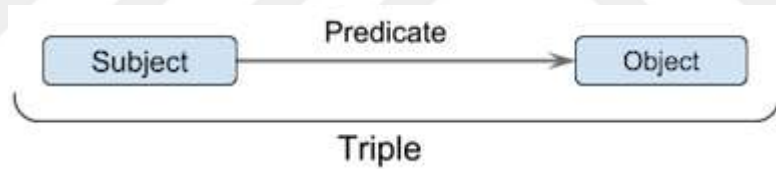
1.5.3. RDF

1999 yılında oluşturulan RDF (Resource Description Framework), meta verilerin işlenmesi amacıyla W3C tarafından önerilmiş bir standarttır. 2004 yılında RDF sadece web kaynaklarını değil, Semantik Web kavramının geliştirilmesiyle gerçek dünya nesnelere ve ilişkilerini tanımlamak için W3C tarafından güncellenmiştir (Yu, 2011).

RDF, RDF'nin birincil amacı, verilerin orijinal anlamını korurken web üzerinde veri alışverişi yapmaktır. Kaynakları tanımlamak için kullanılan bir veri modelidir. RDF, bilgilerin işlenmesine izin verir. RDF, kaynaklar hakkındaki bilgileri graph (çizge) şeklinde temsil eden bir çerçevedir (framework). XML sözdiziminin tavsiye edilmesinden sonra RDF ilk olarak yayınlanmıştır (Gandon vd., 2011). XML ile makineler tarafından anlaşılabilir bilgiler sağlanmaktadır. RDF ise birlikte çalışabilirlik sağlamak amacıyla, belge ve belgede yer alan varlıklar arası ilişkiyi açıklamaktadır (Thuraisingham, 2002).

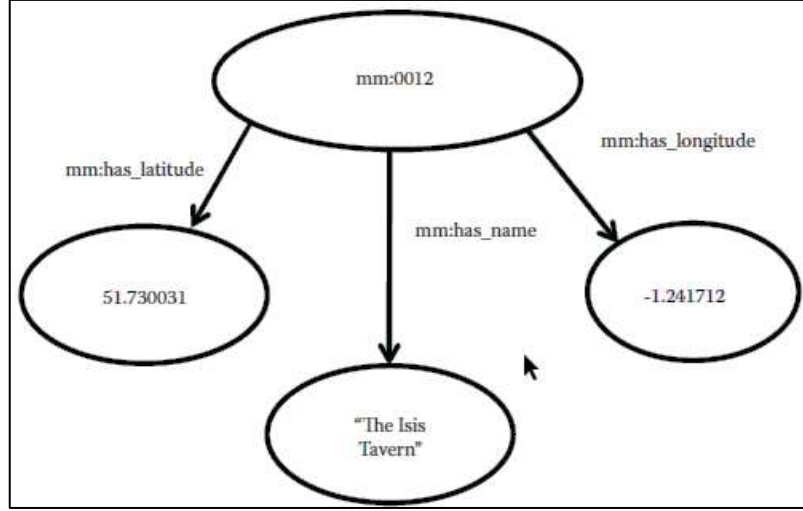
RDF'nin XML tabanlı olduğu, meta verilerin bilgisayar tarafından işlenebilmesi için XML özelliklerinden yararlanması ile açıklanmaktadır. HTML ve XML'in tersine amacı, belgeleri doğru şekilde görüntülemek değil, belgede bulunan bilgilerin daha fazla işlenmesine ve yeniden birleştirilmesine izin vermesidir. Bu özelliği ile RDF'nin Semantik Web'in temeli olduğu açıklanmaktadır (Miller, 1998).

RDF'ye ait olan veri modeli, kaynak (resource), özellik (property) ve ifadeden (statement) meydana gelen grafik veri modelidir ve bir dil değil verileri temsil etmektedir. RDF, kaynaklar hakkında açıklama yapmanın standart bir yoludur. Kaynaklar webdeki şeyleri temsil etmektedir. Kaynağa ait öznitelik ya da ilişki özellik ile ifade edilmektedir. Özne (subject), yüklem (predicate) ve nesne (object) yapısı ile ifade oluşmaktadır. Bu üçlü gruplara RDF çizgesi (graph) adı verilmektedir. Özne (subject), üçlü tarafından açıklanan bir kaynaktır. Yüklem (predicate), özne ile nesne arasındaki ilişkiyi açıklar. Nesne (object), konuyla ilgili bir kaynaktır. Özne ve nesne, şeyleri temsil eden düğümlerdir. Yüklem bir yaydır, çünkü düğümler arasındaki ilişkiyi temsil eder (URL-21). RDF üçlüsü olarak özne-yüklem-nesne yapısı Şekil 7'de görülmektedir.



Şekil 7. RDF üçlüsü (URL-20)

RDF kaynakları özne, yüklem, nesne şeklinde tanımlamaktadır ve RDF üçlüsünün her bileşeni bir URI veya IRI olarak ifade edilebilir. URI, bir web kaynağına işaret eden bir URL olabilir veya rastgele veriler içerebilir. Aynı varlık hakkında birden çok RDF ifadesinin tümü, RDF üçlüleri olacaktır. Aynı özneye, ancak farklı yüklemelere ve nesnelere sahiptirler. Bu üçlülerden bir RDF grafiği oluştururken, özne, farklı yüklemeleri ve farklı nesnelere temsil eden öznenin ayrılan çok sayıda okla bir kez görüntülenebilir (URL-21). RDF grafik veri modeli RDF verilerinin görsel yardımıyla daha kolay anlaşılmasını sağlamaktadır.



Şekil 8. RDF çizgesi örneği (Hart ve Dolbear, 2013)

RDF sözdizimi ile serileştirildikten sonra yayınlanan RDF grafikleri, makine tarafından okunabilir bir biçimde yazılır. Bu işleme, RDF serileştirmesi olarak tanımlanmaktadır. RDF genellikle dört formattan birinde depolanır: N- Triples (.nt), Turtle (.ttl), JSON-LD (.json) veya RDF/XML (.rdf). Tüm bu formatlar ana RDF kitaplıkları ve üçlü depolar tarafından desteklendiğinden, kullanılan esas tercihe bağlıdır (URL-22). İlk RDF serileştirmesi olarak tanımlanan RDF/XML, W3C tarafından kullanılan formattır. RDF belgeleri XML ile yazılır. Sözdiziminin detaylı yazılması nedeniyle okunması zordur. Turtle gibi daha basit olan formatlar geliştirilmesinin nedeni, XML sözdiziminden kaynaklanan zorluklardır (Hyvönen, 2012). RDF/XML'e göre genellikle daha okunabilir ve XML karşılığına göre manuel olarak düzenlenmesi daha kolay olarak kabul edilir. RDF için sorgu dili olan SPARQL, sorgu kalıplarını ifade etmek için Turtle'a benzer bir sözdizimi kullanır (URL-23). N-Triples, verileri depolamak ve iletmek için bir formattır. RDF grafikleri için satır tabanlı, düz metin serileştirme formatı ve Turtle formatının bir alt kümesidir ve Turtle formatından daha basit bir formattır (URL-24).

```

1 <?xml version = "1.0" encoding = "UTF-8"?>
2 <!DOCTYPE
3 <rdf:RDF [!ENTITY xsd http://www.w3.org/2001/XMLSchema#]>
4 <rdf:RDF
5     xmlns:rdf = "http://www.w3.org/1999/02/22-rdf-syntax-ns#"
6     xmlns:meraeMaps = "http://meraeamaps.gov.me/topo/">
7 <rdf:Description
8     rdf:about = "http://meraeamaps.gov.me/topo/0012">
9 <meraeMaps:has_name
10     rdf:datatype = "&xsd:string"/>The Isis Tavern
11 </meraeMaps:has_name>
12 <meraeMaps:has_latitude rdf:parseType = "Resource">
13 <rdf:value
14     rdf:datatype = "&xsd:float"> 51.730031</rdf:value>
15 <http://meraeamaps.gov.me/has_units
16     rdf:resource = "http://meraeamaps.gov.me/units/degrees"/>
17 </meraeMaps:has_latitude>
18 </rdf:Description>
19 </rdf:RDF>

```

Şekil 9. RDF/XML serilizasyonunda RDF üçlüsü (Hart ve Dolbear, 2013)

```

1 <http://meraeamaps.gov.me/topo/0012> <http://www.w3.org/1999/
02/22-rdf-syntax-ns#type> <http://meraeamaps.gov.me/topo/Pub>.
2 <http://meraeamaps.gov.me/topo/0012> <http://meraeamaps.gov.me/ topo/
has_name> "The Isis Tavern".
3 <http://meraeamaps.gov.me/topo/0012> <http://meraeamaps.gov.me/ topo/
has_longitude> "1.241712" ^^<http://www.w3.org/2001/XMLSchema#float>.
4 <http://meraeamaps.gov.me/topo/0012> <http://meraeamaps.gov.me/ topo/
sells> _:beer0.
5 _:beer0 <http://meraeamaps.gov.me/topo/isbrewed_in>
<http://sws.geonames.org/2640726>.

```

Şekil 10. N-Triples serilizasyonunda RDF üçlüsü (Hart ve Dolbear, 2013)

```

1 @prefix rdf: <http://www.w3.org/1999/02/22-rdf-syntax-ns#>.
2 @prefix meraeMaps: <http://meraeamaps.gov.me/topo/>.
3
4 meraeMaps:0012 meraeMaps:has_name "The Isis Tavern".
5 meraeMaps:0012 meraeMaps:has_longitude "-1.241712".

```

Şekil 11. Turtle serilizasyonunda RDF üçlüsü (Hart ve Dolbear, 2013).

JSON-LD, basit bir sözdizimine sahip olan W3C standardıdır. Bağlantılı verilerin serileştirilmesinde kullanılır ve JSON dilinin özelliklerinden faydalanmaktadır. Bağlantılı verilerin serileştirilmesi için JSON (JavaScript Object Notation), veri değişim formatıdır.

İnsanların okuması ve yazması kolaydır. Makinelerin ayrıştırması ve üretmesi kolaydır. Ayrıca JSON verisi ile RDF'ye dönüşüm kolaylıkla yapılmaktadır (URL-25).

```
{ "@context": [
  "https://chalk.coas.unf.edu/champ/files/contexts/champ_substance.jsonld",
  {
    "cao": "http://champ-project.org/images/ontology/champ.owl#",
    "obo": "http://purl.obolibrary.org/obo/",
    "ci": "http://semanticscience.org/resource/"
  },
  {
    "@vocab": "http://www.w3.org/2001/XMLSchema#",
    "analyte": { "@id": "cao:CAO_000004", "@type": "@id" },
    "substanceClass": { "@id": "cao:CAO_000014", "@type": "string" },
    "functionalGroup": { "@id": "obo:CHEBI_33232", "@type": "string" },
    "biologicalProperty": { "@id": "cao:CAO_000019", "@type": "string" },
    "chemicalProperty": { "@id": "cao:CAO_000015", "@type": "string" },
    "physicalProperty": { "@id": "cao:CAO_000016", "@type": "string" },
    "analyzedForm": { "@id": "cao:CAO_000017", "@type": "string" }
  }
] }
```

Şekil 12. JSON-LD serilizasyonunda RDF üçlüsü (URL-26).

Webde bulunan mevcut veri kaynakları ve bu veri kaynaklarının diğer kaynaklar ile arasındaki ilişkiler RDF linkleri ile sağlanmaktadır. RDF linki, iç (internal) ve dış (external) olarak iki türe ayrılmaktadır. İç RDF linkleri, tek bir bağlantılı veri setine ait kaynakları birleştirmekte ve dış RDF linkleri ise bağlantılı veri kaynaklarına ait veri setleri arasında ilişki kurarak yeni veri kaynaklarının keşfedilmesini sağlamaktadır (Kara vd., 2019).

RDFa, RDF verilerini gömmek için HTML, XHTML ve çeşitli XML tabanlı belge türlerine bir dizi öznitelik düzeyinde uzantı ekleyen bir W3C Önerisidir. W3C tarafından XHTML modülü W3C tarafından 2004 yılında oluşturulmuştur. RDFa 1.1 modülü 2012 yılında HTML ve HTML5 için uyumludur ve aynı zamanda RDFa Lite versiyonu oluşturulmuştur. RDF ifadelerine RDFa ile açıklama eklemektedir. Web verilerinin HTML'e gömülmesi ile verilerin bilgisayarlar tarafından okunabilir verilere dönüşümü sağlanmaktadır (URL-27). Böylece, yapılandırılmış meta veri açıklamaları ve HTML için yazma miktarı RDFa ile azaltılmaktadır (Hyvönen, 2012).

```

1 <!DOCTYPE html PUBLIC "-//W3C//DTD XHTML+RDFa 1.0//EN"
2 "http://www.w3.org/MarkUp/DTD/xhtml-rdfa-1.dtd">
3 <html xmlns = "http://www.w3.org/1999/xhtml"
4   xmlns:rdf = "http://www.w3.org/1999/02/22-rdf-syntax-ns#"
5   xmlns:mereaMaps = "http://Mereamaps.gov.me/topo/"
6   xmlns:dc = "http://purl.org/dc/elements/1.1/">
7
8 <head>
9 <meta http-equiv = "Content-Type" content = "application/xhtml+xml;
10 charset = UTF-8"/>
11 <meta property = "dc:creator" content = "CodeMonkey1"/>
12 <title>The Isis Tavern Homepage </title>
13 </head>
14
15 <body>
16 <div about = "mereaMaps:0012" typeof = "mereaMaps:Pub">
17 <span property = "mereaMaps:has_name" content = "The Isis
18 Tavern"/>
19 <span property = "mereaMaps:has_longitude" content =
20 "-1.241712"/>
21 <span property = "mereaMaps:has_latitude" content = "51.730031"/>
22 <h1> Welcome to the Isis Tavern!</h1>
23 </div>
24 </body>
25 </html>

```

Şekil 13. RDFa içeren bir HTML belgesi örneği (Hart ve Dolbear, 2013)

RDFa Lite, RDFa Lite, HTML, SVG ve XML gibi Web belgelerinde makine tarafından okunabilen verileri ifade etmek için kullanılacak birkaç öznelikten oluşan RDFa'nın daha basit bir alt kümesidir. Yalnızca vocab, typeOf, property (özellik), resource (kaynak) ve prefix (önek) gibi beş özellik bulundurmaktadır. RDFa'da öge etiketlerinin temeli olan kelimeler vocab ile tanımlanmaktadır. TypeOf ile bahsedilen ögenin türü belirtilerek, belirli içerikler seçili sözcük dağılımına göre verilmektedir. Ögelere ait özellikler property ile ve farklı öğelere tanımlayıcılar resource ile verilmektedir. Sözcük dağılımı etiketlemede ihtiyaç duyduğu tüm terimlere sahip olmayabilir. Prefix, yetersiz kalan zamanlarda birden fazla kelime belirleme yeteneğidir (URL-28).

1.5.4. RDFS

RDFS (RDF Schema), W3C tarafından geliştirilmiş olan genişletilebilir bir bilgi sunum dilidir. RDF kaynaklarının tanımlanmasında kullanılacak bir terminoloji (vocabulary) oluşturmak amacıyla rdfs:Class, rdfs:Resource, rdfs:Datatype ile sınıflar ve bunlar arasındaki ilişkiler tanımlanmaktadır. Özelliklerin tanımlanmasında ve özelliklerin sınıflarla rdfs:Property, rdfs:domain, rdfs:range ile ilişkilendirilmektedir. rdfs:subClassOf

kullanımı ile taksonomi oluşturulması sağlanır (URL-29). URI ile RDF’de kaynak hakkındaki açıklamalar sağlanır ama anlamlandırma sağlanamamaktadır. Ortak kelime dağarcığının geliştirilmesi ile RDF bilgisi anlamlandırılmaktadır. RDFS kullanıcılara ait sözcük dağarcığını geliştirmeye olanak sağlayan bir dil sunmaktadır. Ayrıca, RDF’ de bulunan basit ontolojileri tanımlamaktadır (Hart ve Dolbear, 2013). RDFS sözcük dağarcığı Şekil 14’de sınıf ve özellik tipi tanımları ile meydana geldiğini göstermektedir.

```

1 <?xml version = "1.0" encoding = "UTF-8"?>
2 <rdf:RDF
3   xmlns:rdf = "http://www.w3.org/1999/02/22-rdf-syntax-ns#"
4   xmlns:rdfs = "http://www.w3.org/2000/01/rdf-schema#"
5   xmlns:meraeMaps = "http://meraeamaps.gov.me/topo/">
6
7   <rdf:Description
6     rdf:about = "http://meraeamaps.gov.me/topo/Pub">
8     <rdf:type
9       rdf:resource = "http://w3.org/2001/01/rdf-schema#Class"/>
10    <rdfs:subClassOf rdf:resource = "meraeMaps:TopographicObject"/>
11  </rdf:Description>
12 </rdf:RDF>

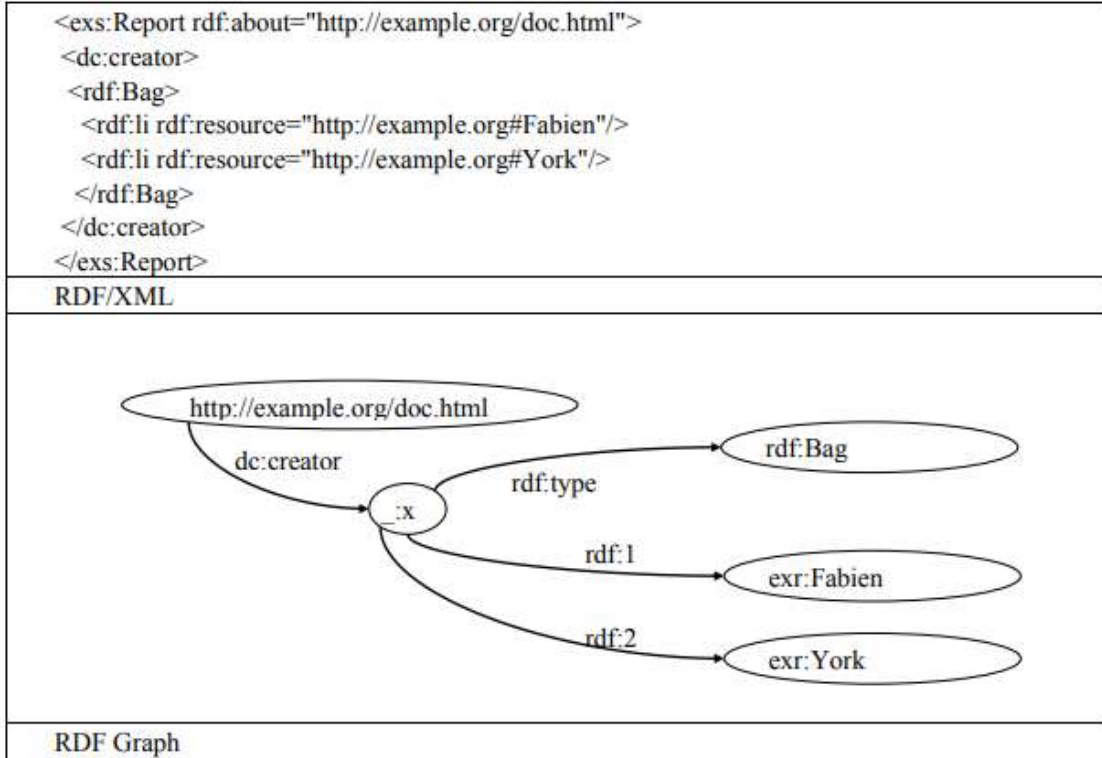
```

Şekil 14. RDFS örneği (Hart ve Dolbear, 2013).

RDFS’in, özelliklerin (properties-predicate) sınıflardan ayrı olarak tanımlanması en önemli özelliğidir. Sınıflardan bağımsız olarak bir özellik oluşturabilir ve rdfs: domain, rdfs:range ile oluşturulan özellik sınıflarla ilişkilendirebilmektedir. Böylece, kolay bir şekilde genişletilebilir ve uyarlanabilir bir yapı elde edilebilmektedir (Devedžić, 2006). Bu yapı, SWT ile ilişkisel veri tabanları birbirinden ayrılmaktadır (Segaran vd., 2009).

1.5.5. RDF Graph

Özne, nesne ve yüklemden oluşan üçlüler RDF ifadeleri temsil edilerek kaynaklar hakkında açıklama yapılmaktadır. Bir ya da birden fazla RDF ifadesi, RDF graph ile düğümler (node) ve çizgiler (arc) ile temsil edilmektedir (Sayers ve Karp, 2004). RDF veri kümeleri RDF çizgesidir (Hayes, 2004).



Şekil 15. RDF çizgesi örneği RDF çizgesine ait RDF/XML serilizasyonu (Gandon vd., 2011)

1.5.6. Ontoloji

Ontoloji, bir alandaki varlıkları tanımlar ve tanımlanan varlıklar arasındaki ilişkileri modeller (Dunwoodie, 2007). Ontolojiler ile verilere ait hiyerarşik yapı ifade edilmektedir. Semantik Web, Web'in bir uzantısıdır. Bilgisayarlar ve insanların beraber çalışmasını sağlamak amacıyla bilgilere tanımlanmış anlamlar kazandırır (Berners-Lee vd., 2001). Ontolojinin Semantik Web'deki rolü, farklı veri kümelerinde kullanılan terimlerde belirsizlikler olduğunda veya ek bilgi ve yeni ilişkilerin keşfedilmesinde veri entegrasyonuna yardımcı olmaktadır. Webin farklı isimler ve yollara sahip olan bir verisi için heterojenlik sorununa çözüm yolu olarak ontolojinin farklı veri kaynakları arasında bilgi alışverişinde kullanılmasıdır (Xue vd., 2015).

Sağlık hizmetleri alanında ortak bir fikir sağlamak amacıyla ontolojilerin uygulanmasını örnek vermek gerekirse tıp uzmanları ontolojileri anlamlı üst düzey bilgilerin oluşturulması için semptomlar, hastalıklar ve tedaviler hakkındaki bilgileri kabul edilmiş kelime hazineleri olarak temsil etmek amacıyla kullanmaktadırlar. İlaç şirketleri ise ilaçlar, dozajlar ve alerjiler hakkında bilgi vermek için kullanır. Tıp ve eczacılık

topluluklarından elde edilen bu bilgileri hasta verileriyle birleştirmek, olası tedavileri arayan karar destek araçları gibi bir dizi akıllı uygulamayı mümkün kılar; ilaç etkinliğini ve olası yan etkileri izleyen sistemler gibi (URL-8).

Geleneksel web dökümanları sadece insanlar için anlam ifade etmektedir. Ontoloji ile kaynakların içindeki kavramlar, kavramlara ait özellikler ve kavramlar ile özelliklerin arasındaki ilişkilerin tanımlanması makineler tarafından anlaşılabilir hale getirilmesini sağlamaktadır (Jacob, 2003). İnsanlar ve bilgisayarlar arasındaki bağlantı miktarı ontoloji yardımı ile artmaktadır (Cardoso, 2007).

Ontoloji içerisinde yer alan her terimin tanımlanmış olmasıdır. Terimlerin anlamlandırılmasında en basit ontoloji olarak adlandırılan taksonomiler ile obje sınıfları ve aralarındaki ilişkiler tanımlanmaktadır (Berners-Lee vd., 2001). Ontoloji, sınıflar (classes), özellikler (attributes), ilişkiler (relations), kurallar (rules), olaylar (events) ve kısıtlamalar (restrictions) olarak bileşenlerden oluşmaktadır. Taksonomiden farklı olarak daha fazla ilişki tanımlanmaktadır (Giri, 2011).

Bilgisayar sistemleri içerisinde bilgiyi temsilen kullanılan ontolojiler ve veri tabanları birbirine benzer özellikler taşıyan iki temel veri modelidir. Eşleşme sağlanarak bu iki veri modeli arasında ilişki kurulabilir. Veri tabanlarına göre bilgilerin daha çok anlamlandırılması için ontolojiler tasarlanmıştır (Martinez-Cruz vd., 2012). Veri tabanı veriyi depolamak için ontolojiler ise insan ve bilgisayar arasındaki ilişkiyi sağlamak tasarlanmıştır. Ontoloji ile mevcut ontolojiler yeniden kullanılmaktadır. Veritabanı kullanımında ise yeni veritabanı oluşturmak gerekmektedir (Sir vd., 2015). Bir alana ait bilgilerin anlamsal birlikte çalışabilirliği, bilgi grafikleri (knowledge graphs) olarak gösterimi ontolojiler ile sağlanabilir. Çeşitli veri kaynakları arasındaki anlamsal ilişkiler ontolojiler ile tanımlanır. Ayrıca, kısıtlamaların ortadan kaldırılması ile veritabanı eksikleri giderilmektedir (Gültepe vd., 2014).

Ontolojinin geliştirilmesi sürecinde tanımlı metotlar kullanılmaktadır. Mevcut ontolojilerin yeniden kullanımında ontolojiler mevcut veriye uygun değilse yeni bir ontoloji oluşturulması gerekir. Ontoloji geliştirme sürecinde amaca uygun metodoloji seçimi yapılır. Metodoloji seçimi yapıldıktan sonra süreçte ontolojiye ait alan ve kapsam belirlenir. Yeniden kullanım için mevcut ontolojilerin sorgulanması gerekmektedir. Önemli olan terimler belirtilerek sınıflar oluşturulmaktadır. Oluşturulan sınıflar arası hiyerarşi belirlenmektedir. Sınıflara ait özelliklerin, özellik yönlerinin tanımlanması ve sınıflara ait örnekler oluşturulmaktadır. Ontolojilerin farklı ontolojiler tarafından kullanılabilir olması

yeniden kullanılabilir ontolojiler ile sağlanmaktadır (Fisseha, 2003). Ontoloji geliştirme sürecinde, mevcut ontolojinin kullanılması ile yeni oluşturulacak ontolojiye göre tasarruf sağlanır ve böylece ontolojinin kalitesi artmaktadır (Lonsdale vd., 2010).

1.5.7. Ontoloji Dilleri

Ontoloji dilleri genelde Semantik Web’de ontoloji geliştirmek ve ontoloji yardımıyla web ortamındaki nesnelere tanımlamak için kullanılmaktadır. Ontoloji dilleri, gerekli sözcük dağarcıklarının oluşturulması ile Semantik Web verilerini anlamsal şekilde belirlemektedir (URL-30). Farklı ölçütlere göre ontoloji dili seçimi yapılmaktadır. Ontolojileri tanımlamakta bu dillerden bazıları RDF, RDFS, SKOS, DAML+OIL ve OWL kullanılmaktadır.

1.5.7.1. RDF/RDFS

RDF, W3C tarafından geliştirilen, web üzerinde bulunan kaynakların tanımlanmasında kullanılan bir standarttır. Web'deki kaynaklar arasındaki ilişkileri tanımlamak için XML dilinde yazılmış bir veri modeli sağlar. Tasarım hedefi olarak verilerin anlam kazanması için model ve sözdizimi sağlamakta standart olarak gösterilmesidir (Kalibatiene ve Vasilecas, 2011). RDF sözdizimi kullanılması ile sınıf hiyerarşilerini, özellik hiyerarşilerini ve aralarındaki ilişkileri (üçlü) tanımlanması RDFS ile sağlanmaktadır (URL-31). RDFS sınırlı bir ontoloji dili olması ile ontoloji tanımlanmasında daha gelişmiş olan DAML+OIL ve OWL gibi ontoloji dilleri ortaya çıkmıştır (Hart ve Dolbear, 2013).

1.5.7.2. SKOS

SKOS, sözlükler, sınıflandırma şemaları, taksonomiler veya diğer herhangi bir yapılandırılmış kontrollü kelime dağarcığının temsili için tasarlanmış bir veri paylaşım standardı olan W3C tavsiyesidir. SKOS, bir RDF ve RDFS’ye dayanan Semantik Web standartları olarak ifade edilmektedir ve temel amacı, bu tür sözlüklerin bağlantılı veri olarak kolay yayınlanmasını ve kullanılmasını sağlamaktır (URL-32). SKOS, RDF söz

dizimi ile üretilmiştir ve böylece temsil edilenler makineler tarafından okunabilmekte ve webde yayımlanabilmektedir (URL-33).

1.5.7.3. DAML+OIL

DAML+OIL, Web kaynakları için anlamsal bir biçimlendirme dilidir. RDF ve RDF Schema gibi daha önceki W3C standartlarını temel alır ve bu dilleri daha zengin modelleme ilkeleriyle genişletir (URL-34). DAML, 2000 yılında ABD hükümeti desteği ile geliştirilmiştir. OntoKnowledge projesi kapsamında OIL ontoloji dili tasarlanmıştır (Gómez-Pérez ve Corcho, 2002). DAML+OIL, çerçeve tabanlı dillerde yaygın olarak bulunan modelleme temel öğelerini sağlar. DAML+OIL (Mart 2001), DAML+OIL'i (Aralık 2000) XML Şeması veri türlerinden gelen değerlerle genişletir. DAML+ OIL, OIL'in birçok dil bileşenini birleştirmek amacıyla orijinal DAML ontoloji dili DAML-ONT'den (Ekim 2000) oluşturulmuştur (URL-34). DAML+OIL dili RDFS dilinin üzerine kurulması ile tanımlamaların makineler tarafından okunabilmesi RDFS ile sağlanmaktadır (Connolly vd., 2001).

1.5.7.4. OWL

OWL (Web Ontology Language), DAML+OIL dilinden türetilmiştir. W3C tarafından DAML+OIL dilinin devamı olarak 2004 yılında önerilmiş standart bir dildir. (McGuinness ve Van Harmelen, 2004). OWL, XML, RDF, and RDF Schema (RDF-S) tarafından desteklenen ağ içeriğinin bilgisayarlar tarafından daha iyi yorumlanabilmesini, biçimsel bir semantik ile birlikte ek sözcük kümeleri sunarak kolaylaştırır. XML, RDF ve RDFS 'ye göre daha fazla anlam gücü sağlayan OWL, RDF/XML sözdizimi formatını temel almaktadır (URL-35).

OWL Lite, OWL DL ve OWL Full olmak üzere OWL dili üç alt dilden oluşmaktadır (McGuinness ve Van Harmelen, 2004). OWL dili, RDF gibi XML tabanlı olup, RDFS'e yeni özellikler ekleyen yapıya sahiptir. OWL ilk olarak, OWL 1 adı ile 2004 yılında tanıtılmıştır. OWL 1 dili 2007 yılında daha geniş kapsamlı olarak OWL 2 versiyonu ile güncellenmiştir. Bu yeni versiyonda OWL dilinin alt dilleri olarak OWL 2 EL, OWL 2 QL ve OWL 2 RL tanımlanmıştır (URL-36). OWL 2 QL ve OWL 2 RL dillerine göre OWL 2

EL daha fazla kavram, iliřiki ve sınıf ifadesine sahip olarak tasarlanmıřtır. OWL 2 QL, Sorgu yanıtıma ve veri miktarının fazla olduđu durumlarda kullanılmaktadır. OWL 2 RL ise ifade etme gúcünü kullanarak mantıklı dűřünmeyi gerekli kılan uygulamalarda kullanılmaktadır (URL-37).

OWL alt dilini ilk olarak OWL Lite oluřturmaktadır (Horridge vd., 2004). Kısıtlama desteđi ve basit sınıflandırma sađlamaktadır. Kısıtlama desteđi 0 ve 1 olarak OWL Lite içerisinde sınırlandırılmıřtır. İkinci olarak OWL DL, mantıksal tanımlamaları içerir. Kısa sürede karar verme, dođru hesaplama ve maksimum anlamlılık desteđi sunmaktadır. OWL DL maksimum anlamlılık desteđi sunar. OWL içerisindeki tüm yapıları belirli sınırlılıklarla kullanmaktadır (Devedđić, 2006). Alt dili olarak OWL Full bilinmektedir (McGuinness ve Van Harmelen, 2004). RDF sözdiziminde kısıt olmadan ifadelerin kullanılmasıyla OWL DL ile getirilmiř kısıtlamalar OWL Full alt dilinde ortadan kaldırılmıřtır. Bu nedenle OWL Full diline ait belge yapısı RDF belge yapısı ile eř tutularak, RDF ile tam uyum sađlanmaktadır (Maier vd., 2009). RDF'den bađımsız söz dizimi desteđi ile maksimum anlamlılık sunmaktadır. Ancak hesaplama ve karar verilebilirlik desteđini sađlayamayabilir (Devedđić, 2006). Kullanıcının ihtiyacına göre OWL Lite belgeleri OWL DL belgeleri olarak ve tüm OWL DL belgeleri de OWL Full belgeleri olarak belirlenmektedir (McGuinness ve Van Harmelen, 2004).

Tablo 1. RDF, RDFS ve OWL dillerinin karşılaştırılması (Ehimwenma vd., 2020)

| RDF | RDFS | OWL | OWL 2 |
|---|--|---|---|
| <p>Belirli bir alandan bağımsızdır.</p> <p>Kaynakları üçlüler biçiminde belirtir ve özne nesne arasındaki ilişkiyi kurar.</p> | <p>Belirli alanı tanımlamak için mekanizma sağlar.</p> <p>Sınıf ve özellik ilişkisini ifade eder.</p> <p>Sınıflandırmada sınıf ve alt sınıfları ifade eder, özellik ve alt özellik, alan (domain) ve aralık (range) kısıtlamasını destekler.</p> <p>RDFS kullanımının ötesinde mantıksal kombinasyon sağlar.</p> | <p>OIL, DAML-OIL gibi birkaç mevcut ontoloji ile uyumludur.</p> <p>RDF kaynak belirtme, RDF sınıfı ve özelliği belirtme yeteneğini genişletir.</p> <p>Sınıflandırma hiyerarşisinde sınıf ve alt sınıfları ifade eder. Sınıflar diğer sınıfların mantıksal kombinasyonları (kesişim, birleşim, eksiklik) ya da belirli nesnelerin listesi olabilir.</p> <p>Özellikleri geçişli, simetrik, işlevsel veya ters olarak ifade ederek RDFS'yi genişletir.</p> <p>Ayrıklığı, eşdeğerliliği, nesnenin bireyselliğini, nicelleştirme ve değer kısıtlamasını ifade eder.</p> <p>Veri tiplerini listelemek için XML şemasına (xsd) dayanır.</p> <p>SHOIN(D) ve NExpTime-complete temellidir.</p> | <p>OWL ile uyumludur. Veri tipi tutmayı iyileştirir ve genişletir.</p> <p>Nitelikli kardinalite kısıtlamalarına sahip ek özellik ve adresler sağlar.</p> <p>Basitleştirilmiş veri modelleme ve genişletilmiş açıklama sağlar.</p> <p>Temel değişim sözdizimi RDF/XML'dir. OWL gibi OWL2'de Turtle, XML, RDF çizge, Manchester gibi diğer sözdizimlerine sahiptir.</p> <p>İfade için değiş tokuşlu OWL2 EL, OWL2 QL, OWL2 RL profillerine sahiptir.</p> <p>STROIQ(D) ve 2NExpTime-complete temellidir.</p> |

1.5.8. SPARQL – GeoSPARQL

SPARQL (SPARQL Protocol and RDF Query Language), W3C tarafından 2008 yılında bir standart haline getirilmiş olan bir sorgulama dili ve protokolüdür. Web'de RDF formatında sunulan veri setlerini sorgular. Güncellemelerle birlikte 2013 yılında SPARQL 1.1 versiyonu yayımlanmıştır (URL-38).

RDF veri modeli, Semantik Web verilerinin bağlantılı hale getirilebilmesi ve yayınlanabilmesi için kullanılmaktadır ve bu verilerin sorgulanması ve üzerinde değişiklik yapılması amacı ile modele özgü SPARQL sorgu dili oluşturulmuştur. (URL-39).

SPARQL standardı Semantik Web ile tamamen uyumlu olarak çalışmaktadır ve RDF sorgulama dilini farklı sorgulama araçları desteklemektedir. Yapısal ya da söz dizimsel olarak SPARQL, ilişkisel veri tabanlarını sorgulayan SQL diline benzetilmektedir (Valle ve Ceri., 2011). Web’de verilerini RDF olarak sunmuş tek bir kaynağı veya birden fazla kaynağı aynı anda sorgulayabilmekte ve LOD (Linked Open Data) dosyaları ve grafik veri tabanlarında sorgulama yapabilmektedir (Sikos, 2015).

SPARQL dilinin 4 tane sorgu formları bulunmaktadır.

SELECT, sorgusu ile sorgulanmak istenen verileri ve kaynaklar içerisinde sonuç çıktısı olarak alınmak istenen değişken değerler tanımlanmaktadır. En çok kullanılan sorgu türüdür.

CONSTRUCT, üçlüler veri kaynağından değişime uğramadan alınabilmektedir. Alınan bu veriler ile yeni üçlüler oluşturulabilmektedir (DuCharme, 2013). Yeni bir RDF grafiği olarak sorguya cevap vermekte, grafikler diğer grafikler ile birleştirilebilmekte ve RDF verilerini kopyalama ve dönüşümüne imkân vererek RDF depolarına eklenebilmektedir.

ASK, Veri setinde yapılacak sorguya ‘true’ veya ‘false’ olarak yanıt verilip verilemeyeceğini belirtir. Genellikle iki farklı veri setinde aynı bilginin olup olmadığını denetlemek için kullanılmaktadır. SELECT ve CONSTRUCT sorguları kullanılmadan endpointlere ait sorgulamalar yapılabilmektedir (Hebeler vd., 2011).

DESCRIBE sorgusu ile cevapta kaynağı tanımlayan ayrıntılar tek bir RDF grafiğinde istenmektedir (Sakr vd., 2018). Sorguya yanıt olarak hangi üçlülerin sorguya yanıt olarak istenildiği kararı sorgu işlemcisine ait olduğundan kararsızlıklar oluşmakta ve böylece yaygın kullanılmaktadır (DuCharme, 2013).

SPARQL (SPARQL Protocol and RDF Query Language) sorgu dilinde XML biçimi ilk versiyonunda sorgu sonuçları için desteklenmiştir (URL-38). JSON, CSV ve TSV formatlarında SPARQL 1.1 versiyonu ile sorgu alınabilmektedir. SPARQL 1.1 versiyonunda dört temel sorgu formlarından farklı grafik güncelleme ve yönetim işlemleri eklenmiştir (URL-40). Bazı matematiksel işlemler ile ortalama, minimum ve maksimum değer bulunması gibi SPARQL 1.0 sürümünde desteklenmemektedir. Fakat matematiksel sorgulamalara SPARQL 1.1 sürümü ile sağlanmaktadır. SPARQL 1.1 versiyonunda bu ana sorgu formları dışında birçok farklı form da bulunmaktadır. Grup içindeki değerler SUM fonksiyonu toplanır. En küçük ve en büyük sayısal değer MIN/MAX ile bulunmaktadır. SAMPLE fonksiyonu ile elde edilen verilerinden sorgu sonucunda rastgele bir değer

ulaştırılması sağlanmaktadır. Ortalama değerin hesaplanması AVG ile sağlanmaktadır (URL-16). Sorgu cevaplarının belirli bir sıraya konulması 'ORDER BY' ile yapılmaktadır. 'LIMIT' ile sorgulamalara verilen cevaplar kısıtlanabilir. 'SERVICE' sorgu formu ile SPARQL 1.1 özelliklerini sağlayan diğer SPARQL Endpoint (sorgulamaya olanak sağlayan bir arayüz) servislerine erişilir. Sorgulanan verinin sadece bir kez gösterilmesi ve tekrarların silinmesi 'DISTINCT' ile sağlanır. Farklı veri setlerinin birleştirilmesi 'UNION' ile sağlanmaktadır (URL-38). SPARQL veri setleri üzerinde matematiksel/fonksiyonel işlemler içeren sorgulamalar da yapılabilir. Veri kümelerine ait verilerin SPARQL Endpoint olarak yayınlaması ile SPARQL sorgularını doğrudan çalıştırılarak sorgulamalar direkt yapılabilir. WHERE ve FROM, verilerin hangi veri setlerinden seçileceğini belirlemek için kullanılmaktadır. Semantik Web'e ait tüm kaynakları sorgulayabilecek SPARQL Endpoint yoktur. Ama oldukça büyük veri kapasitesine sahip veri kümelerini sorgulayabilen bazı uç noktalar bulunmaktadır. Datahub, DBpedia, Geonames, Sindice ve URIBurner veri kümeleri örnek verilebilir (Sikos, 2015). DBpedia'ya sorgulayan Vikipedi'nin oluşturduğu arayüzler en çok kullanılanlar arasındadır (URL-41). OS de verilerini DBpedia'ya benzer şekilde SPARQL Endpoint olarak sunmuştur (URL-42). Sunulan verilerin sorgulanması ve uygulamalarda kullanılması bu arayüzler aracılığıyla sağlanmaktadır. Sunulan, sorgulanan verilerin uygulamalarda kullanılması hizmeti veri tabanlarında SQL tarafından, Semantik Web'de ise SPARQL tarafından sunulmaktadır. Şekil 16'da Avrupa'nın farklı ülkelerinde 20/01/2021 tarihine kadar toplam COVID-19 pozitif vaka sayısını almak için SPARQL sorgusunu göstermektedir.

```

SELECT ?countryname xsd:integer (?totalPositiveCases)
WHERE {
  ?node dbpprop:countryName ?countryname.
  ?node covid:total_positive_cases ?totalPositiveCases.
  ?node covid:continent ?continent.
  ?node dc:date ?mydate.
  FILTER regex (?mydate, "2021-01-20") .
  FILTER regex (?continent, "Europe").
  FILTER (?totalPositiveCases > 300000).
}

```

Şekil 16. SPARQL örnek sorgusu (Ali vd., 2022).

Semantik Web ile GeoSPARQL, konumsal verilerin sunulması ve sorgulanması için OGC (Open Geospatial Consortium) tarafından 2011 yılında geliştirilmiştir. OGC

GeoSPARQL standardı, RDF veri modelinde coğrafi detayları, geometrileri ve konumsal ilişkileri tanımlamaktadır. GeoSPARQL standardı ile SPARQL'e sınıf, ilişki ve fonksiyon tanımlamaları eklenmiştir. Coğrafi detayları sunmak için bir sözlük, konumsal filtreleme fonksiyonları ve sorgu dönüşüm kuralları olarak üç ana bileşeni bulunmaktadır. GeoSPARQL, konumsal verilerin RDF veri modeli olarak temsil edilebilmesi için ontolojiler tanımlamaktadır. SPARQL dilinin eklentisi olarak GeoSPARQL ile veriler sorgulanır ve filtreleme işlemi yapılabilmektedir (Perry ve Herring, 2012). SPARQL sorgularının geliştirilmesi ile konumsal verilere ulaşılabilir (Osman, 2018).

GeoSPARQL, nitel ve nicel hesaplamaları kullanan sistemlere sahiptir. Nicel sistemlerde özelliklere ait geometriler ile birlikte topolojik ilişkiler hesaplanır. Geometrinin bilinmediği özellikler için nitel sistemler topolojik çıkarımlar yapılmaktadır. GeoSPARQL tarafından nicel uygulamalardan elde edilen sonuçlar nitel uygulamalar tarafından kullanılmaktadır (Perry ve Herring, 2012).

GeoSPARQL, 'Konumsal Nesne', 'Coğrafi Detay' ve 'Geometri' sınıflarından oluşmaktadır. 'Konumsal Nesne' sınıfının alt sınıfı olarak tüm coğrafi detaylar ve geometriler temsil edilmektedir. 'Coğrafi Detay' tanımı olarak ISO 19156:2011, 'Geometri' tanımlamaları olarak da ISO 19107 standardının tanımlamaları kullanılmıştır. geo:SpatialObject sınıfı ile geo:Feature ve geo:Geometry alt sınıflarını içeren ontoloji konumsal verilerin oluşturulmasını ve sorgulanmasını sağlar. Ayrıca geo:hasGeometry özelliği ile geometri ve özellikler arasındaki bağlantılar, geometrinin iki boyutlu olarak GML veya WKT (ogc:asGML, ogc:asWKT) veri tipleri ile ifade edilmektedir. Daha sonraki yıllarda KML, GeoJSON gibi çok kullanılan formatların sözlüğe ekleneceği ifade edilmiştir. Konumsal veriler arasında ilişkilerin kurulması ve aralarındaki ilişkilerden daha fazla anlamlar ortaya çıkarılması GeoSPARQL sağlamaktadır (Battle ve Kolas, 2011).

GeoSPARQL'in ikinci bileşeni olarak ele alınan fonksiyonlar, konumsal filtreleme fonksiyonlarıdır. GML literalleri (ogc:GMLLiteral) ve WKT literalleri (ogc:WKTLiteral) ile SPARQL FILTER fonksiyonu desteklenerek bu fonksiyonlar kullanılmaktadır. Destek sağlandığında, uzaklığı belirten fonksiyon 'ogcf:distance', bir çember oluşturan fonksiyon ogcf:buffer, kesişimleri sunan fonksiyon ogcf:intersection, birleşimleri sunan fonksiyon ogcf:union ve daha birçok fonksiyon kullanılabilir duruma gelmektedir.

GeoSPARQL'de sorgu dönüşüm kuralları son bileşendir. Örnekleemelerin 'ogc:SpatialObject' sınıfı ile tanımlanması ve 'ogc:asWKT' ile 'ogc:asGML' veri ilişkilerinin örneklemelerin geometri temsilinde kullanılması gerekmektedir.

Kullanıldığında OGC tarafından belirlenen dokuz konumsal ilişki (eşittir, ayrıktır, kesişir, dokunur, keser, içindedir, içerir ve çakışır) işler duruma gelmektedir (Kara, 2014).

1.5.9. Unifying Logic

Birleştirici mantık (unifying logic) katmanı, W3C çalışma grupları tarafından standart geliştirme çalışmaları devam etmektedir. Daha çok çıkarsama yapma süreciyle ilgili bir katmandır. Çıkarsama sürecinde yazılımlar, ontolojide tanımlanan kısıtlamaları ve sınırlılıkları ya da Rule Interchange Format (RIF) dikkate almaktadır. Webde bulunan mevcut verilerden yeni anlamlar çıkarılmaktadır. Örneğin, bir yazılım web üzerinden elde etmiş olduğu farklı iki kaynağın e-posta adreslerinin aynı olduğunu belirlerse, bu iki kaynağı (owl:sameAs) özelliği ile birbirine bağlayabilir ve bu iki kaynağın aynı kişiyi tanımlayan kaynaklar olduğu çıkarılması yapılabilmektedir (Karalar ve Özdemir, 2013).

1.5.10. Proof

Kanıt (proof), W3C çalışma grupları tarafından standart geliştirme çalışmaları devam etmektedir. Web üzerinden erişimi sağlanan kaynakların doğruluğuyla ve sonucun çıkarsama sürecinde elde edilen yeni bilgilerin hangi kanıtlara dayandırıldığı ile ilgilenen katmandır. Proof için güvenilir girdiler, Semantik Web mimarisinin üst katmanında gösterilen sonuçlara güvenilebileceği anlamına gelir (Karalar ve Özdemir, 2013).

1.6. Bağlantılı Veri

2006 yılında Tim Berners-Lee tarafından bağlantılı veri düşüncesi önerilmiştir ve W3C tarafından desteklenmiştir. Semantik Web, bilgiye iyi tanımlanmış anlamlar verilen, bilgisayarların ve insanların işbirliği içinde daha iyi çalışmasını sağlayan mevcut Web'in bir uzantısı olarak tanımlanmaktadır. Bilgi, bilgisayarların kaynaklar arasında ilişkiler kurabileceği şekilde yapılandırılmaktadır. Bu ilişkiler, bağlantılı veri olarak adlandırılır (URL-43).

Webde veriler farklı formatlarda (PDF, TIFF, CSV, Excel ya da farklı düz metin vb.) yayınlanmıştır. Farklı formatlarda yayınlanan bu verilere erişmek ve kullanmak

zorlaşmaktadır (Wood vd., 2014). Web’de yapılandırılmamış olan metin belgeleri HTML sayfaları şeklinde yayımlanır ve web sayfaları üzerinde sorgulama anahtar kelimeye dayalı arama motorları kullanılarak yapılmaktadır. Webde yayınlanan yapılandırılmış veriler bağlantılı veri sayesinde daha yüksek anlama sahip sorgulamalar yapılmaktadır. Yapılandırılmış verilerin kısıtlamaları bağlantılı veri ile kaldırılarak herkesin kullanımına izin verilmektedir. Böylece, farklı kaynaklardan olan veriler kolaylıkla paylaşır, sorgulanır veya birleştirilir (Ngomo vd., 2014).

Semantik Web ile sadece veri Webde paylaşılmaz, aynı zamanda bağlantı kurulması da sağlanır. Böylece bir kişi veya bilgisayarlarca Webde bulunan veriler keşfedebilir. Verinin bir kısmına erişim sağlandığında diğer ilgili verilerinde bulunması Bağlantılı veri yaklaşımı ile gerçekleştirilmektedir (Berners-Lee, 2021). SWT kullanılarak verilerin yayınlanması ve birbirleri ile bağlantılı hale getirilebilmesi için gerekli olan kurallar;

- “Şey”leri adlandırmak için URI’ler kullanılır,
- İnsanlar http URI’leri kullanarak bu adları arar,
- Yararlı bilgilerin sağlanması bir URI arandığında, RDF ve SPARQL gibi standartları kullanılır,
- URI'lere bağlantılar eklenerek daha çok şey keşfedilir (Berners-Lee, 2006).

Bağlantılı Veri yaklaşımında birinci ilke; URI referansları kullanılarak klasik Web belgeleri, tüm gerçek dünya nesnelere ve soyut kavramları tanımlanmaktadır. Örnek olarak insanlar, coğrafi yerler, araçlar gibi somut şeyler ve ‘birini tanımak’ gibi soyut bir ilişki URI’lerle tanımlanabilir. URI tanımlayıcıları, Geleneksel Web’de sadece dökümanları tanımlamaktadır. Semantik Web’de ise adı olan her şeyin tanımlanmaktadır. HTTP protokolleri, Web’in küresel erişim mekanizmasıdır. İkinci ilke, HTTP protokolü kullanımı ile URI’lerle tanımlanan gerçek dünya nesnelere, kavramlara ve ilişkilere erişebilmektedir. Farklı formatlarda kaynaklara sunulabilir (Kara, 2014). HTTP protokolü üzerinden tanımlı nesneye ait dokümanlara erişim HTTP URI’leri kullanılarak sağlanır (Heath ve Bizer, 2011). Üçüncü kural ile URI aranmasında bilgilere erişim anlamsal ağ standartları kullanılır. Veriler için değerlendirme veya seçim yapılması yayınlanan veriler ile ilgili meta verilere istemciler tarafından erişim sağlandıktan sonra yapılmaktadır. Standart bir veri modeli kullanılarak yapılandırılmış dokümanlar Web’de sunulur. Bunun için makineler tarafından anlaşılabilen RDF dilinin kullanımı önerilmektedir (Kara, 2014). Bağlantılar sayesinde gerçek bir veri ağının oluşumunun sağlanması son kuraldır. Kurulan RDF bağlantıları ile veri ağında gezinerek kullanıcılar istediği bilgiye ulaşabilmektedir

(Sikos, 2015). RDF dokümanlarını sorgulamak için SPARQL dilinin kullanılmasını, veri setlerinin insanlar tarafından da anlaşılabilirliği için HTML olarak da sunulması gerektiği vurgulanmaktadır (Kara, 2014).

Bağlantılı veriyi tanımlamak için verileri bağlantılı veri ilkelerine uygun şekilde yayınlamak gerekmektedir. Bağlantılı veri, farklı veri kaynakları arasında birlikte işlenebilirliği, yeniden kullanılabilirliği ve semantik sorgular yapabilmeyi sağlamaktadır. Verilerin semantik olarak işlenmesi, görselleştirilmesi, Webde yayınlanması ve paylaşılmasını Bağlantılı Veri teknolojileri ile sağlanmaktadır. Bir alan için ortak kavramsallaştırmalar olarak tanımlanan ontolojilere ya da harici Web kaynaklarına, bağlantılar kurularak verilerin semantik tanımları oluşturulur. Kaynak veri sözcükleri ile global olarak yayınlanmış kaynakların kavramları arasında bağlantı kurmak, farklı veri kaynaklarını ve veri kümelerini birleştirilir. Böylece, verilere daha anlaşılır, keşfedilebilir ve kullanılabilir hale getirilir. Verilerin birlikte işlerliğini ve entegrasyonunun geliştirilmesi ile otomatik anlam çıkarsaması sağlanır ve veri tekrarı önlenmektedir (Ulutaş Karakol vd., 2018).

Devlete ait, açık ve tekrar kullanılabilir biçimde bağlantılı veri prensiplerine uygun olarak yayınlanan veriler bağlantılı açık veriler olarak tanımlanmaktadır. Bağlantılı veriler insanların desteklenmesi amacıyla 5 yıldızlı bir derecelendirme sistemi Tim Berners-Lee tarafından geliştirilmiştir (Berners-Lee, 2006).

- ★ : Açık lisansla mevcut herhangi bir formatta Web üzerinde veriler olmalı,
- ★★ : Makine tarafından okunabilir yapılandırılmış veriler olarak bulunmalı,
- ★★★ : Özel olmayan bir biçimde veriler olmalı,
- ★★★★ : Şeylerin tanımlanmasında W3Cnin açık olan standartlarına göre veriler yayınlanmalı,
- ★★★★★ : Diğer kullanıcılara ait verilere bağlantılar kurulmalıdır.

Geleneksel Web’de sadece Web dokümanları hiperlinkler ile ilişkilendirilir. Ayrıca bu hiperlinkler ile farklı sunuculara ait veriler tek bir bilgi alanına bağlanabilmektedir. Sir Tim Berners-Lee tarafından öne sürülen Semantik Web kavramı ile sadece dokümanların birbirlerine bağlanabildiği, veri setleri arasında bağlantı sağlamayan 1990’ların son döneminde Geleneksel Web’e yeni bir bakış açısı getirilmiştir. Semantik Web, sadece dokümanların değil, veri setlerinin de birbiriyle bağlanmasına olanak vermesi amacıyla

tasarlanmıştır (Bizer vd., 2011). Semantik Web, Geleneksel Web’de farklı servislerden sunulan farklı formattaki verilerin aynı anda değerlendirilmesine olanak sağlayan küresel bir veri tabanı olarak düşünülebilir. Birçok farklı veri formatının kullanılabilmesine olanak sağlayan RDF veri modeli ile veriler arasında bağlantı kuramama, verilerin birlikte çalışmaması gibi sorunlara çözüm sağlanmaktadır (Hart ve Dolbear, 2013). Bağlantılı veri sayesinde RDF linkleri ile veri kümelerinde bulunan veriler arasında bağlantı kurularak bağlantılara sahip veriler elde edilmektedir. URI ile benzersiz bir tanımlama sağlanmaktadır. HTML, belgelerde içerik formatı olarak kullanılmaktadır. HTTP erişim mekanizması olarak kullanılmaktadır. Webde bulunan belgelerde standart format HTML iken bağlantılı verilerde yapılandırılmış olan veriler tek bir formata sahip biçimde RDF veri modeli kullanılarak yayınlanmaktadır. Hiperlinkler, geleneksel Web’de sadece Web dokümanlarını ilişkilendirmektedir. Semantik Web ve Bağlantılı Veri yaklaşımında ise ilgili olduğu düşünülen her şeyin ilişkilendirilmesinde kullanılmaktadır. Bu linkler, Bağlantılı veri kapsamında RDF linki olarak adlandırılmaktadır. Farklı ad alanlarındaki (namespace) URI’ler birbirine bağlanır ve farklı veri setlerindeki kaynaklar birbiriyle ilişkilendirilir. RDF üçlüsü ile RDF linkleri ifade edilmektedir. Bu üçlüye ait özne, veri setinin ad alanındaki bir URI’yi ifade etmektedir. İlişki ve nesne URI’leri, diğer veri setlerinin ad alanlarındaki URI’leri işaret etmektedir (Health ve Bizer, 2011).

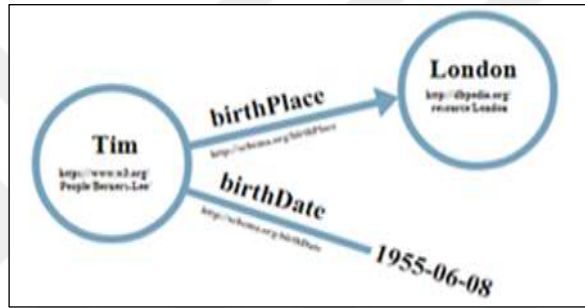
Şekil 17’de Tim kişisine ait doğum tarihi ve doğum yeri hakkında açıklama yapılmıştır. Her ifadenin tabloda kendi hücresi vardır. Bağlantılı verilerde bu ifadeler genellikle üçlüler (triples) denir. Bunun nedeni, her üçlü ifadenin özne (subject), yüklem (predicate) ve nesne (object) olmak üzere üç bölümü olmasıdır. RDF, bağlantılı veriler için standart veri formatıdır.

| Subject | Predicate | Object |
|------------|-------------------|---------------|
| <u>Tim</u> | <u>birthPlace</u> | <u>London</u> |
| <u>Tim</u> | <u>birthDate</u> | 1955-06-08 |

Şekil 17. Üçlü ifadenin tablosal sunumu (URL-44).

Üçlü bir tablo kullanmak yerine, RDF verilerini bir grafik olarak görselleştirebiliriz. Şekil 18’de verilen bağlantı örneği Tim kişisine ait doğum yerini göstermektedir. London, ilk üçlüdeki nesne olarak farklı kaynaklara başka bağlantılar ile bağlanmaktadır. Böylece farklı kaynaklardan gelen verilerin bağlanması ve sorgulanması sağlanır. Diğer üçlüde değişmez bir değeri “155-06-08” nesnesi ifade etmektedir. Herhangi bir özellik taşıması, kaynak olmamasından dolayıdır. Burada;

- Özne URI, <https://www.w3.org/People/Berners-Lee/> ile,
- Nesne URI, <http://dbpedia.org/resource/London> ile,
- Yükleme URI <http://schema.org/birthPlace> ile temsil edilmektedir.



Şekil 18. RDF verilerinin grafik sunumu (URL-44).

1.7. Konumsal Bağlantılı Veri

Konumsal veriler, dünyadaki varlıkların koordinat, adres vb. konum bilgisi içeren veya harita üzerinde bir lokasyona bağlanmış her türlü veri olarak tanımlanabilir (URL-45). Meta veri, konumsal veri ile sunulmaktadır. Kullanıcıya verinin kaynağını, verinin oluşumunda kullanılan standartları ve kim tarafından ne zaman oluşturulduğu bilgilerini sağlamaktadır (URL-46). Birçok kamu ya da özel kuruluşlar tarafından toplanan konumsal veriler, kullanıcılar tarafından kolay erişim sağlanması için açık veri olarak yayınlanmaktadır. Güvenilir ve açık kriterlerle hazırlanan ve paylaşılan mekansal veri ise sağlık ve sosyal politikalarının doğru şekil, zamanlama ve ölçeğini tespit etmek için oldukça önem arz etmektedir. Bu noktada verileri düzenli toplayabilen sistemlerin ve büyük veri madenciliğinin hayati öneme sahip olduğunu söylemek mümkündür (URL-47).

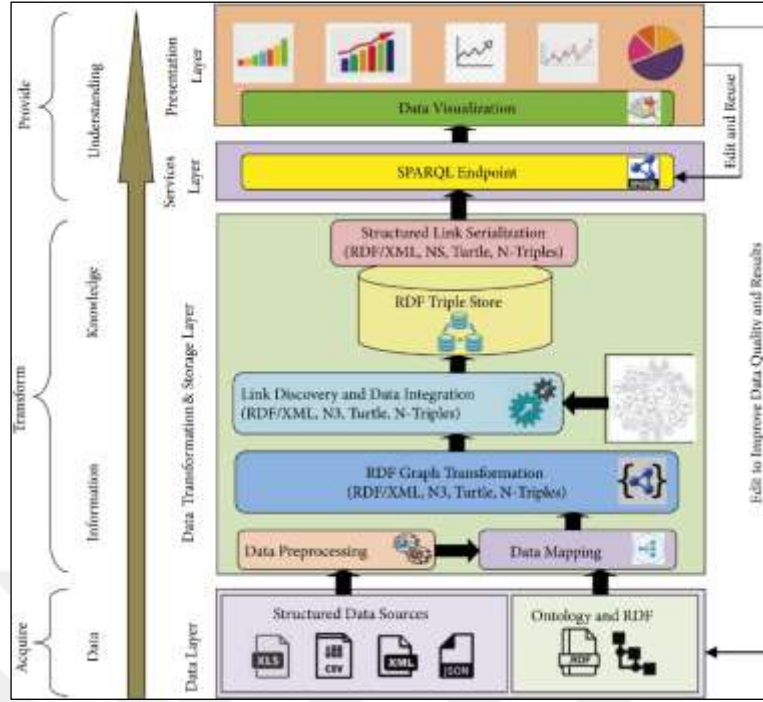
Çoğunlukla, webde açık veri kümeleri, meta verilerle tablo biçimlerinde (CSV, Excel, PDF vb.) temsil edilir ve paylaşılır. Fakat birbirinden farklı formatlarda yayınlanan

bu veriler, veri entegrasyonunu, karşılaştırmasını ve yeniden kullanımını zorlaştırmaktadır. Farklı veri kümelerine ve gerekli tüm bilgilere erişim Bağlantılı Veri yaklaşımı ile sağlanmaktadır (Ali vd., 2022). Büyük ölçekli verilerin birlikte çalışabilirliğini kolaylaştıran Bağlantılı Veriler birbiriyle ilişkili veri kümelerinin toplanmasını sağlar. Konumsal Bağlantılı Veri, konumsal ve zamansal özniteliklerle zenginleştirilen Bağlantılı Veri demektir. Bu veriler, veriyi daha erişilebilir, paylaşılabilir ve anlaşılır hale getirerek kendi değerini artırmaya yardımcı olmaktadır. Ayrıca sorgulama ve akıl yürütmeyi sağlamaktadır (Potnis ve Durbha, 2016). Link kurulacak veri setlerinin bulunması ve mevcut konumsal veriler ile veri setleri arasındaki ilişkilerin tespit edilmesi konumsal verileri daha anlamlı hale getirmektedir. Konumsal verilerin bağlantılı veri olarak yayınlanması ile verilerin sorgulanabilmesi için GeoSPARQL sorgu dili kullanılmaktadır ve webde bu sonuçlar görüntülenebilir. Triple Store ismindeki veri tabanlarında bağlantılı konumsal veriler saklanarak RDF sorgulamaları yapılmaktadır (Kara ve Cömert, 2016).

1.8. COVID-19 Bağlantılı Veri Uygulamaları

Uluslararası düzeyde COVID-19 verilerinin semantik tanımlarını oluşturmak için ilgili literatürde mevcut proje, girişim ve akademik çalışmalar incelenmiştir. Mevcut çalışmalar içinde SWT ve Bağlantılı Veri İlkeleri kullanılarak gerçekleştirilen çalışmalara bakıldığında, ontoloji geliştirme ve bağlantılı veri yayınlamaya yönelik akademik çalışmalara rastlanmıştır. Bu bölümde yapılan çalışmalar değerlendirilerek tez çalışmasının bu çalışmalardan farklı yönleri vurgulanmıştır.

Linked Open COVID-19 Data (LOCD) , LOD Cloud alan uzmanları, bilim adamları ve web geliştiricileri tarafından herhangi bir kısıtlama olmaksızın COVID-19 veri kümelerine erişilmesini ve yeniden kullanılmasını sağlamak için Bağlantılı Veri ilkelerini kullanarak COVID-19 veri kümelerini toplamak, yayınlamak ve paylaşmak için bir referans noktası sağlamayı amaçlamaktadır. LOD Cloud diyagramı, sağlık hizmetleri ve klinik deney araştırmalarının uygulanması için nispeten yenidir. Gelişmiş COVID-19 veri izleme ve görselleştirme için LOCD sisteminin tasarımını ve geliştirilmesini sağlamak için Şekil .. de verildiği gibi çok katmanlı bir mimari önerilmiştir.



Şekil 19. LOCD sisteminin çok katmanlı mimarisini (Ali vd., 2022).

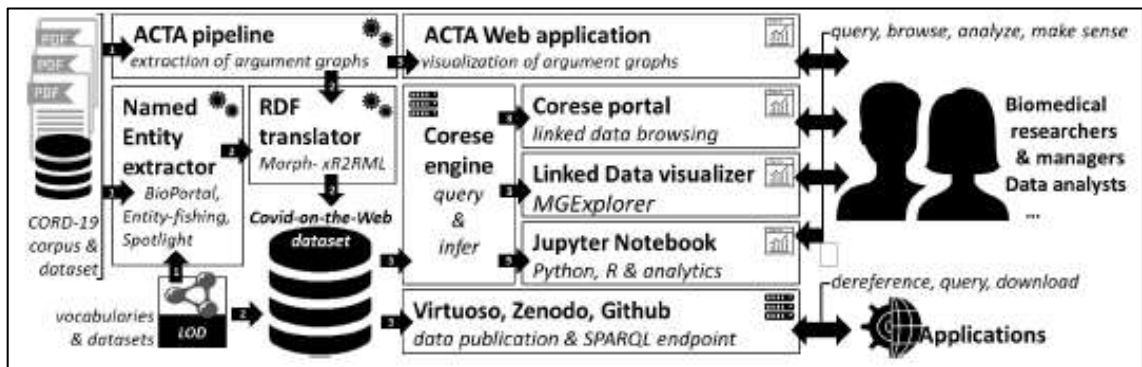
LOCD, SWT ve Bağlantılı Veri İlkelerini kullanarak COVID-19 verilerine kullanıcıların erişimini sağlamak ve kullanıcılar tarafından yeniden kullanımına imkân vermek için COVID-19 verilerinin toplanması, yayınlanması ve paylaşılması için bir referans noktası sağlamayı hedeflemektedir (Ali vd., 2022) Farklı COVID-19 verilerini eşleştirmek için kullanılan ontoloji kullanıcı erişimine açık değildir. Kullanım senaryoları sadece istatistiksel sonuçları değerlendirmeye yöneliktir.

Belirtilen sınırlamalardan birincisi, sistem mevcut veri boyutu ve örnek sorgular ile tatmin edici sonuçlar gösterdiği ancak COVID-19 veri boyutu her geçen gün artmakta olduğudur. Bu nedenle, yaygın uyarlamalarından önce farklı açılardan (örn. performans ve sorgu sonuçları vb.) daha sağlam ve titiz değerlendirmelere ihtiyaç duyulmaktadır. İkinci sınırlama ise referans ontoloji olarak OntoCOVID ontolojisinin ek kavramlar ve özelliklerle genişletilmediğidir. Gelecek çalışmalarında referans ontolojinin ek kavramlar ve özelliklerle genişletilmesi ve ilgili ontolojiler ve sözlüklerle birbirine bağlanmasıdır (Ali vd., 2022).

Michel ve arkadaşları (2020), biyomedikal araştırmacıların erişimini, sorgulamasını ve COVID-19 ile ilgili literatürü anlamlı yapmak amacıyla Covid-on-the-Web Projesi tarafından sağlanan veri ve yazılım kaynaklarını tanımlamıştır. Bu kapsamda The Covid-19 Open Research Dataset (CORD-19) veri seti kullanılmıştır. Ayrıca CORD-19 korpusu;

DBpedia, Wikidata ve diğer BioPortal kelime hazineleri ile bağlantılıdır. Diğer yandan karar vermeyi desteklemek için çizge görselleştirmelerini üreten ve argümanları otomatik olarak çıkararak klinik deneylerinin analizinde klinisyenlerin işini kolaylaştırmak için tasarlanan ACTA platformu uyarlanarak geliştirilmiştir. Proje kapsamında Bilgi temsili, metin, veri ve argüman madenciliği ile veri görselleştirme ve arama yetenekleri geliştirilerek RDF veri seti ve yazılım kaynakları üretilmiştir. Veri setleri CORD-19 Named Entities Knowledge Graph ve CORD-19 Argumentative Knowledge Graph olmak üzere iki temel knowledge graph tan oluşur. Knowledge graph, Argument Model Ontology (AMO) kullanır ancak ontoloji erişilebilir değildir. Knowledge Graph and Services to Advance COVID-19 Research Dataset (Covid-on-the-Web Dataset), COVID-19 Open Research Dataset (CORD-19) bilimsel makalelerini analiz ederek üretilen iki ana bilgi grafiği sağlayan bir RDF veri kümesidir. CORD-19, DBpedia, Wikidata ve diğer BioPortal sözlükleriyle bağlantılı olan adlandırılmış varlıkları açıklar. Ek olarak argümantatif grafiklerin çıkarılmasını ve görselleştirilmesini otomatikleştiren bir araç olan The Argumentative Clinical Trial Analysis (ACTA) kullanılarak çıkarılan argümanları tanımlar (Michel vd., 2020).

Yapılan çalışmalar çoğunlukla biyomedikal ontolojilere odaklanmaktadır. Sonuç olarak, tanımlanan adlandırılmış varlıklar genler, proteinler, ilaçlar, hastalıklar, fenotipler ve yayınlarla ilgilidir. Bu yaklaşımda, yalnızca BioPortal'daki biyomedikal ontolojiler dikkate alınmamıştır. Kapsamı Web of Data'daki ana merkezler olan DBpedia ve Wikidata iki genel bilgi tabanları ile genişletilmiştir (Michel vd., 2020). Fakat CORD-19, bugüne kadar toplanan en büyük COVID-19 bilgi koleksiyonu olarak kabul edilse de, veri setinin çoğunluğu yapılandırılmamış veriler içermektedir (Wang vd., 2022).



Şekil 20. Covid-on-the-Web overview (Michel vd., 2020)

Wang ve diğeri (2022), Linked Data Set of COVID-19 Knowledge Graphs (OpenKG-COVID19) ile bilginin kullanım değerini iyileştirmek ve araştırmacılara COVID-19 ile mücadelede yardımcı olmak amacıyla, COVID-19 ile ilgili mevcut en büyük KG'lerden biri olan OpenKG-COVID19 adlı birleşik bir bağlantılı veri seti oluşturulmuş ve yayınlanmıştır. Bir KG, akıllı soru cevaplama, semantik aramalar, öneri sistemleri, görselleştirme analizi ve karar verme desteği için kolaylık sağlamaktadır. OpenKG-COVID19, araştırma yayınları, tıbbi kılavuzları ve ansiklopedi web siteleri gibi farklı kaynaklardan elde edilen encyclopedia, concept, medical, research, event, health, epidemiology, goods, prevention ve karakter konularını kapsayan altçizgeler (subgraphs) içermektedir. OpenKG-COVID19, 2572'den fazla concepts, 329.600 entities, 513 properties ve 2.687.329 facts'a sahiptir ve veri seti sürekli olarak güncellenecektir. Ayrıca, çalışmada kullanıcılar için daha yararlı bilgiler sunmak amacıyla COVID-19'un daha zengin ve daha çeşitli konularını kapsamak amacıyla, mevcut girişimler, COVID-19 KG'lerinden oluşan bağlantılı bir veri seti olan OpenKG-COVID19'u oluşturmak için genişletilmiştir. OpenKG-COVID19 kapsamında kullanılan temel yüklemeler (predicates), rdfs:label, rdfs:subClassOf, rdfs:domain, rdfs:range, owl:sameAs olarak belirlenmiştir. Fakat OpenKG-COVID19 sürekli olarak güncellense de, güncelleme sıklığı günlük değildir, bu da bazı bilgilerin güncel olmamasına neden olabilir ve aşağı akış görevleri için uygunsuzluğa neden olabilir. Geliştirilen her KG graph gibi OpenKG-COVID19 çizgesinin kalitesinin değerlendirilmesi önemlidir. Ancak değerlendirmek için background bilgitabanı veya şema olmadığı için değerlendirme uzmanlar tarafından belirlenen örnek gerçeklere dayanarak manuel olarak gerçekleştirilmiştir. OpenKG-COVID19 mevcut durumda erişime açık değildir.



Şekil 21. OpenKG-CoVID19'a genel bakış

Representing COVID-19 information in collaborative knowledge graphs: the case of Wikidata, COVID-19 pandemisi hakkında genel bilgiler için bir bilgi tabanı olarak hizmet etmektedir. Daha geniş bilgisayar destekli işbirlikçi çalışma ve bağlantılı açık veriler için web ölçeğinde SPARQL aracılığıyla Resource Description Framework (RDF) formatında kaydedilen veritabanlarından bilgi almak ve işlemek için kullanılan anlamsal bir sorgulama dili de dâhil olmak üzere uzmanlar, otomatik araçlar ve halk tarafından neredeyse gerçek zamanlı olarak birçok yolla sorgulanabilir. Wikidata ile COVID-19 için oluşturulan zengin bilgi grafiği, karar desteğinin yanı sıra eğitim ve bilimsel araştırma gibi amaçlarla görselleştirilebilir, araştırılabilir ve analiz edilebilir (Turki vd., 2021).

RDF formatında açık bir semantik kaynak olarak Wikidata, büyük dış kaynaklarla uyumu ve çok geniş disiplinliliği nedeniyle COVID-19 bilgisi için bir merkez haline gelmiştir. Bağlantılı Açık Veri biçiminde bilgilerin eklenmesi, COVID-19 verilerinin birçok yönünden gelen verileri COVID-19 dışı verilerle entegre etme esnekliği sağlar. Çok dilli yapısı sayesinde bu girdiler, dünyanın her yerinden farklı geçmişlere sahip insanlara katkıda bulunur ve onlar tarafından yeniden kullanılır (Turki vd., 2021). Dezavantajları, özellikle açık bilgi grafiklerinde COVID-19 bilgilerinin sürdürülebilirliğini sağlamanın zorluğu, çelişen anlamsal verilerin doğrulanmasının zor olduğu, analiz edilen pandeminin

için kaynağı T.C. Sağlık Bakanlığı tarafından yayınlanan TURCOVID19 veri kümelerindeki tüm veri öğelerini anlamsal olarak modellemek amacıyla uygun ontoloji geliştirilmiştir. Açık veri olarak yayınlanan veri kümeleri tablosal biçimlerde (CSV, Excel, PDF vb.) temsil edilir ve dağıtılır. COVID-19 veri setinin çoğunluğu yapılandırılmamış veriler içermektedir. Tez kapsamında yapılan çalışmada ise açık veriler ontoloji ile ilişkilendirilip yapılandırılmış RDF formatına dönüştürülerek COVID-19 veri entegrasyonu, karşılaştırması ve yeniden kullanımı kolaylaşacaktır. Elde edilen RDF COVID-19 veri seti ile mevcut veri kaynakları arasında kurulacak olan bağlantılarla Linked Open Data (LOD) bulutunda yayınlanacaktır. Ülke karşılaştırmalarına yönelik içeriklerde diğer ülkelere göre farklı kavramlar kullanılmaktadır. Böylece Türkiye'ye ait sağlık hizmetleri hakkında, veri setinde bulunmayan bilgileri sağlamak için kullanım senaryoları sağlayacaktır. Ulusal bazlı ve güncel olması ile mevcut bilgi grafiklerine göre COVID-19 bilgilerinin sürdürülebilirliği ve analiz edilen pandeminin çeşitli yönlerinin kapsanması sağlanacaktır.

1.9. Bağlantılı Veri Teknolojileri

Bağlantılı veri teknolojileri, webde bulunan mevcut verilerin bağlantılı olarak yayınlanması amacı ile veri webine dönüşümünü sağlamak için kullanılmaktadır. Geliştirilmiş olan bağlantılı veri teknolojileri bu bölümde incelenmiştir.

1.9.1. RDF Veri Modeline Dönüşüm Araçları

Webde bulunan COVID-19 verileri Excel, PDF vs. gibi biçimlerde yayınlanmıştır. Farklı kaynaklarda ve biçimlerde yayınlanmış verilerin birleştirilmesi ve anlamsal olarak temsil edilmesi gerekmektedir. Bu kapsamda veriler RDF veri modeline dönüştürülerek tek bir çatı altında toplanır. COVID-19 verilerinin bağlantılı veri olarak yayınlanmasında ve bilgisayar uygulamaları için kullanılabilir bir hale getirilmesinde belirli bir veri modeline dayandırılarak RDF verilerinin oluşturulması gerekmektedir. RDF veri modeline dönüşümde geliştirilmiş araçlardan bazıları bu bölümde incelenmiştir.

1.9.1.1. Geometry2RDF

Geometry2RDF, GML veya WKT olarak mevcut olan konumsal bilgiden RDF dosyalarını oluşturmak için GeoLinkedData (.es) girişimi kapsamında 2011 yılında oluşturulan Java tabanlı bir yazılımdır. İlk araç olarak konumsal veriyi RDF üçlülerine dönüştürmek için tasarlanmıştır. Dönüşüm için girdi dosyası olarak konumsal veriler kullanılmaktadır. Konumsal veri formatları, ESRI shapefile, GML ve konumsal veri tabanı yönetim sistemi (Oracle, PostgreSQL, MySQL, vb.) kullanılabilir. Coğrafik içerikler NeoGeo Kelime Hazinesine (NeoGeo Vocabulary) uygun RDF'de yapılandırılmış bilgiler olarak belirtilirken varsayılan referans koordinat sistemi WGS84 kullanılmaktadır (Kara ve Cömert, 2016).

1.9.1.2. shp2GeoSPARQL

shp2GeoSPARQL, ISO19152 (Land Administration Domain Model) ve GeoSPARQL kullanarak kadastro alanında ESRI shapefile formatındaki dosyaları RDF ye dönüştürmek amacıyla geliştirilmiştir (Kara ve Cömert, 2016). RDF verilerinin oluşturulmasında GeoSPARQL desteği sağlanmaktadır. Geometry2RDF aracını temel alır ve eksik yönlerini tamamlamaktadır. (Saavedra vd., 2014).

1.9.1.3. GeomRDF

Datalift platformunun bir modülü olarak geliştirilen GeomRDF, geleneksel CBS formatlarındaki konumsal veri setlerinin RDF'ye dönüşümü için kullanılmaktadır. GeomRDF, koordinat referans sistemleri arasındaki dönüşümleri sağlayarak kaynak veriler için ESRI Shapefile, konumsal VTYS ve GML formatları yanı sıra CSV gibi koordinat bilgilerini içeren metin dosyaları da girdi dosyası olarak desteklenmektedir (URL-47).

1.9.1.4. TripleGeo

TripleGeo, Geometry2RDF den genişletilmiş olup bir ETL (Extract-Transform-Load) aracı olarak GeoKnow projesi kapsamında geliştirilmiştir. Farklı kaynaklarda

bulunan konumsal öznitelikler TripleGeo ile elde edilmektedir. Elde edilen konumsal özniteliklerin RDF'ye dönüşümü gerçekleştirilir ve dönüştürüldükten sonra RDF depolarında saklanabilmektedir. Farklı formatlardaki dosyalardan TripleGeo ile elde edilen coğrafi veriler, RDF veri biçimine dönüştürülmesi esnasında girdi olarak konumsal tabloların yanı sıra ESRI Shapefile dosyalarını da almaktadır (Kara ve Cömert, 2016). Açık kodlu ve java dilinde ücretsiz bir yazılımdır. TripleGeo ile RDF/XML, N3, N-Triples, Turtle formatlarında çıktı verileri elde edilmektedir. GeoSPARQL, Virtuoso RDF ve WGS84 RDF Geoposition kelime hazineleri üçlülerin temsilinde kullanılmaktadır. GeoSPARQL standardını desteklediği için GML ve Simple Feature da yer alan bütün geometri tiplerini desteklemektedir (URL-49).

1.9.1.5. GeoTriples

GeoTriples, konumsal verileri Bağlantılı Açık Konumsal Veri olarak yayınlamak için geliştirilen açık kaynak kodlu ve yarı otomatik bir araçtır. MELODIES ve LEO projeleri kapsamında geliştirilmiş belirli bir kelime hazinesine bağlı olmayan, GeoSPARQL ve stSPARQL gibi kelime hazinelerini kullanmaktadır. Kullanılan girdi verileri için Shapefile (ESRI shapefile), KML, PostgreSQL gibi konumsal veri tabanı dosyaları desteklenmektedir. GeoTriples, Mapping Generator (Eşleştirme Üreticisi), Mapping Processor (Eşleştirme İşlemcisi), stSPARQL/GeoSPARQL Evaluator (stSPARQL/GeoSPARQL Değerlendirici) olarak üç ana bileşene sahiptir. Mapping Generator bileşeni ile veri kaynağını alır ve onu bir RDF çizgesine dönüştüren R2RML eşleştirmesini otomatik yapmaktadır. Eşleştirme sonucu oluşturulan R2RML/RML dosyası kullanılarak Mapping Processor bileşeni ile RDF çizgesi oluşturulur. İlişkisel veri tabanı üzerinde, stSPARQL/GeoSPARQL sorguları, Ontology Based Data Access (Ontoloji Tabanlı Veri Erişimi) modülü ile değerlendirilmektedir.

GeoTriples'in kullanılması ile konumsal veriler bağlantılı veriye dönüştürülür. GeoSPARQL ve stSPARQL gibi kelime hazineleri kullanılmaktadır. R2RML ve RML dilleri verilerin RDF'ye dönüşümünde desteklenmektedir. İlişkisel verilerin RDF'ye dönüşümünde R2RML dili kullanılmaktadır. RML, R2RML dilinin genişletilmesidir. GeoTriples ile ilk aşamada R2RML ya da RML eşleştirilmeleri oluşturulmaktadır. Kullanıcılar ihtiyaçlarına göre oluşturulan eşleştirmeler, konumsal verinin özelliklerini dikkate almak ve GeoSPARQL Kelime Hazinesi ile uyumlu bir RDF çizgesi üretmek için

ihtiyaç duyulan bütün dönüşümleri sağlamak için subject-object-predicate eşleştirmeleri ile zenginleştirilir. Alternatif olarak yazılım, R2RML eşleştirme dosyasının farklı kelime hazineleri kullanılacak şekilde kullanıcı gereksinimlerine izin vermektedir (Kara ve Cömert, 2016). Konumsal RDF depolarında oluşturulan RDF dosyaları sorgulanabilir, saklanabilir ya da diğer veriler ile aralarında bağlantılar oluşturulabilir (Kyzirakos vd., 2018).

GeoTriples, kullanıcıların kaynaklara erişimini sağlayabilmesi için grafik ara yüzü sunmaktadır. Ontolojiler kullanıcılar tarafında yüklenebilmektedir. Ayrıca, RDF/XML, N3, N-Triple, Turtle formatlarında eşleştirmelerin görüntülenmesi sağlanır ve üzerinde değişiklik yapılabilir ya da RDF çıktısı için işleme devam edilebilir (Kara vd., 2020).

1.9.1.6. OpenRefine RDF Extension

OpenRefine, dağınık verilerle çalışmak için ücretsiz ve açık kaynak kodlu ve CSV, TSV, XML, JSON, metin dosyaları gibi farklı formatlarla çalışabilen web uygulamasıdır (URL-50). OpenRefine ile verilerin keşfedilmesi, temizlenmesi dönüştürülmesi ya da veri kümeleri arasında bağlantılar oluşturmak amacıyla kullanılmaktadır (URL-51). OpenRefine RDF Extension ile RDF verilerine dönüşümde, veriler ve öznitelikleri eşleştirilmektedir. Farklı veri kaynakları arasında linkler SPARQL Endpoint ya da RDF dosyaları yardımıyla keşfedilmektedir (Angelis ve Kotis, 2020).

1.9.1.7. Karma - A Data Integration Tool

Karma, kullanıcıların elektronik tablolar, sınırlandırılmış metin dosyaları, veri tabanları, JSON, XML, KML ve Web API'leri dahil olmak üzere farklı veri kaynaklarından verileri hızlı ve kolay bir şekilde ilişkilendirmelerini sağlamaktadır. Kullanıcılar, bir grafik ara yüzü kullanarak seçmiş oldukları ontolojiye göre verileri modelleyerek semantik tanımlarını oluşturmaktadır. Karma, verilerin ilgili ontoloji sınıfları ile olan ilişkilerini yapılan eşleştirmeler ile öğrenmektedir. Daha sonra ilişkilendirmelerde otomatik olarak veri için ilgili ontoloji sınıflarını ve özniteliklerini önerir. Ontoloji ve verilerin ilişkilendirilmesinde sonra kullanıcılar entegre edilen verileri RDF olarak yayınlatabilir ya

da bir veri tabanında saklayabilmektedir (URL-52). Tablo 2’de RDF veri modeline dönüşüm araçlarının karşılaştırılması verilmiştir.

Tablo 2. RDF veri modeline dönüşüm araçlarının karşılaştırılması

| Yazılım Araçları | Tematik Veri Desteği | Doğrudan Eşleştirme | R2RML | RML | Otomatik Eşleştirme | GeoSPARQL Uyumu | RDBMS | ESRIShapefile dosyası | *.csv, *.xlsx |
|------------------|----------------------|---------------------|-------|-----|---------------------|-----------------|-------|-----------------------|---------------|
| Geometry2RDF | - | ✓ | - | - | ✓ | - | ✓ | ✓ | - |
| shp2GeoSPARQL | - | - | - | - | ✓ | ✓ | - | ✓ | - |
| GeomRDF | ✓ | ✓ | - | - | ✓ | ✓ | ✓ | ✓ | - |
| TripleGeo | ✓* | ✓ | - | - | - | ✓ | ✓ | ✓ | - |
| GeoTriples | ✓ | - | ✓ | ✓ | ✓ | ✓ | ✓ | ✓ | - |
| Karma | ✓ | ✓ | - | - | ✓ | ✓ | ✓ | - | ✓ |
| OpenRefine | ✓ | ✓ | - | - | - | ✓ | - | - | ✓ |

1.9.2. SPARQL Endpoint

Webde RDF dokümanları ya da SPARQL Endpoint olarak yayınlanan bağlantılı verilerin sorgulanması için SPARQL sorgu dili kullanılmaktadır. SPARQL Endpoint’lerin Semantik Web’e ait tüm dokümanları olmasa da büyük miktarda verilere veri kümeleri üzerinde sorgulamalar yapılabilmektedir. Bazı SPARQL Endpoint’ler bu bölümde incelenmiştir.

1.9.2.1. OpenLink Virtuoso

OpenLink Virtuoso Endpoint, SPARQL protokolünü uygulayarak RDF biçiminde webde verileri sorgulamaktadır. HTTP ile SPARQL gibi sorgu dillerini kullanılarak sorgulama yapılır. Kullanıcılar tarafından en çok temsil edilen bu Endpoint, veri tanımlama ve işleme odaklı işlemler için sistemlerden gelen verilerin bilgi grafiklerine dönüştürülmesini sağlar (URL-53).

1.9.2.2. Fuseki

Fuseki, bir java web uygulaması ve Apache Jena'ya ait SPARQL sunucusudur (URL-54). Dosya yükleme ve RDF veri setindeki tüm grafikleri okumak ve güncellemek gibi çeşitli hizmetler sunulmaktadır. SPARQL sorgusu ve güncellemesi SPARQL protokolü ile sağlamaktadır (URL-55).

1.9.2.3. D2R Server

D2R Server, Semantik Web'de ilişkisel veritabanları yayınlamak için bir araçtır veriler RDF'de modellenir ve temsil edilir. D2R Server, veri tabanı içeriğini bu formatta eşleştirmek için özelleştirilebilir bir D2RQ eşleşmiş veri tabanı kullanır ve SPARQL sorgu dilini kullanarak veri tabanını sorgulamaya izin verir. XML ya da JSON biçimlerinde sorgu sonuçları elde edilmektedir (URL-56).

1.9.2.4. 4store SPARQL Server

4store, ölçeklenebilir ve kararlı bir RDF veritabanıdır. Açık kaynak kodlu geliştirilmiştir. SPARQL HTTP sorgu protokolü kullanılmaktadır (URL-57).

1.9.2.5. PublishMyData

PublishMyData, verileri webde bağlantılı veri olarak yayınlamak için oluşturulmuş kurulum gerektirmeden doğrudan erişim sağlanan platformdur. Bu platform ile veri yayıncıları kendi veri kümelerini oluşturup yönetimini gerçekleştirebilir. Kullanıcılar SPARQL Endpoint ara yüzü ile verilere erişebilir, sorgulayabilir ve bu verileri indirebilmektedir (URL-58).

1.9.3. Eşleştirme Araçları

Bağlantılı Açık Veri oluşturmak için en iyi uygulamalardan biri, RDF formatında oluşturulan veri setini webde bulunan farklı veri setlerine bağlamaktır. RDF COVID-19 veri setini zengileştirmek için diğer ilgili veri setleri ile arasında bağlantılar oluşturulur. Verilerin tanımlanmasında farklı kelime hazinelerinin kullanılması ve yoğun veri akışının olması sebebi ile veri setlerinin arasında benzer ve eşleşen bağlantı türlerini belirlemekte zorluklar ortaya çıkmaktadır. Farklı terminolojiler kullanılmasından dolayı seçilen veri kümelerinden gelen veriler, ülkelerin adları eşleştirilerek bağlanması mümkün olabilmektedir. Örneğin, `dbpprop:countryName owl:sameAs skos:preflabel, dbpprop:countryName owl:sameAs madsrdf:authoritativeLabel` vb. linkler ile RDF COVID-19 veri kümesi diğer farklı veri kaynaklarıyla yayınlanabilir (Ali vd, 2022).

Bağlantılı verilerin Web'de farklı veri setleri ile bağlanması amacıyla kullanılan araçlar bu bölümde incelenmiştir.

1.9.3.1. Silk-The Linked Data Integration Framework

Silk, heterojen veri kaynaklarını entegre etmek için açık kaynaklı bir çerçevedir. Silk, iki basit fikir üzerine kurulu Bağlantılı Veri paradigmasına dayanmaktadır. Birincisi, RDF, yapılandırılmış bilgiyi temsil etmek için anlamlı bir veri modeli sağlar. İkinci olarak, farklı veri kaynaklarındaki varlıklar arasında RDF bağlantıları ayarlanır. Silk'in birincil kullanım durumları şunları içerir:

- Farklı Bağlantılı Veri kaynakları içindeki ilgili veri öğeleri arasında bağlantılar oluşturur.
- Bağlantılı Veri yayıncıları, veri kaynaklarından Web'deki diğer veri kaynaklarına RDF bağlantıları ayarlamak için Silk'i kullanabilir.

Veri dönüşümlerini yapılandırılmış veri kaynaklarına uygular (URL-59).

1.9.3.2. LIMES (Link Discovery Framework For Metric Spaces)

LIMES, Web of Data için bir bağlantı bulma çerçevesidir. Metrik uzayların özelliklerine dayalı büyük ölçekli bağlantı keşfi için zaman açısından verimli yaklaşımlar

uygular. Bir konfigürasyon dosyası ve ayrıca grafiksel bir kullanıcı arayüzü aracılığıyla kolayca konfigüre edilebilir. LİMES, bağlantı keşfi gerçekleştirmek için bağımsız bir araç olarak veya bir Java kitaplığı olarak indirilebilir (URL-60).

1.9.3.3. The xCurator Project

xCurator, girdi kaynaklarının hem yapısından hem de verilerinden yararlanarak yarı yapılandırılmış verileri bağlantılı verilere dönüştürür. Yarı yapılandırılmış veriler Web'de bol miktarda bulunur. Birçok Web veri kaynağı ve API, verileri Web uygulaması geliştiricileri tarafından kolayca erişilebilir ve kullanılabilir hale getirmek amacıyla verilerini XML, JSON veya alana özgü yarı yapılandırılmış bir biçimde sunar. Bu tür veri biçimleri, salt metin belgelerinden daha fazla makine tarafından işlenebilir olsa da, bu tür verilerin büyük ölçekte yönetilmesi ve analiz edilmesi zordur. Bahsedilen zorluk temel olarak, veri formatlarında iyi tanımlanmış bir yapı ve semantik eksikliğinden kaynaklanmaktadır. Bu da zaman içinde verilerin kalitesinin düşmesine neden olabilir.

xCurator projesi ile varlıkların türleri ve ilişkileri ayıklanır. Yinelenen varlıklar tanımlanarak birleştirilir. İlgili varlıklar birbirine bağlanarak ve sonuçları Web'de yayınlanır. Böylece, bu tür verilerin yapısı eklenir ve kalitesi artırılmış olur (URL-61).

1.9.4. Bağlantılı Verinin Yayınlanması

Uygun ontoloji ile yapılandırılmış ve RDF veri modeline dönüştürülmüş olan veri, mevcut bağlantılı veri setleri ile eşleştirildikten sonra bağlantılı veri olarak yayınlanmaktadır. Bağlantılı verilerin yayınlanmasında kullanılan araçlar bu bölümde incelenmiştir.

1.9.4.1. Pubby

Pubby mevcut SPARQL özellikli üçlü mağazalara Bağlantılı Veri arayüzü ekleyen bir RDF sunucusudur. Açık kaynaklı olarak Java tarafından geliştirilmiş bir web uygulamasıdır. Bu araç ile kaynağa ait bilgiler, ilişkiler görüntülenebilir ya da meta

verilerde deęişiklik yapabilir. Kullanıcılar, web arayüzü aracılığıyla bağlantılı verilerin HTML görünümü ile kaynaklar arasında gezinebilir (URL-62).

1.9.4.2. CKAN

CKAN, 2006 yılında açık verilerin depolanması ve paylaşılması için oluşturulmuş ücretsiz açık veri platformudur. CKAN ara yüz ile kullanıcıya çok yönlü arama özellięi imkânı sunmakta ve verileri yayınlamayı, paylaşmayı ve verilerle çalışmayı kolaylaştırmaktadır. Kullanıcılar verileri görüntülemek için harita, grafik ve tabloları kullanabilmektedirler (URL-63).

1.9.4.3. DataHub

DataHub, Datopian ve Open Knowledge International tarafından başlatılmış kurumlara ait verilerin yayınlanması, kaydedilmesi ve paylaşılmasını sağlayan bir platformdur (URL-64).

1.9.4.4. LOD Cloud

LOD Cloud veri kümelerinin bağlantılı veri ilkelerine uygun biçimde LOD Cloud diyagramında yayınlanan veri kümesine RDF linkleri ile bağlanarak yayınlandığı platformdur. 2007 yılında yalnızca bir veri kümesine sahip biçimde oluşturulan LOD Cloud Mayıs 2020 itibarıyla, 16.174 bağlantıya sahip 1.255 veri kümesi içermektedir. Tüm veri kümesine erişim RDF taraması, RDF dökümü veya bir SPARQL Endpoint aracılığıyla mümkün olmaktadır (URL-65).

1.9.5. Konumsal Üçlü Depoları

RDF Üçlü depoları ya da Triple Store, bir sorgu dili aracılığıyla sorgulanabilen RDF verilerinin depolandığı veri tabanlarıdır. Hâlihazırda bulunan bazı konumsal üçlü depolar bu bölümde incelenmiştir.

1.9.5.1. Parliament

Parliament, Haziran 2009'da Semantik Web için tasarlanmış açık kodlu yüksek performanslı bir triple deposudur. W3C'nin RDF, OWL ve SPARQL standartlarıyla uyumludur. Parliament kural motoru tüm RDFS çıkarımını ve OWL RL'nin seçili öğelerini uygulamaktadır ve veriler arasındaki ilişkileri sahip olduğu kural motoru sayesinde anlamlandırarak sorgu sonuçlarını güçlendirebilmektedir. Parliament depolama katmanı C++ ile yazılmıştır, geri kalanı ise Java kodu kullanılarak oluşturulmuştur. Parliament'in amacı, özellikle Semantik Web'in ihtiyaçlarına göre optimize edilmiş bir depolama mekanizmasının oluşturulmasıdır. Parliament üçlü bir depo ve bir kural motorudur, ancak bir sorgu işleyicisi içermemektedir. Bu nedenle, Jena veya Sesame gibi bir üçüncü taraf sorgu işlemcisi ile eşleştirilmektedir (Emmons, 2018). R-Tree uygulamasına dayanan bir konumsal indeks kullanılır. R-tree indeksi konumsal veri tabanlarında kullanılan indeks türüdür. Bu indeks ile uzamsal bilgiler içeren SPARQL sorguları birden fazla parçaya bölünür. Sorgunun uzamsal bileşenleri çok parçalı olarak bölünürken, uzamsal olmayan bileşenleri arasında en uygun olan sorgulama yapılmaktadır. Karmaşık coğrafi sorguların yapılması GeoSPARQL standardına dayandırılmıştır (Battle ve Kolas, 2012).

1.9.5.2. Strabon

Strabon, RDF'de ifade edilen bağlantılı jeo-uzamsal verileri depolayan veritabanı sistemidir. SPARQL uzantısı kullanarak bu verileri sorgulamaktadır. Strabon, farklı Koordinat Referans Sistemi ve WKT ve GML gibi geometrik nesnelere için yaygın olarak kullanılan serileştirmeleri desteklemektedir. Strabon Sesame genişletilerek oluşturulmuştur.

Strabon'da stSPARQL sorgu dili ve GeoSPARQL'in alt kümesi kullanılarak sorgular gerçekleştirilmektedir. Strabon'a ait sorgu diline stSPARQL denir. stSPARQL, stRDF adı verilen bir RDF uzantısında temsil edilen verileri sorgulamak için kullanılabilir. stRDF ve stSPARQL, zaman içinde değişen coğrafi verileri temsil etmek ve sorgulamak için tasarlanmıştır (URL-66).

1.9.5.3. Fuseki

Fuseki, SPARQL 1.1 ile entegre RDF verisi için bir depolama ve sorgulama aracıdır. HTTP üzerinden SPARQL sorgularına yanıt sağlar ve veri setinin SPARQL Endpoint üzerinden yayını yapılabilir (URL-54).

1.9.6. Bağlantılı Verinin Doğrulaması

Bağlantılı Veri yayınlamada doğrulama işlemi yapılmaktadır. Fakat yapılan işlemlerin doğruluğu URI'ler aracılığı ile söz dizim hataları ve erişilebilirlik gibi konularda test edilmelidir. Geliştirilmiş bazı araçlar bu bölümde incelenmiştir.

1.9.6.1. RDFSshape-RDF Playground

RDFSshape- RDF playground, ShEx (Shape Expressions) ve SHACL (Shapes Constraint Language) dillerini kullanarak RDF verilerini doğrulamak için tasarlanmıştır. Kullanıcılar RDF graphlarında SPARQL sorguları ile temel düzeyde çıkarımlar gerçekleştirebilir (URL-67).

1.9.6.2. Vapour

Vapour, W3C standartına dayalı Python dilinde yazılmış açık kodlu bir web tabanlı uygulamadır. Yayımlanan Semantic Web verisinin Bağlantılı Veri İlkelerine uygunluğunu denetlemektedir (URL-68).

1.9.6.3. W3 RDF Validator

W3 RDF Validator, Shape Expressions (ShEx) dilinde yazılmış doğrulama yaparak geçerli RDF belgeleri oluşturmaya yardımcı olan web tabanlı bir araçtır. Sorgulama doğrudan URI ya da doküman ile yapılabilir. Sonuçlar farklı formatlarda biçimlendirilebilmektedir (URL-69).

2. YAPILAN ÇALIŞMALAR

Tez çalışması kapsamında, TURCOVID19 verilerinin semantik tanımları oluşturulmuştur. Semantik tanımların oluşturulması için en önemli gereksinim mevcut ontolojilerin bulunması ve ontoloji geliştirmektir. Bu bağlamda ilgili literatürdeki mevcut ontolojiler incelenmiştir. Mevcut ontolojiler değerlendirilerek TURCOVID19 verilerinin semantik tanımlarında kullanılacak ontolojiler belirlenmiştir. Ancak mevcut ontolojiler, ülkemizdeki COVID19 açık verilerinin semantik tanımları için yeterli değildir ve bu nedenle TURCOVID19 Ontolojisi geliştirilmiştir. TURCOVID19 ontolojisinin geliştirilmesi için gerekli ontolojier Bölüm 2.2’de verilmiştir. TURCOVID19 Ontolojisi’nin geliştirilmesinden sonra TURCOVID19 açık verileri bu ontoloji ile ilişkilendirilmiştir. Veriler ve TURCOVID19 Ontolojisi arasında ilişkilerin belirlenmesi için ya da verilerin semantik tanımları ile referanslandırılması için KARMA 2.2 kullanılmıştır.

COVID19 açık verilerinin semantik tanımlarının oluşturulması, bu verilerin Semantik Web uygulamalarında kullanılabilirliği ve SWT ile bulunabilirliği sağlanmıştır. Ayrıca semantik tanımlar, mevcut sınıflar ve ilişkiler kullanılarak yeni bilgilerin çıkarılması hedeflenmiştir. Tez çalışmasının ana hedefleri doğrultusunda hazırlanan araştırma sorularına yönelik sorgular belirlenmiştir ve SPARQL dili kullanılarak ifade edilmiştir. Semantik tanımlar kullanılarak sorgulamaların yapılması için Fuseki kullanılmıştır.

Bütün kategorilerde yer alan COVID-19 açık verilerinin semantik tanımları RDF olarak yayınlanmıştır. RDF olarak yayınlanan COVID19 açık verileri ve oluşturulan TURCOVID19 Ontolojisi, Silk Link Discovery Framework kullanılarak LOD Cloud’da yer alan mevcut COVID-19 veri setleri arasında bağlantılar kurulmuştur. Bağlantıların kurulması aşamasından sonra TURCOVID19 verileri LOD Cloud üzerinden bağlantılı veri olarak yayınlanmıştır.

2.1. TURCOVID19 Verilerinin Elde Edilmesi

Sağlık hizmetlerinde birey düzeyinde hastaların monitörize edilmesi sürecin yönetimini mümkün kılmaktadır. Halk sağlığı sorunlarında da epidemiyolojik verilerin monitörizasyonunu sağlamak toplum sağlığını yönetebilmek için büyük önem taşımaktadır. Bu öneme istinaden Google bulut bilişim hizmetleri kullanılarak pandeminin toplumda hızlı ilerleyişini gerçek zamanlı izlemek ve değerlendirmek için Türkiye COVID-19 Pandemi İzlem Ekranı (TURCOVID19) adıyla interaktif bir platform oluşturulmuştur (Ucar vd., 2020).

21 Mart 2020’de yayına giren TURCOVID19’da Türkiye’deki hasta, vefat, iyileşen, uygulanan test sayıları, vaka ölümlülük oranı (VÖÖ), yoğun bakım yatışı, entübe hasta kategorilerindeki veriler günlük ve kümülatif değerler olarak görselleştirilmekte ve tablo halinde sunulmaktadır. Veri kaynağı olarak T.C. Sağlık Bakanlığı’nın açıkladığı günlük tablo ve PDF dosyası olarak yayınlanan durum raporlarından alınan veriler ekran denetiminden geçirilerek çalışma ekibi tarafından veritabanına manuel olarak girilmektedir. Veriler günde bir kez güncellenmekte olup verilerin son güncellenme tarihi ilgili grafikler üzerinde sunulmaktadır (Ucar vd., 2020). Grafikler ve bilgiler ücretsiz olarak TURCOVID19 sayfasında yayınlanarak yapılacak olan salgın modelleme çalışmalarına katkı sunması amaçlandığı için veriler .XLSX ve .CSV formatlarında indirilerek tez kapsamında kullanılmasına karar verilmiştir.

30 Eylül 2020 tarihinde yapılan Sağlık Bakanlığı’nın açıklamasında, “vaka sayısı” tanımının sadece semptomatik vakaları, yani “hasta sayısı”nı ifade ettiği, asemptomatik vakaları kapsamadığı ifade edilmiştir. Bu tanıma istinaden 2 Ekim 2020 itibariyle TURCOVID19 içeriklerinde Türkiye verilerine yönelik “vaka” ifadesi “hasta” ifadesi ile değiştirilmiştir. 30 Eylül 2020 tarihli resmi açıklamada ise tabloda sunulan “günlük vaka sayısı”nın, esasında “günlük hasta sayısı” olduğu ifade edilmiştir. Bunun üzerine vaka sayısının açıklanması yönünde kamuoyunda oluşan talep oluşmuştur. Bu nedenle, 25 Kasım 2020 tarihinden itibaren de günlük vaka sayıları açıklanmaya başlanmıştır. 25 Kasım 2020 tarihinden önceki döneme ait vaka sayıları ise bir süre açıklanmamıştır. 10 Aralık 2020 tarihinde tablodaki “toplam hasta” başlığı “toplam vaka” olarak değiştirilmiştir. 9 Aralık’ta 558.517 olarak sunulan rakam 10 Aralık’ta 1.748.567 olarak sunulmuştur. Meydana gelen bu farkın ne olduğuna dair resmi ve net bir açıklama yapılmamıştır. Bir günde 1.190.050 hasta veya vaka olamayacağı için bu büyük fark, 25

Kasım öncesinde açıklanmamış olan vaka sayılarının toplamı olduğu düşünülmektedir. Bu ekleme baz alınarak toplam vaka sayısına günlük vaka sayılarını ekleyerek tablo güncellenmektedir (URL-70).

TURCOVID19 sayfasında yayınlanan bazı veriler, Türkiye İstatistik Bölge Birimleri Sınıflandırması (İBBS)'na ve cinsiyet ve yaş gruplarına göre verilmektedir. Türkiye İBBS, Avrupa Birliği ülkelerinin Türkiye İstatistik Bölge Birimleri Sınıflandırması içinde Türkiye için kullanılan sınıflandırmadır. İBBS, Avrupa'da 1970'lerde başlayan bir coğrafi kodlama sistemidir. Amacı, bölgesel olarak istatistikleri toplamak, sosyo-ekonomik analizler yapmak ve topluma yönelik bölgesel politikaların çerçevesini oluşturmaktır (URL-71). Şekil 23'te İBBS-1 bölgeleri gösterilmiştir.



Şekil 23. Türkiye İBBS-1 bölgeleri (URL-72)

2.2. TURCOVID19 Verilerinin Semantik Tanımları İçin Ontolojilerin Değerlendirilmesi

2.2.1. CODO

CODO'nun geliştirilmesi ile üretilen verilerin semantik olarak sorgulanıp geri alınabilmesi ve ayrıca girdi olarak alınabilmesi için COVID-19 vaka verilerinin günlük organizasyonunun desteklenmesi amaçlanmıştır. CODO ayrıca hasta verilerinin temsili, hastalar arasındaki ilişkileri, hasta ve lokasyon arasındaki ilişkileri, zaman içindeki değişiklikleri vb. kolaylaştırmayı amaçlamaktadır. Bu ağ verileri, hastalığın davranış

Tablo 3. TURCOVID19 verilerinin CODO'daki karşılığı

| TURCOVID19 | | CODO |
|--|---------------------|-----------------------|
| Veri Setleri | Öznitelik | Data Property |
| COVID-19 Pandemisi Türkiye Günlük Verileri | Tarih | covid-19 case on |
| | Toplam Vaka | cases |
| | Günlük Vaka | daily increased cases |
| | Toplam Hasta | - |
| | Günlük Hasta | - |
| | Toplam Vefat | deceased |
| | Günlük Vefat | - |
| | Toplam İyileşen | recovered |
| | Günlük İyileşen | - |
| | Toplam Test | - |
| | Günlük Test | - |
| | YBU | - |
| | Entube | - |
| | Ağır Hasta | - |
| | Yatak Doluluk Oranı | - |
| Erişkin Yoğun Bakım Doluluk Oranı | - | |

2.2.2. ROC Ontology

Mevcut ontolojiler, COVID-19 vakalarıyla ilgili kavramsallaştırmaları ve kısmen örnek verileri içermektedir. Mevcut ontolojilerin öncelikleri soru-cevap veya gösterge tablolarını izleme gibi yazılım uygulamalarına yöneliktir. Ancak ROC ontolojisinin temel amacı, COVID-19 pandemisinde uygulanan politikaların ulusal bazlı yanıtlarını dikkate alan ve bunları vaka verileri gibi epidemiyolojik verilerle ilişkilendirmektir. Heterojen veri kaynaklarından gelen verilerin entegrasyonunu sağlamak ve COVID-19'a karşı ülkelerin almış olduğu tedbirlerin etkililiğini ve yan etkilerini araştırmak ve değerlendirmek için istatistiksel analizi kolaylaştırması hedeflenmiştir. ROC ontolojisinin merkezi kavramları, Oxford Covid-19 Government Response Tracker (OxCGRT) tarafından tanımlanan göstergelerdir. Bu göstergelerin her biri için değerler elde edilmekte ve bir veri kaydında saklanmaktadır. ROC ontolojisinde bir kayıt, CODO sınıfı CountryWiseStatistics'in bir alt sınıfı olan ResponseStatistics sınıfının bir örneği olarak temsil edilmektedir. Şekil 26'da gösterilen OxCGRT kodlama kategorilerinin (kapsama ve kapatma (C), ekonomik

müdahale (E), sağlık sistemleri (H) ve çeşitli (M)) sınıflandırmasını yansıtan bir veri özelliği hiyerarşisi kurulmuştur (Qundus vd., 2021).

| Property Assertion | Value | Control Buttons |
|--|------------------------|-----------------|
| roc:ch4_emergency_investment_in_healthcare | 0 | [+][x][o] |
| roc:c1_school_closing | 3 | [+][x][o] |
| roc:h1_public_information_campaigns | 2 | [+][x][o] |
| roc:c6_stay_at_home_requirements | 0 | [+][x][o] |
| roc:c4_restrictions_on_gatherings | 4 | [+][x][o] |
| roc:h2_testing_policy | 3 | [+][x][o] |
| roc:h5_investment_in_vaccines | 0 | [+][x][o] |
| roc:c2_workplace_closing | 2 | [+][x][o] |
| roc:c3_cancel_public_events | 2 | [+][x][o] |
| roc:c5_close_public_transport | 0 | [+][x][o] |
| roc:c8_international_travel_controls | 3 | [+][x][o] |
| roc:h6_facial_coverings | 2 | [+][x][o] |
| roc:h3_contact_tracing | 2 | [+][x][o] |
| roc:e3_fiscal_measures | 0 | [+][x][o] |
| roc:e1_income_support | 2 | [+][x][o] |
| roc:c7_restrictions_on_internal_movement | 2 | [+][x][o] |
| roc:e2_debt_contract_relief | 1 | [+][x][o] |
| roc:e4_international_support | 0 | [+][x][o] |
| roc:responseActiveOnDate | "2020-06-23"^^xsd:date | [+][x][o] |

Şekil 25. Alınan yanıtta sahip istatistiklerin değerleri örneği (Qundus vd., 2021).



Şekil 26. OXCGRT yanıtını temsil eden veri özelliklerinin hiyerarşisi (Qundus vd., 2021).

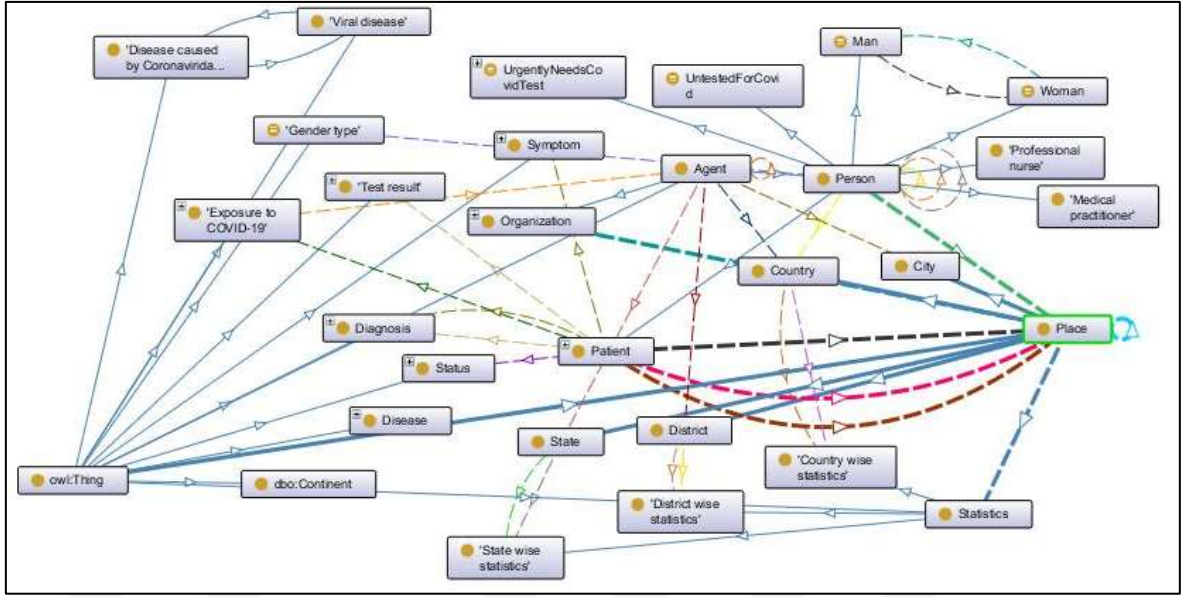
Ontoloji 27 OWL sınıfı, 10 nesne özelliği, 42 veri özelliği ve 3 açıklama özelliğinden oluşur. Sınıflar ve aralarındaki ilişkiler Şekil 27’de verilmiştir.

Tablo 4. TURCOVID19 verilerinin ROC ontolojisi'ndeki karşılığı

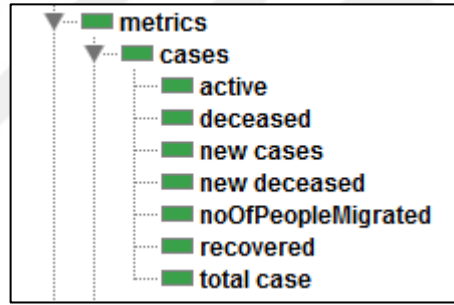
| TURCOVID19 | | ROC |
|--|-------------------------------------|-----------------------|
| Veri Setleri | Öznitelik | Datatype property |
| Türkiye İl Düzeyi Haftalık Vaka Sayıları | Tarih başlangıç | startDate |
| | Tarih bitiş | endDate |
| | il | - |
| | 100 bin kişide haftalık vaka sayısı | - |
| COVID-19 Pandemisi Türkiye Günlük Verileri | Tarih | 'covid-19 case on' |
| | Toplam Vaka | cases |
| | Günlük Vaka | daily increased cases |
| | Toplam Hasta | - |
| | Gunluk Hasta | - |
| | Toplam Vefat | deceased |
| | Gunluk Vefat | - |
| | Toplam İyileşen | recovered |
| | Gunluk İyileşen | - |
| | Toplam Test | - |
| | Gunluk Test | - |
| | YBU | - |
| | Entube | - |
| | Ağır Hasta | - |
| | Yatak Doluluk Oranı | - |
| | Erişkin Yoğun Bakım Doluluk Oranı | - |
| Ventilator Doluluk Oranı | - | |

2.2.3. WHOnto

WHOnto, COVID-19 verilerini Web'de Grafik olarak yayınlamak için anlamsal bir model oluşturulmuştur. Ontolojide COVID-19 vakalarını temsil eden bir bilgi alanı oluşturulmuştur. Bu alan içerisinde Şekil 28'de gösterildiği gibi COVID-19 vakalarına ait, yeni vakalar (new cases), toplam aktif (active), iyileşen (recovered), ölen (deceased), bir günde yeni ölen (new deceased) olarak bilgiler yer almaktadır. Ontoloji, vakalara ait bilgilerin coğrafi konum (bölge, eyalet, ülke ve kıta) içindeki istatistiklerini günlük olarak desteklenmesine odaklanmıştır (URL-7). TURCOVID19 ontolojisinde ise WHOnto'dan farklı olarak hasta, vefat, iyileşen, uygulanan test sayıları, yoğun bakım yatışı, entübe hasta, aşı, hastane yatışları, taburcu edilenler kategorilerindeki verilerin coğrafi konuma göre istatistikleri ve cinsiyet-yaş grubuna göre de ölüm ve hasta sayıları istatistikleri olarak günlük ve kümülatif şekilde desteklenmesi amaçlanmıştır.



Şekil 28. WHOnto Ontolojisi (URL-7)



Şekil 29. WHOnto veri özellikleri hiyerarşisi (URL-7)

TURCOVID19 verilerine ait tarih, toplam vaka, günlük vaka, toplam vefat, günlük vefat özniteliklerinin ontolojideki karşılığı covid case on, new cases, deceased ve new deceased olarak tespit edilmiştir ve Tablo 5.'de gösterilmiştir.

Tablo 5. TURCOVID19 verilerinin WHOnto'daki karşılığı

| TURCOVID19 | | WHOnto |
|---|-----------------------------------|-------------------|
| Veri Setleri | Öznitelik | Datatype property |
| COVID-19 Pandemisi Türkiye Günlük Verileri | Tarih | covid case on |
| | Toplam Vaka | total case |
| | Günlük Vaka | new cases |
| | Toplam Hasta | - |
| | Günlük Hasta | - |
| | Toplam Vefat | deceased |
| | Günlük Vefat | new deceased |
| | Toplam İyileşen | - |
| | Günlük İyileşen | - |
| | Toplam Test | - |
| | Günlük Test | - |
| | YBU | - |
| | Entube | - |
| | Ağır Hasta | - |
| | Yatak Doluluk Oranı | - |
| | Erişkin Yoğun Bakım Doluluk Oranı | - |
| | Ventilator Doluluk Oranı | - |

ROC ve WHOnto ontolojilerinde göstergeler tarafından tanımlanan kavramlar vakalara ait verilerle ilişkilendirilmiştir. Ontolojilerin geliştirme sürecine bakıldığında CODO'nun ilham kaynağı olduğu ve kullanıldığı tespit edilmiştir. TURCOVID19 ontolojisinin geliştirilmesi ve TURCOVID19 platformu tarafından yayınlanan verilere ait özniteliklerin modellenmesinde en uygun ontolojinin CODO olduğuna karar verilmiştir.

2.3. TURCOVID19 Verileri İçin Kullanılacak Ontolojilerin Belirlenmesi

Mevcut ontolojiler değerlendirildiğinde TURCOVID19 verilerinin semantik tanımlarının temsili için geliştirilebilir ontolojinin CODO olduğuna karar verilmiştir.

CODO, COVID-19 salgınına ilişkin veri toplama için standartlara dayalı ve kapsamlı bir açık kaynak veri modeli sağlamaktadır. Ontoloji, COVID-19 verilerinin bilgi grafiği olarak yayınlanmasını kolaylaştırması sonucunda devlet kurumları, hastaneler, veri yayıncıları vb. çeşitli kuruluşların COVID-19 hastalığı ile ilgili bilgilere açıklama eklenmesini sağlayacaktır. Böylece mevcut pandemi ile başa çıkmaya yardımcı ve

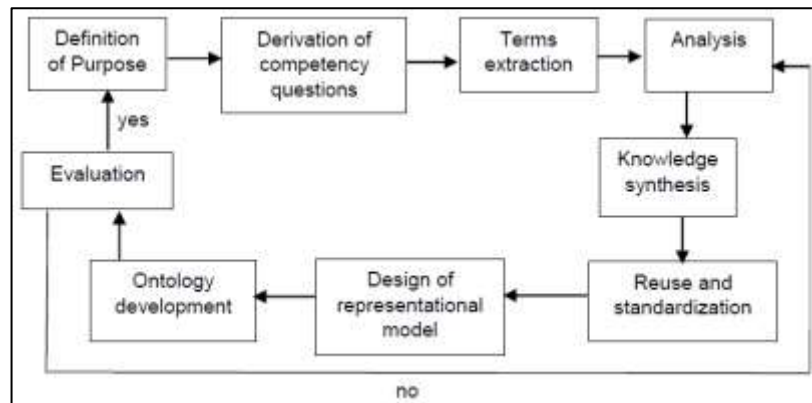
gelecekteki olası krizlere hazırlanmak için bir araç olması ile çalışma için en önemli ilham kaynaklarından birini temsil ettiği tespit edilmiştir.

COVID-19 verilerinin konumsal ilişkilerinin semantik olarak tanımlanması için GeoSPARQL Ontolojisi kullanılmıştır.

2.4. TURCOVID19 Ontolojisinin Geliştirilmesi

Ontoloji geliştirme süreci için, bir yandan mevcut verilerle, diğer yandan ontoloji tarafından cevaplanabilecek ilgi çekici sorularla yönlendirilen bir yaklaşım izlenmiştir. Ontolojiyi tasarlamak için Protégé kullanılmıştır. CODO ontolojisi import edilerek TURCOVID19 ontolojisi geliştirilmiştir.

Bir ontoloji geliştirmek için literatürde birçok metodoloji mevcuttur. METHONTOLOGY (Fernandez vd., 1997), TOVE (Gruninger ve Fox, 1995), DILIGENT (Vrandecic ve diğerleri, 2005), NeOn (Suárez-Figueroa vd., 2012), UPON (De Nicola vd., 2005), YAMO (Dutta vd., 2015), vb. modern popüler yaklaşımlardan bazılarıdır. CODO'nun tasarımında öncelikle, resmi olarak tanımlanmış adım adım bir yaklaşım olan YAMO metodolojisi baz alınmıştır. YAMO metodolojisi ayrıca oldukça benzersiz olan bir dizi ontoloji geliştirme kılavuz ilkeleri sağlar. CODO ontolojisi tasarım sürecinin adımları Şekil 30'da gösterilmiştir (Dutta ve DeBellis, 2020). TURCOVID19 ontolojisi tasarım sürecinde de bu adımlar izlenmiştir.



Şekil 30. CODO Ontolojisi geliştirme sürecinin adımları (Dutta ve DeBellis, 2020).

2.4.1. Amacın Tanımı (Definition of Purpose)

Kaynağı T.C. Sağlık Bakanlığı olan TURCOVID-19 platformu tarafından yayınlanan COVID-19 veri setleri mevcut kaynaklardan ayrıntı dereceleri bakımından farklılık göstermektedir. Geliştirilen TURCOVID19 ontolojisi ile verilerin semantik tanımlarının oluşturulması ve veri setlerindeki veri öğelerinin anlamsal olarak modellenmesi, COVID-19 alanına ait bilgilerin bilgisayarlar tarafından anlaşılabilir hale getirilmesini sağlayacaktır. Ayrıca, COVID-19 verilerinin bir bilgi grafiği (knowledge graph) olarak yayınlanması ve birden çok kaynaktan veri setleri arasında sorguların yapılması kolaylaşacaktır.

2.4.2. Yetkinlik Sorularının Tanımı (Derivation of Competency Questions)

TURCOVID19 ontolojisinin tanımlanan amacına göre yetkinlik soruları oluşturulmuştur. Yeterlilik sorularından bazıları şunlardır:

- T tarihinde P yerinde kaç tane COVID-19 vakası vardır?
- T tarihinde P yerinde kaç COVID-19 yeni vakası bulunmuştur?
- T tarihine kadar P yerinde COVID-19 hastalarından kaç kişi iyileşmiştir?
- C ülkesinde kaç kişi ölmüştür?
- T tarihinde kaç kişi ölmüştür?
- Hangi yerde maksimum vaka sayısı vardır?
- T tarihinde illere ait uygulanan aşı sayıları ne kadardır?
- Cinsiyete göre hangi yaş grubunda ölüm hızı daha yüksektir?
- COVID-19 vakalarının İstatistikî Bölge Birimleri Sınıflandırması (İBBS) içindeki konumsal ilişkileri nasıldır?

Ontolojinin amacına uygun şekilde belirlenen yetkinlik soruları ile RDF formatına dönüştürülmüş mevcut veriler SPARQL sorgu dili kullanılarak verilere tek tip formatta erişimi ve sorgulamayı sağlayacaktır. Ayrıca, mevcut verilerden yeni bilgiler çıkarılacaktır.

2.4.3. Kavramların Çıkarılması (Term Extraction)

Ontoloji kavramlarını çıkarırken amaca uygun olarak TURCOVID19 veri deposundaki veri setlerine atıfta bulunulmuştur. Veri kaynağı olarak T.C. Sağlık Bakanlığı tarafından yayınlanan günlük tablo, durum raporları ve COVID-19 Aşısı Bilgilendirme Platformu kullanılmaktadır. TURCOVID19 ontolojisinde, TURCOVID19 tarafından yayınlanan veri setlerindeki tablo isimleri sınıf (class) ve sınıflar arasındaki ilişkiler nesne özelliği (object property) olacak şekilde kavramlar çıkarılmıştır. Ayrıca, verilere ait öznitelikler ise veri özellikleri (data properties) olarak çıkarılmıştır.

2.4.4. Analiz (Analysis)

Kavramların çıkarılmasının ardından, analiz adımında türetilmiş bileşik ve karmaşık kavramları analiz etmeyi ve bunları temel varlıklarına ayırmayı içermektedir. Analiz, kavramların her birinin tanımına ve özelliğine göre yapılır ve ardından benzerliklerine göre gruplandırılır. Kavram (concept) kelimesi bazen sınıf yerine kullanılır. Sınıflar kavramların somut bir temsilidir (a concrete representation of concepts).

Ontolojide, Patient semptomatik vakaları temsil eden sınıfı ifade etmektedir. Semptomatik vakalar olarak tanımlanan hastaların cinsiyetlerine göre yaş grupları oluşturulmuştur. İBBS bölgeleri, düzey bölgelerine göre Country sınıfının alt sınıfları olarak temsil etmektedir. TURCOVID19 veri setleri ise CODO sınıfı Country Wise Statistics'in bir alt sınıfı olarak temsil edilmektedir. Oluşturulan sınıflardaki gruplandırma ve sınıflara ait alt sınıflar düzey ilişkisi ile 1. Düzey, 2. Düzey. 3. Düzey ve 4. Düzey olarak Tablo 6.'da gösterilmiştir. Böylece, karmaşık kavramları analiz etmek kolaylaşacaktır.

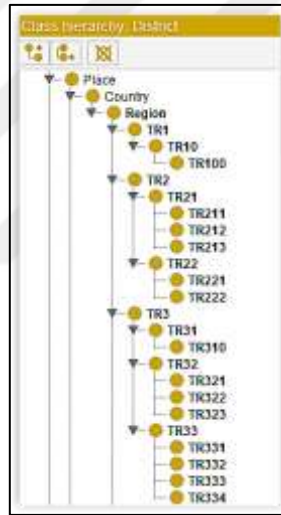
Tablo 6. TURCOVID19 Ontolojisi'nin farklı düzeydeki sınıfları

| 1.Düzyey | 2. Düzey | 3.Düzey | 4.Düzey |
|----------------------------|-------------------------|--|-------------------------------|
| Statistics | Country Wise Statistics | GenderandAgeGroup_Statistics | NumberOfNewPatients1 |
| | | | NumberOfNewPatients-1M |
| | | | NumberOfTotalPatients1-100K |
| | | | MortalityRateInConfirmedCases |
| | | | NumberOfTotalDeceased1 |
| | | ProvincialLevelDailyCOVID19VaccineData | |
| | | PivotTable-Total | |
| | | NumberOfProvincialLevelCases | |
| | | DailyData-COVID-19 | |
| | | PivotTable | |
| | Region_Statistics | NumberOfNewHospitalAdmissions | |
| | | NumberOfNewPatient | |
| | | NumberOfPatientsDischargedFromHospital | |
| | | NumberOfPatientsInTheLast7Days | |
| | | NumberOfPatientsInTheLast7Days-100K | |
| | | NumberOfTotalDeceased-100K | |
| | | NumberOfTotalPatients | |
| NumberOfTotalPatients-100K | | | |
| Patient | DiagnosedWith Covid | FemaleAgeGroup | K_0-2 |
| | | | K_2-4 |
| | | | K_5-14 |
| | | | K_15-24 |
| | | | K_25-49 |
| | | | K_50-64 |
| | | | K_65-79 |
| | | | K_80+ |
| | | MaleAgeGroup | E_0-2 |
| | | | E_0-2 |
| | | | E_2-4 |
| | | | E_5-14 |
| | | | E_15-24 |
| | | | E_25-49 |
| | | | E_50-64 |
| | | | E_65-79 |
| E_80+ | | | |
| Place | Country | Region | TR1 |
| | | | TR2 |
| | | | TR3 |
| | | | TR4 |
| | | | TR5 |
| | | | TR6 |
| | | | TR7 |
| | | | TR8 |
| | | | TR9 |
| | | | TRA |
| TRB | | | |
| TRC | | | |

2.4.5. Bilgi Sentezi (Knowledge Synthesis)

Bilgi sentezi (knowledge synthesis) adımı, kavramlar arasındaki ilişkileri tanımlayarak bilginin sentezlenmesini ve düzenlenmesini içerir. Böylece kavram hiyerarşilerinin keşfedilmesi sağlanmaktadır.

TURCOVID19 ontolojisinde bir kayıt Covid-19 vakalarına ait istatistikleri ifade eden Statistics sınıfının bir örneği olarak temsil edilir. İstatistiklerin kapsadığı alanları temsil eden sınıflar arasında hiyerarşi kurulmuştur. Şekil 31’de verilen girinti hiyerarşiyi göstermektedir. Türkiye İBBS’yi, Düzey 1 (12 bölge), Düzey 2 (26 alt bölge) ve Düzey 3 (81 il) kodlarına göre ifade etmektedir. Hiyerarşide de 12 bölgenin alt bölgeleri ve alt bölgelerin de illeri kapsadığı gösterilmektedir.



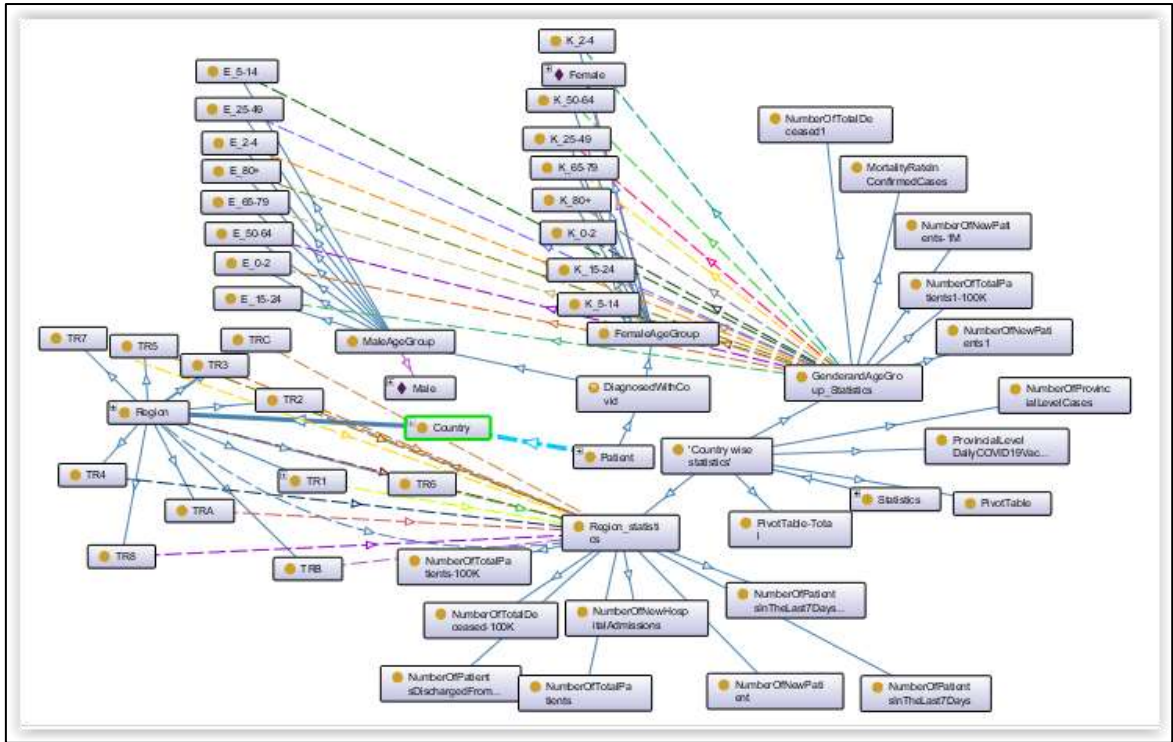
Şekil 31. İBBS Bölgeleri'ne ait sınıflar ve hiyerarşik gösterimi

2.4.6. Yeniden Kullanım ve Standardizasyon (Reuse and standardization)

Gerekli olan alan kelime dağarcıklarını geliştirmek ve yeniden kullanmak için CODO ve GeoSPARQL ontolojileri import edilmiştir. Schema.org; cinsiyet ve konum gibi ortak kavramları, FOAF; kişi ve kuruluş sınıfları ve ilgili özellikleri, SNOMED CT ve OBO; klinik bulgu ve semptomları modellemek için kullanılmıştır.

2.4.7. Temsili Modelin Tasarımı (Design of Representational Model)

Yeniden Kullanım ve Standardizasyon adımıyla üretilen alan bilgisinin yapılandırılmasını ve modellenmesini içerir. Bu adımdaki amaç, sınıflar, özellikler ve bunların ilişkileri gibi çeşitli bileşenlerini gösteren alan bilgisini modellemektir. TURCOVID19 veri setlerindeki her bir tabloya ait sınıflar oluşturulmuştur. CODO Ontolojisi genişletilerek oluşturulan TURCOVID19 Ontolojisi Şekil 32’de verilmiştir.



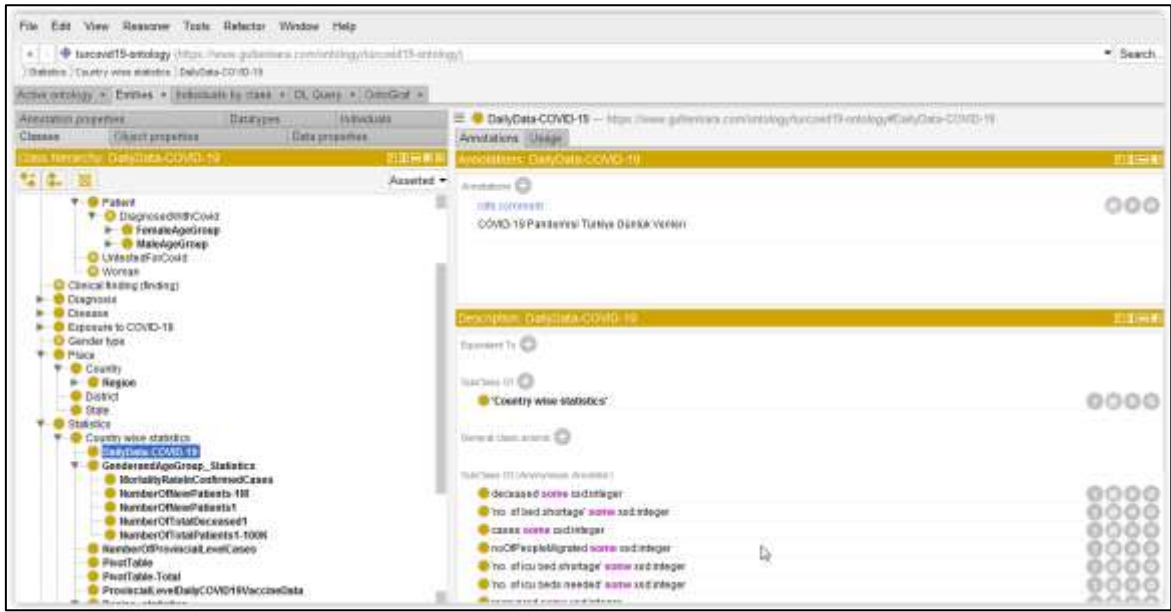
Şekil 32. TURCOVID19 Ontolojisi ontograph gösterimi

2.4.8. Ontoloji Geliştirme (Ontology development)

YAMO metodolojisinde ontoloji geliştirme adımı, biçimsel bir mantık dili kullanarak model geliştirmeyi içermektedir. TURCOVID19 ontolojisini geliştirmek için bir Açıklama Mantığı (Description Logic) ontoloji dili olan OWL-DL'yi kullanılmıştır. Ontoloji, Protégé -5.5.0 ontoloji editörü kullanılarak tasarlanmıştır.

2.4.8.1. Sınıflar (Classes)

TURCOVID19 veri setlerine ait tablo isimleri Şekil 33.'de görüldüğü gibi ontolojide CODO sınıfı olan CountryWiseStatistics'in bir alt sınıfları olarak temsil edilmektedir. Veri setlerinin kapsadığı İBBS bölgeleri, Country sınıfının alt sınıfları olarak belirtilmektedir. Hastaların ait oldukları cinsiyet ve yaş grupları ise virüs teşhisi konan herkesin o sınıfın üyesi olması için gerekli ve yeterli koşulları tanımlanmış olan DiagnosedWithCovid sınıfının alt sınıfları olarak temsil edilmektedir.

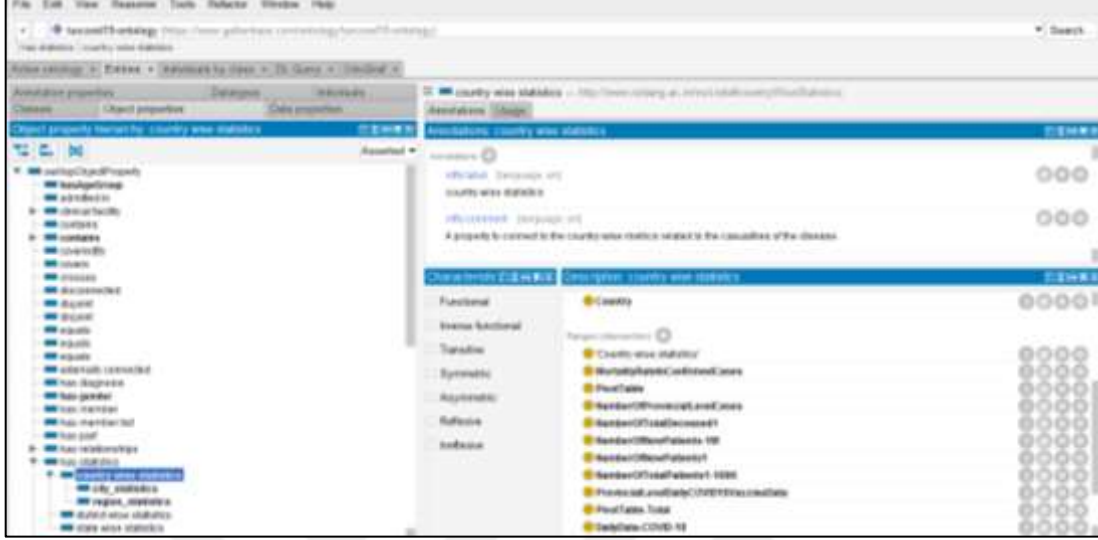


Şekil 33. TURCOVID19 Ontolojisi'ne ait sınıflar

2.4.8.2. Özellikler (Object-Data Properties)

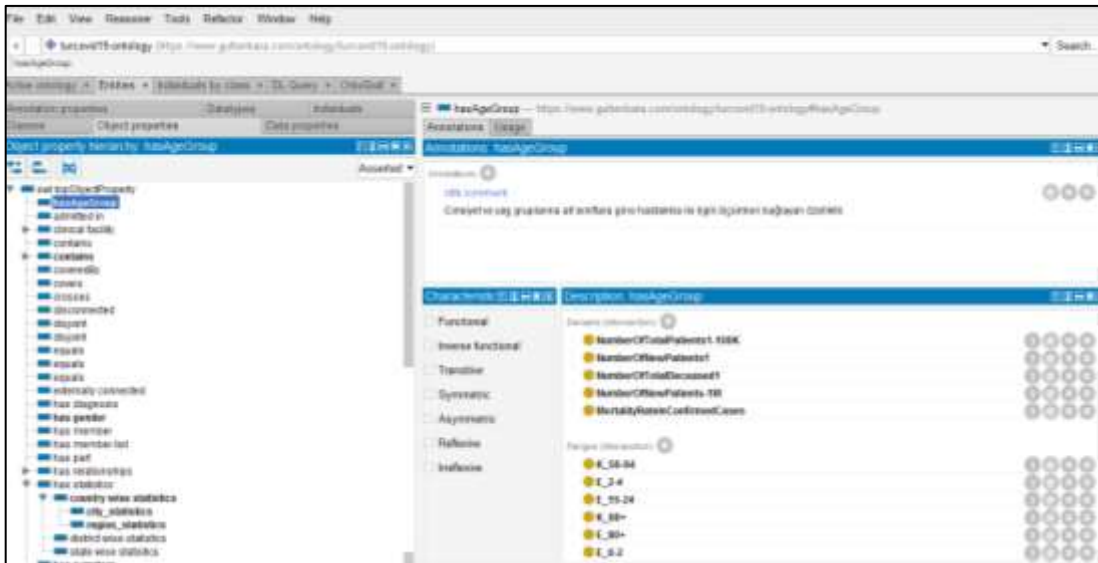
Geliştirilen ontolojide tanımlı özellikler gösterilmektedir. TURCOVID19 ontolojisi için belirli özelliklerin eklenebilir olması, ontolojinin uygulamanın gereksinimine göre genişletilebilmesini sağlamaktadır. TURCOVID19 ontolojisinde hastalıkla ilgili ülke geneli ölçümleri ilişkilendirmek için country wise statistics özelliği kullanılmıştır. Bölgesel olarak ilişki kurmak için bir alt özellikler olan region_statistics ve city_statistics özellikleri geliştirilmiştir. İBBS bölgelerini temsil eden sınıflar ile bölgelere ait ölçümleri temsil eden sınıflar arasındaki ilişkiler için region_statistics ve city_statistics özellikleri kullanılmıştır.

İlgili ontoloji sınıfları ve aralarındaki ilişkileri temsil eden nesne özellikleri (object properties) Şekil 34’te verilmiştir.



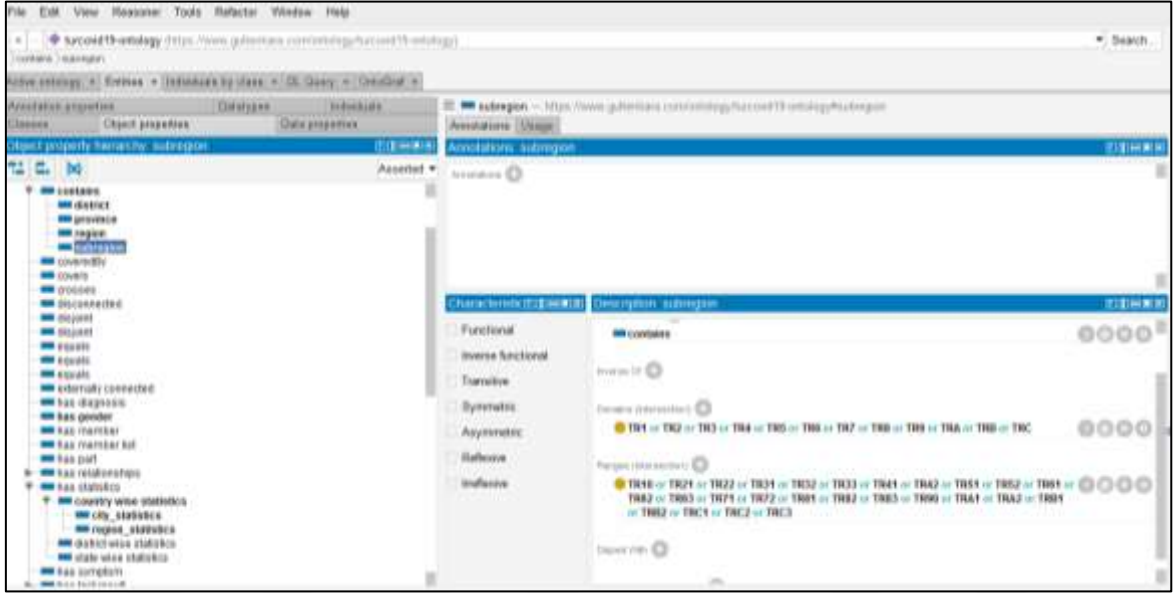
Şekil 34. Ontoloji sınıfları ve aralarındaki ilişkileri temsil eden nesne özellikleri

Şekil 35’te gösterilen GenderandAgeGroup_Statistics sınıfı ve alt sınıflarının hangi cinsiyet ve yaş gruplarına ait olduklarını belirten hasAgeGroup ilişkisi kurulmuştur. Ayrıca Türkiye geneline ait olan cinsiyet ve yaş grupları ile Country sınıfı arasındaki ilişkiyi location özelliği temsil etmektedir.



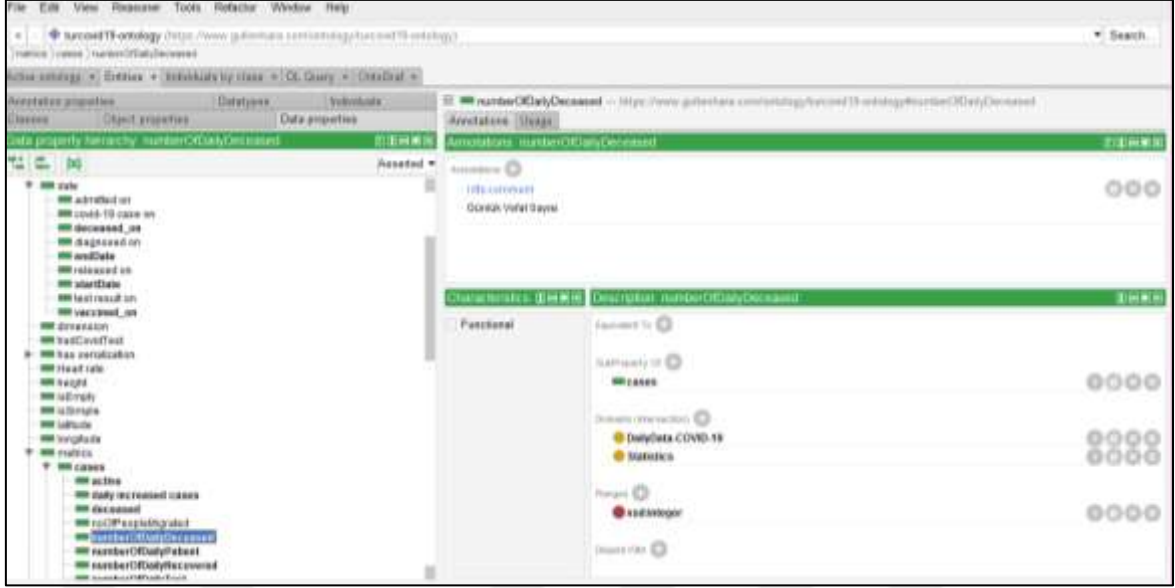
Şekil 35. hasAgeGroup ilişkisinin gösterimi

Türkiye İBBS’de 26 bölgenin 12 alt bölgeyi, 12 alt bölgenin de 81 ili kapsadığını ifade eden GeoSPARQL tarafından tanımlanan contains nesne ilişkisi kullanılmıştır. Şekil 36.’da gösterildiği gibi contains nesne ilişkisi ile “Batı Marmara (TR2), Tekirdağ alt bölgesi (TR21)’ni içerir”; “Tekirdağ alt bölgesi (TR21), Tekirdağ (TR211), Edirne (TR212), Kırklareli (TR213)’ni içerir.” konumsal çıkarsama sağlanabilmektedir.



Şekil 36. Contains ilişkisinin gösterimi

Sınıflara ait veri öznitelikleri (data properties), COVID-19 hastalığı ile ilgili tüm ölçümleri depolamak için genel bir veri özelliği olan Metrics’in alt özellikleri olarak oluşturulmuştur. Sınıflar için data propertyler tanımlanmıştır. Şekil 37’de gösterildiği gibi her bir veri özelliği için domain-range ve annotation alanlarında ilgili bilgiler sınıf ve literal değerler ile ilişkilendirilmiştir.



Şekil 37. Veri özelliklerinin domain-range ve annation alanlarındaki ilişkilerin gösterimi

Ontolojide, TURCOVID19 platformu tarafından yayınlanan veri setlerine ait tablo isimleri sınıf olarak tanımlanmıştır. Tablo 7.'de her bir sınıfa ait verilerin özniteliklerine karşılık gelecek şekilde data propertyler gösterilmiştir. Data propertylere ait domain alanı verileri temsil eden sınıf, range alanı ise veri tipine göre seçilmiştir.

Tablo 7. Sınıflara ait veri özelliklerinin gösterilmesi

| Veri Setleri | Classes | Öznitelikler | Data Properties |
|--|--|---------------------|--|
| Türkiye İl Düzeyi Anlık COVID19 Aşı Verileri | ProvincialLevelDailyCOVID19VaccineData | İl | cityName |
| | | Tarih | vaccined_on |
| | | Toplam Doz | numberOfTotalDoses |
| | | Toplam 1. Doz | numberOfTotalFirstDoses |
| | | Toplam 2. Doz | numberOfTotalSecondDoses |
| | | Toplam 3. Doz | numberOfTotalThirdDoses |
| | | Toplam 4. Doz | numberOfTotalFourthDoses |
| | | Günlük Toplam Doz | numberOfDailyDose |
| | | Günlük 1. Doz | numberOfDailyFirstDose |
| | | Günlük 2. Doz | numberOfDailySecondDose |
| | | Günlük 3. Doz | numberOfDailyThirdDose |
| Günlük 4. Doz | numberOfDailyFourthDose | | |
| COVID-19 Özet Tablo, | | Tarih | 'covid-19 case on' |
| | | Yeni Test | numberOfNewTest |
| | | Yeni Hasta | numberOfNewPatient |
| | | Yeni Hastane Yatışı | numberOfNewHospitalAdmissions |
| | | Yeni Entübe | numberOfNewIntubated |
| | | Yeni Taburcu | numberOfPatientsDischargedFromHospital |

Tablo 7'nin devamı

| | | | | | |
|---|------------------------|--|----------------------|-------|------------------|
| İBBS-1'e Göre Yeni COVID-19 Hasta Sayısı, Türkiye | NumberOfNewPatient | Tarih | 'covid-19 case on' | | |
| | | İstanbul | TR1 | | |
| | | Batı Marmara | TR2 | | |
| | | Ege | TR3 | | |
| | | Doğu Marmara | TR4 | | |
| | | Batı Anadolu | TR5 | | |
| | | Akdeniz | TR6 | | |
| | | Orta Anadolu | TR7 | | |
| | | Batı Karadeniz | TR8 | | |
| | | Doğu Karadeniz | TR9 | | |
| | | Kuzeydoğu Anadolu | TRA | | |
| | | Orta Doğu Anadolu | TRB | | |
| | | Güneydoğu Anadolu | TRC | | |
| | | Cinsiyet ve Yaş Grubuna Göre Yeni COVID-19 Hasta Sayısı, Türkiye | NumberOfNewPatients1 | Tarih | covid-19 case on |
| E_<2 | E_0-2 | | | | |
| E_2-4 | E_2-4 | | | | |
| E_5-14 | E_5-14 | | | | |
| E_15-24 | E_15-24 | | | | |
| E_25-49 | E_25-49 | | | | |
| E_50-64 | E_50-64 | | | | |
| E_65-79 | E_65-79 | | | | |
| E_80+ | E_80+ | | | | |
| K_<2 | K_0-2 | | | | |
| K_2-4 | K_2-4 | | | | |
| K_5-14 | K_5-14 | | | | |
| K_15-24 | K_15-24 | | | | |
| K_25-49 | K_25-49 | | | | |
| K_50-64 | K_50-64 | | | | |
| K_65-79 | K_65-79 | | | | |
| K_80+ | K_80+ | | | | |
| Cinsiyet ve Yaş Grubuna Göre 1.000.000 Kişiye Düşen Yeni COVID-19 Hasta Sayısı, Türkiye | NumberOfNewPatients-1M | | | Tarih | covid-19 case on |
| | | | | E_<2 | E_0-2 |
| | | E_2-4 | E_2-4 | | |
| | | E_5-14 | E_5-14 | | |
| | | E_15-24 | E_15-24 | | |
| | | E_25-49 | E_25-49 | | |
| | | E_50-64 | E_50-64 | | |
| | | E_65-79 | E_65-79 | | |
| | | E_80+ | E_80+ | | |
| | | K_<2 | K_0-2 | | |
| | | K_2-4 | K_2-4 | | |
| | | K_5-14 | K_5-14 | | |
| | | K_15-24 | K_15-24 | | |
| | | K_25-49 | K_25-49 | | |
| | | K_50-64 | K_50-64 | | |
| | | K_65-79 | K_65-79 | | |
| | | K_80+ | K_80+ | | |

Tablo 7'nin devamı

| | | | |
|---|--|-----------------------|-------------------------------------|
| İBBS-1'e Göre Hastane Hizmetleri ile İlgili COVID-19 Göstergeleri, Yeni hastane yatışları, Türkiye | NumberOfNewHospital Admissions | Tarih | 'covid-19 case on' |
| | | İstanbul | TR1 |
| | | Batı Marmara | TR2 |
| | | Ege | TR3 |
| | | Doğu Marmara | TR4 |
| | | Batı Anadolu | TR5 |
| | | Akdeniz | TR6 |
| | | Orta Anadolu | TR7 |
| | | Batı Karadeniz | TR8 |
| | | Doğu Karadeniz | TR9 |
| | | Kuzeydoğu Anadolu | TRA |
| | | Orta Doğu Anadolu | TRB |
| | | Güneydoğu Anadolu | TRC |
| İBBS-1'e Göre Hastane Hizmetleri ile İlgili COVID-19 Göstergeleri, Hastaneden Yeni taburcu edilenler, Türkiye | NumberOfPatientsDischargedFromHospital | Tarih | 'covid-19 case on' |
| | | İstanbul | TR1 |
| | | Batı Marmara | TR2 |
| | | Ege | TR3 |
| | | Doğu Marmara | TR4 |
| | | Batı Anadolu | TR5 |
| | | Akdeniz | TR6 |
| | | Orta Anadolu | TR7 |
| | | Batı Karadeniz | TR8 |
| | | Doğu Karadeniz | TR9 |
| | | Kuzeydoğu Anadolu | TRA |
| | | Orta Doğu Anadolu | TRB |
| | | Güneydoğu Anadolu | TRC |
| Türkiye COVID-19 özet tablosu – Toplam | PivotTable-Total | Tarih | 'covid-19 case on' |
| | | Toplam Test | numberOfTotalTest1 |
| | | Toplam Hasta | numberOfTotalPatients |
| | | Toplam Ölüm | numberOfTotalDeceased |
| | | Toplam Hastane Yatışı | numberOfTotalDeceased |
| | | Toplam Entübe Hasta | numberOfTotalIntubated |
| | | Toplam Taburcu | numberOfTotalDischargedFromHospital |
| | | Toplam İyileşen | numberOfTotalRecovered |
| İBBS-1'e göre laboratuvar onaylı COVID-19 Toplam Hasta sayısı | NumberOfTotalPatients | Tarih | covid-19 case on |
| | | İstanbul | TR1 |
| | | Batı Marmara | TR2 |
| | | Ege | TR3 |
| | | Doğu Marmara | TR4 |
| | | Batı Anadolu | TR5 |
| | | Akdeniz | TR6 |
| | | Orta Anadolu | TR7 |
| | | Batı Karadeniz | TR8 |
| | | Doğu Karadeniz | TR9 |
| | | Kuzeydoğu Anadolu | TRA |
| | | Orta Doğu Anadolu | TRB |
| | | Güneydoğu Anadolu | TRC |
| Türkiye | TR | | |

Tablo 7'nin devamı

| | | | |
|---|-------------------------------------|-------------------|--------------------|
| İBBS-1'e göre laboratuvar onaylı COVID-19 Hasta / 100.000 nüfus | NumberOfTotalPatients-100K | Tarih | 'covid-19 case on' |
| | | İstanbul | TR1 |
| | | Batı Marmara | TR2 |
| | | Ege | TR3 |
| | | Doğu Marmara | TR4 |
| | | Batı Anadolu | TR5 |
| | | Akdeniz | TR6 |
| | | Orta Anadolu | TR7 |
| | | Batı Karadeniz | TR8 |
| | | Doğu Karadeniz | TR9 |
| | | Kuzeydoğu Anadolu | TRA |
| | | Orta Doğu Anadolu | TRB |
| | | Güneydoğu Anadolu | TRC |
| | | Türkiye | TR |
| İBBS-1'e göre laboratuvar onaylı COVID-19 Son 7 Günlük Hasta Sayısı | NumberOfPatientsInTheLast7Days | Tarih | 'covid-19 case on' |
| | | İstanbul | TR1 |
| | | Batı Marmara | TR2 |
| | | Ege | TR3 |
| | | Doğu Marmara | TR4 |
| | | Batı Anadolu | TR5 |
| | | Akdeniz | TR6 |
| | | Orta Anadolu | TR7 |
| | | Batı Karadeniz | TR8 |
| | | Doğu Karadeniz | TR9 |
| | | Kuzeydoğu Anadolu | TRA |
| | | Orta Doğu Anadolu | TRB |
| | | Güneydoğu Anadolu | TRC |
| | | Türkiye | TR |
| İBBS-1'e göre laboratuvar onaylı COVID-19 Son 7 Gün İnsidansı (100.000 nüfusta) | NumberOfPatientsInTheLast7Days-100K | Tarih | covid-19 case on |
| | | İstanbul | TR1 |
| | | Batı Marmara | TR2 |
| | | Ege | TR3 |
| | | Doğu Marmara | TR4 |
| | | Batı Anadolu | TR5 |
| | | Akdeniz | TR6 |
| | | Orta Anadolu | TR7 |
| | | Batı Karadeniz | TR8 |
| | | Doğu Karadeniz | TR9 |
| | | Kuzeydoğu Anadolu | TRA |
| | | Orta Doğu Anadolu | TRB |
| | | Güneydoğu Anadolu | TRC |
| | | Türkiye | TR |
| Cinsiyet ve Yaş Grubuna Göre 100.000 Kişiye Düşen COVID-19 Hasta Sayısı | NumberOfTotalPatients1-100K | Tarih | 'covid-19 case on' |
| | | E_<2 | E_0-2 |
| | | E_2-4 | E_2-4 |
| | | E_5-14 | E_5-14 |
| | | E_15-24 | E_15-24 |
| | | E_25-49 | E_25-49 |
| | | E_50-64 | E_50-64 |
| | | E_65-79 | E_65-79 |

Tablo 7'nin devamı

| | | | | | |
|---|-------------------------------|---|----------------------------|----------|--------------------|
| Cinsiyet ve Yaş Grubuna Göre 100.000 Kişiye Düşen COVID-19 Hasta Sayısı | | E_80+ | E_80+ | | |
| | | K_<2 | K_0-2 | | |
| | | K_2-4 | K_2-4 | | |
| | | K_5-14 | K_5-14 | | |
| | | K_15-24 | K_15-24 | | |
| | | K_25-49 | K_25-49 | | |
| | | K_50-64 | K_50-64 | | |
| | | K_65-79 | K_65-79 | | |
| | | K_80+ | K_80+ | | |
| Cinsiyet ve yaş grubuna göre onaylanmış vakalar içindeki ölüm hızı, (%) | MortalityRateInConfirmedCases | Tarih | 'covid-19 case on' | | |
| | | E_<2 | E_0-2 | | |
| | | E_2-4 | E_2-4 | | |
| | | E_5-14 | E_5-14 | | |
| | | E_15-24 | E_15-24 | | |
| | | E_25-49 | E_25-49 | | |
| | | E_50-64 | E_50-64 | | |
| | | E_65-79 | E_65-79 | | |
| | | E_80+ | E_80+ | | |
| | | K_<2 | K_0-2 | | |
| | | K_2-4 | K_2-4 | | |
| | | K_5-14 | K_5-14 | | |
| | | K_15-24 | K_15-24 | | |
| | | K_25-49 | K_25-49 | | |
| | | K_50-64 | K_50-64 | | |
| | | K_65-79 | K_65-79 | | |
| | | K_80+ | K_80+ | | |
| | | Cinsiyet ve yaş grubuna göre Toplam ölüm sayıları | NumberOfTotalDeceased 1 | Tarih | 'covid-19 case on' |
| | | | | E_<2 | E_0-2 |
| E_2-4 | E_2-4 | | | | |
| E_5-14 | E_5-14 | | | | |
| E_15-24 | E_15-24 | | | | |
| E_25-49 | E_25-49 | | | | |
| E_50-64 | E_50-64 | | | | |
| E_65-79 | E_65-79 | | | | |
| E_80+ | E_80+ | | | | |
| K_<2 | K_0-2 | | | | |
| K_2-4 | K_2-4 | | | | |
| K_5-14 | K_5-14 | | | | |
| K_15-24 | K_15-24 | | | | |
| K_25-49 | K_25-49 | | | | |
| K_50-64 | K_50-64 | | | | |
| K_65-79 | K_65-79 | | | | |
| K_80+ | K_80+ | | | | |
| İBBS-1'e göre toplam ölüm sayıları dağılımı | NumberOfTotalDeceased | | | Tarih | 'covid-19 case on' |
| | | | | İstanbul | TR1 |
| | | Batı Marmara | TR2 | | |
| | | Ege | TR3 | | |
| | | Doğu Marmara | TR4 | | |
| | | Batı Anadolu | TR5 | | |
| Akdeniz | TR6 | | | | |

Tablo 7'nin devamı

| | | | |
|---|------------------------------|------------------------------------|------------------------------------|
| İBBS-1'e göre toplam ölüm sayıları dağılımı | | Orta Anadolu | TR7 |
| | | Batı Karadeniz | TR8 |
| | | Doğu Karadeniz | TR9 |
| | | Kuzeydoğu Anadolu | TRA |
| | | Orta Doğu Anadolu | TRB |
| | | Güneydoğu Anadolu | TRC |
| | | Türkiye | TR |
| İBBS-1'e göre ölüm insidansı, (100.000 nüfusta) | NumberOfTotalDeceased-100K | Tarih | 'covid-19 case on' |
| | | İstanbul | TR1 |
| | | Batı Marmara | TR2 |
| | | Ege | TR3 |
| | | Doğu Marmara | TR4 |
| | | Batı Anadolu | TR5 |
| | | Akdeniz | TR6 |
| | | Orta Anadolu | TR7 |
| | | Batı Karadeniz | TR8 |
| | | Doğu Karadeniz | TR9 |
| | | Kuzeydoğu Anadolu | TRA |
| | | Orta Doğu Anadolu | TRB |
| | | Güneydoğu Anadolu | TRC |
| | | Türkiye | TR |
| Türkiye İl Düzeyi Haftalık Vaka Sayıları | NumberOfProvincialLevelCases | tarih_baslangic | startDate |
| | | tarih_bitis | endDate |
| | | il | hasName |
| | | 100bin_kiside_haftalik_vaka_sayisi | numberOfWeeklyCases-100K |
| COVID-19 Pandemisi Türkiye Günlük Verileri | DailyData-COVID-19 | Tarih | 'covid-19 case on' |
| | | Toplam Vaka | cases |
| | | Günlük Vaka | 'daily increased cases |
| | | Toplam Hasta | numberOfPatients |
| | | Gunluk_Hasta | numberOfDailyPatient |
| | | Toplam Vefat | deceased |
| | | Gunluk_Vefat | numberOfDailyDeceased |
| | | Toplam iyileşen | recovered |
| | | Gunluk_iyileşen | numberOfDailyRecovered |
| | | Toplam_Test | numberOfTotalTest |
| | | Gunluk_Test | numberOfDailyTest |
| | | YBU | numberOfICU |
| | | Entube | numberOfIntubated |
| | | Ağır Hasta | numberOfSeriouslyIll |
| | | yatak_doluluk_orani | hasAdultIntensiveCareOccupancyRate |
| | | eriskin_yogun_bakim_doluluk_orani | hasAdultIntensiveCareOccupancyRate |
| ventilator_doluluk_orani | hasVentilatorOccupancyRate | | |

2.4.9. Değerlendirme (Evaluation)

Bu adım, ontolojinin tasarım hedeflerini ne kadar yakından karşıladığını değerlendirmeyi içermektedir. Ontolojinin teknik yeterliliğini ölçer. TURCOVID19 ontolojisinin tasarlandığı amaca hizmet ettiğini doğrulamak için, Karma 2.2 kullanarak TURCOVID19 platformundan sağlanan veriler ontoloji ile ilişkilendirilmiştir. Ayrıca 2.4.2'de açıklanan yetkinlik sorularına dayalı olarak SPARQL sorguları yazılmıştır.

İlişkisel veritabanlarında SQL sorguları gibi, SPARQL sorgu motoru da kabaca OWL'ye benzemektedir. Ancak, OWL'nin altında yatan yapı tablolardan ziyade grafikler olduğundan, SPARQL grafik desenleri oluşturur ve ardından grafik deseniyle eşleşen tüm bireyler için bilgi grafiklerini arar. SQL gibi, SPARQL de sorgulamaktan fazlasını yapabilir. Verileri silebilir, ekleyebilir ve dönüştürebilir (DuCharme, 2011).

SPARQL ile birden çok heterojen veri kaynağından gelen veriler entegre edilir. Böylece, birleşik bir sanal bilgi tabanı sağlanmakta ve birden çok kaynaktan geniş veri kümeleri arasında sorgular yapılabilmektedir. Örnek bir SPARQL sorgusu Şekil 38'de gösterilmiştir. Önekler (prefixes) önce sorgunun kullanacağı çeşitli ad alanlarını ve bunların IRI'lerini tanımlamaktadır. SPARQL sorgusunu kullanmanın bir avantajı, illere ait sorguyla eşleşen aşılama miktarındaki toplam doz sayısını (?toplamdoz sütunu) görmenin yanı sıra, 100 bin kişide haftalık vaka sayısını (?haftalıkvaka sütunu) görebilmemizdir. Örnek sorguda aşı sayılarına göre vaka sayılarındaki değişimi değerlendirmek mümkün olabilir.

SPARQL sorgularının Fuseki ile gerçekleştirilmesi için kullanılan URI adresleri örnekleri ile birlikte aşağıda sıralanmıştır.

- PREFIX geo: <http://www.opengis.net/ont/geosparql#>
- PREFIX foaf: <http://xmlns.com/foaf/0.1/>
- PREFIX fo: <http://purl.org/ontology/fo/>
- PREFIX rdf: <http://www.w3.org/1999/02/22-rdf-syntax-ns#>
- PREFIX owl: <http://www.w3.org/2002/07/owl#>
- PREFIX xsd: <http://www.w3.org/2001/XMLSchema#>
- PREFIX rdfs: <http://www.w3.org/2000/01/rdf-schema#>
- PREFIX codo: <http://www.isibang.ac.in/ns/codo#>
- PREFIX

turcovid19-ontology: <https://www.gultenkara.com/ontology/turcovid19-ontology#>

```

18 - where {
19   ?a turcovid19-ontology:vaccined_on ?date.
20   ?a turcovid19-ontology:numberOfTotalDoses ?toplamdaz.
21   ?a turcovid19-ontology:cityName ?city.
22   ?b turcovid19-ontology:numberOfWeeklyCases-1000 ?haftalikvaka.
23   ?b turcovid19-ontology:cityName ?city.
24   ?b turcovid19-ontology:startDate ?bas.
25   ?b turcovid19-ontology:endDate ?bit.
26
27   FILTER (?bas="15.05.2021")
28   FILTER (?date="21.05.2021 23:46:56")
29
30 }
31 ORDER BY (?city )

```

| city | toplamdaz | date | haftalikvaka | bas | bit |
|------------|-----------|---------------------|--------------|------------|------------|
| 1 Adana | 665.640 | 21.05.2021 23:46:56 | 18,90 | 15.05.2021 | 21.05.2021 |
| 2 Adıyaman | 148.285 | 21.05.2021 23:46:56 | 40,52 | 15.05.2021 | 21.05.2021 |
| 3 Afyon | 260.197 | 21.05.2021 23:46:56 | 82,64 | 15.05.2021 | 21.05.2021 |
| 4 Aksaray | 117.445 | 21.05.2021 23:46:56 | 58,86 | 15.05.2021 | 21.05.2021 |
| 5 Amasya | 165.219 | 21.05.2021 23:46:56 | 79,29 | 15.05.2021 | 21.05.2021 |
| 6 Ankara | 2.294.500 | 21.05.2021 23:46:56 | 123,02 | 15.05.2021 | 21.05.2021 |
| 7 Antalya | 974.137 | 21.05.2021 23:46:56 | 47,01 | 15.05.2021 | 21.05.2021 |
| 8 Ardahan | 94.199 | 21.05.2021 23:46:56 | 110,23 | 15.05.2021 | 21.05.2021 |

Şekil 38. İllere ait yapılan toplam doz aşı sayısına göre 100 bin kişideki haftalık vakasayısı sorgusu

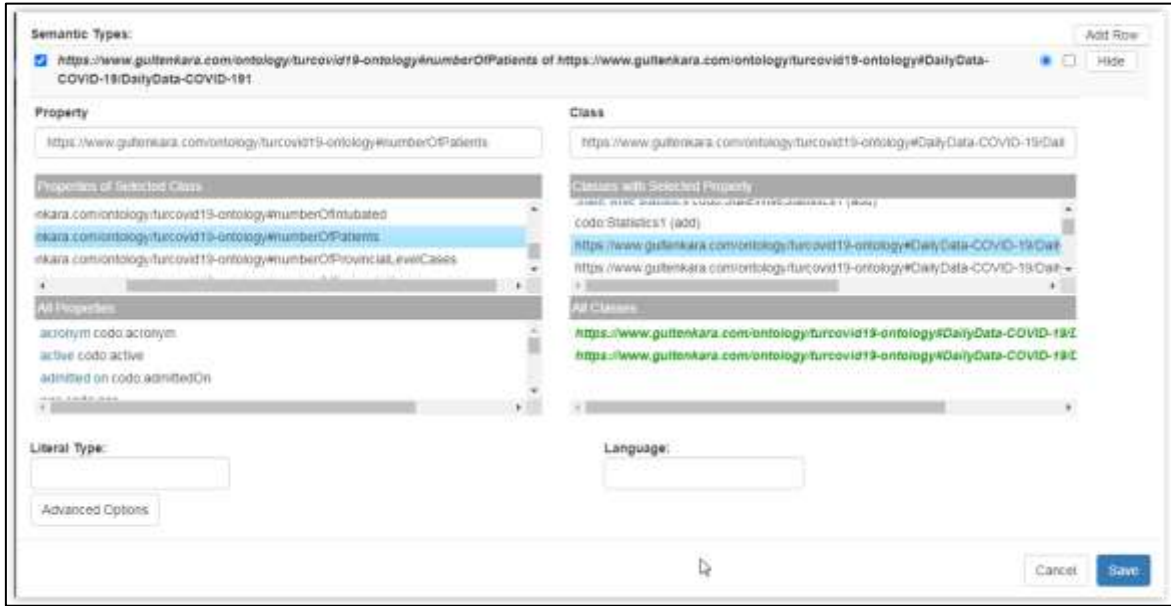
2.5. COVID-19 Verilerinin Semantik Tanımlarının Oluşturulması

TURCOVID19 platformundan temin edilen COVID-19 verilerinin semantik tanımlarının oluşturulması için gerekli ontolojilerin veriler ile ilişkilendirilmesinde Karma yazılımı kullanılmıştır. Özellikle ontolojiler ve veriler arasındaki ilişkilerin tanımlanmasında Karma yazılımı büyük kolaylık sağlamaktadır. Import -> From File yolu ontolojiler ve veriler import edilmektedir. Şekil 39'de verildiği gibi ontolojiler ve veri import edildikten sonra veri tablosal biçimde görüntülenmektedir.

| Suggest | Toplam_Vaka | Gunluk_Vaka | Takip_Hizni | Gunluk_Hizni | Toplam_Vefat | Gunluk_Vefat | Toplam_Syismem | Gunluk_Syismem |
|------------|-------------|-------------|-------------|--------------|--------------|--------------|----------------|----------------|
| 11.03.2020 | 1 | 1 | 1 | | | | | |
| 12.03.2020 | 1 | 0 | 0 | | | | | |
| 13.03.2020 | 5 | 2 | 1 | | | | | |
| 14.03.2020 | 5 | 5 | 3 | | | | | |
| 15.03.2020 | 19 | 16 | 15 | | | | | |
| 16.03.2020 | 47 | 47 | 29 | | | | | |
| 17.03.2020 | 98 | 98 | 51 | 1 | 1 | | | |
| 18.03.2020 | 191 | 191 | 95 | 2 | 1 | | | |
| 19.03.2020 | 369 | 369 | 168 | 4 | 2 | | | |
| 20.03.2020 | 670 | 670 | 311 | 9 | 5 | | | |

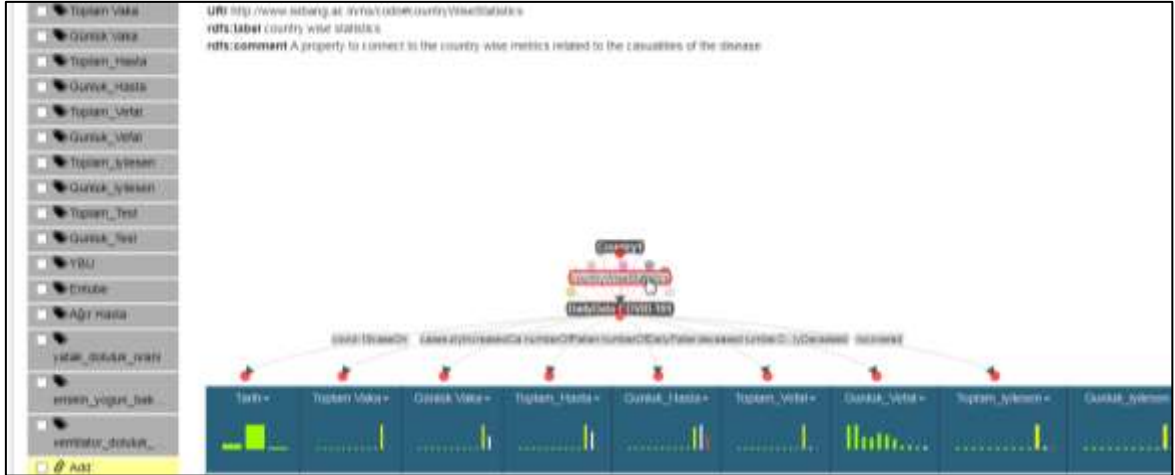
Şekil 39. Karma ara yüzünde covid19_gunlukveriler.csv dosyasının görüntülenmesi

Verilerin semantik tanımlarının oluşturulmasında için CODO ve TURCOVID19 Ontoloji'leri kullanılmıştır ve bu ontolojiler Karma'da import edilmiştir. Tablo'da her bir öznitelik üzerine tıklanarak ontolojiler ile semantik tipler ilişkilendirilmiştir. Şekil 40'da görüldüğü gibi covid19_gunlukveriler.csv dosyasındaki toplam hasta özniteliği TURCOVID19 Ontolojisi'ndeki DailyData-COVID-19 sınıfı ile arasında numberOfPatients ilişkisi olacak şekilde tanımlanmıştır. Her bir sütun için tablodaki verilerin semantik tip tanımlaması bu şekilde gerçekleştirilmiştir.



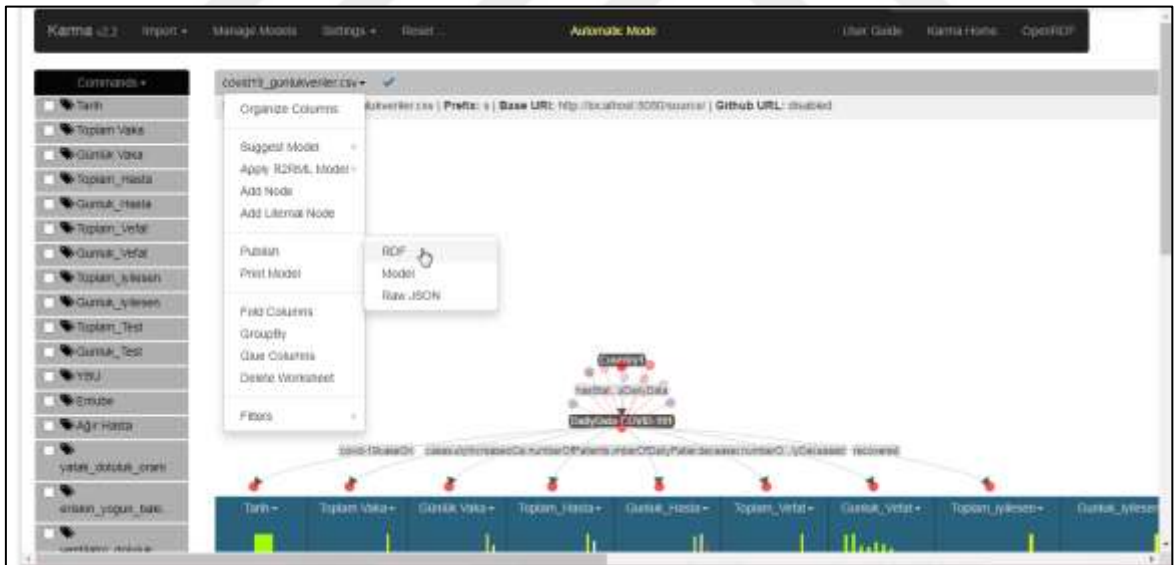
Şekil 40. Karma ara yüzünde semantik tiplerin tanımlanması

Tabloda bulunan öznitelikler için ilişkilerin tanımlanmasının ardından veri modellemenin ikinci aşaması olarak sınıflar arasındaki ilişkiler belirlenmiştir. Oluşturulan model sınıfı, özellikleri ve ilişkileri ile Şekil 41'deki gibi gösterilmektedir.



Şekil 41. Karma ara yüzünde semantik ilişkilerin tamamlanması

Veriler ile ontolojiler ilişkilendirildikten sonra Şekil 25'te verildiği gibi Karma'da covid19_gunlukveriler.csv üzerine tıklanarak Publish ->RDF yolu ile RDF dosyası oluşturulmuştur.



Şekil 42. Karma ara yüzünde RDF dosyasının oluşturulması

Dosya adının sağ tarafındaki RDF üzerine tıklanarak oluşturulan RDF dosyası açılır.



Şekil 44. LOD cloud üzerinden yayınlanan mevcut COVID-19 veri kaynakları

SILK ile kullanıcı tarafından sağlanan bağlantı özellikleri ve koşullarını kullanarak mevcut farklı veri kaynakları içindeki varlıklar arasında zaman açısından RDF eşleştirmelerinin bulunması için araçlar sağlamaktadır. Bağlantılı veri yayınlamak için, RDF'ye dönüştürülmüş verilerin webdeki veri kaynaklarına bağlantılar kurmak için SILK kullanılmaktadır (Ali vd., 2022). Uygulama senaryosu kapsamında linklerin kurulması için SILK yazılımı seçilmiştir. Oluşturulan RDF verisi ve TURCOVID19 ontolojisi SILK Workbench yazılımına yüklenerek DBpedia, Nuts, Wikidata, Covid-on-the-web ve WHOnto veri kaynakları girdi kaynak ve hedef veri setleri olarak tanımlanmıştır. Bağlantılı veri kaynaklarında bulunan owl:sameAs ilişkisi kurularak verilerle link kurulması işlemi bu aşamada gerçekleştirilir.

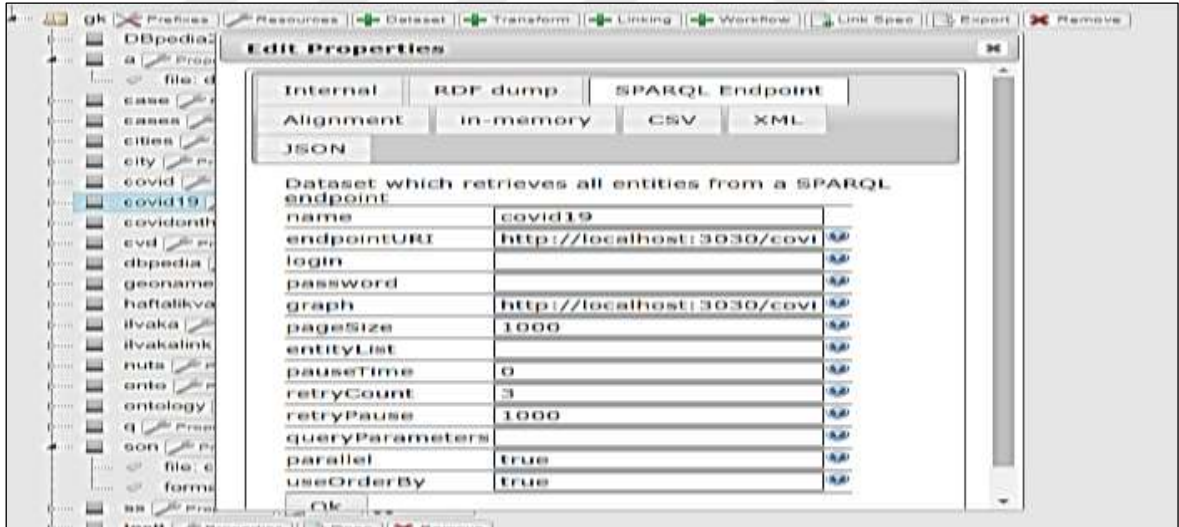
2.6.1. Workspace

Projede kullanılan URI prefixler belirlenmiştir. Şekilde gösterildiği gibi kullanılan prefixler workspace alanında Prefixes butonuna tıklayarak tanımlanmıştır.



Şekil 45. Prefixlerin tanımı

Dataset butonuna tıklayarak veri kaynağında yer alan varlıkların alınması sağlanmaktadır. Çalışma kapsamında oluşturulan RDF verileri ve TURCOVID19 ontolojisi, Fuseki Triple Store'a yüklenmiştir ve yayınlanmıştır. Şekil 46'da gösterildiği gibi veri kaynaklarının özellikleri düzenlenerek eklenmiştir.



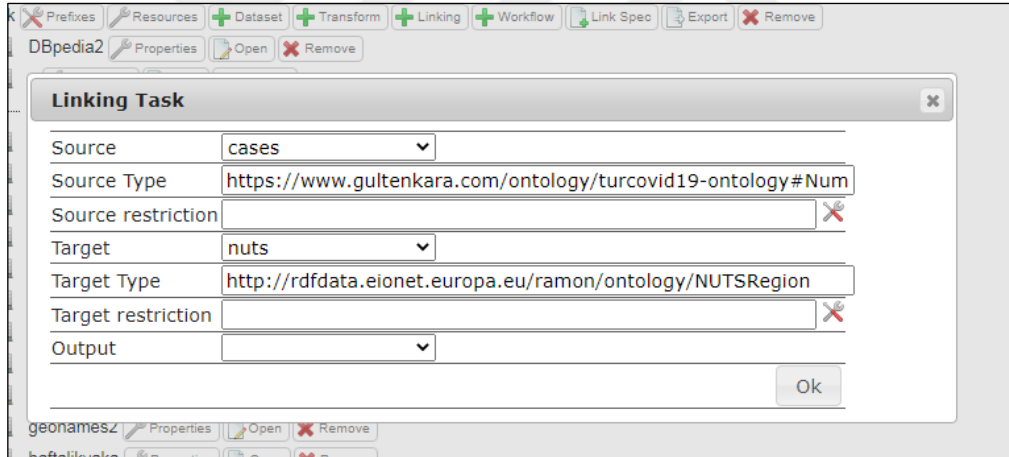
Şekil 46. SPARQL Endpoint ile veri kaynağının tanımlanması

Lod Cloud'da farklı veri kaynakları TURTLE ve RDF/XML formatında yer almaktadır. Şekil 47'de gösterildiği gibi RDF dump sekmesine tıklayarak veri kaynakları eklenmiştir.



Şekil 47. RDF dump ile veri kaynağının tanımlanması

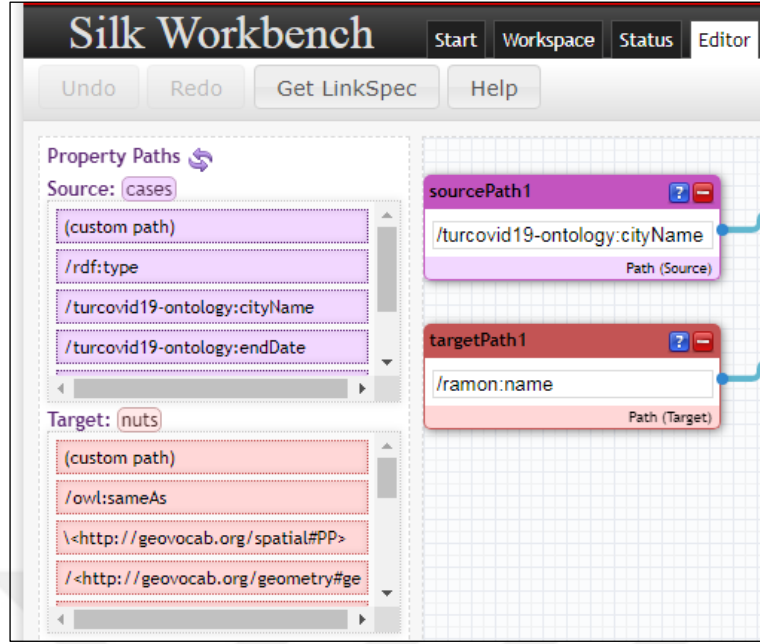
Şekil 48’de gösterildiği gibi Linking task ile aralarında link kurulacak kaynak veri seti ile hedef veri seti seçimi yapılmaktadır. Source type ve target type ile eşleştirmenin sınıf veya property düzeyinde olduğuna dair seçim yapılmaktadır.



Şekil 48. Linking Task gösterimi

2.6.2. Linkage Rule Editör

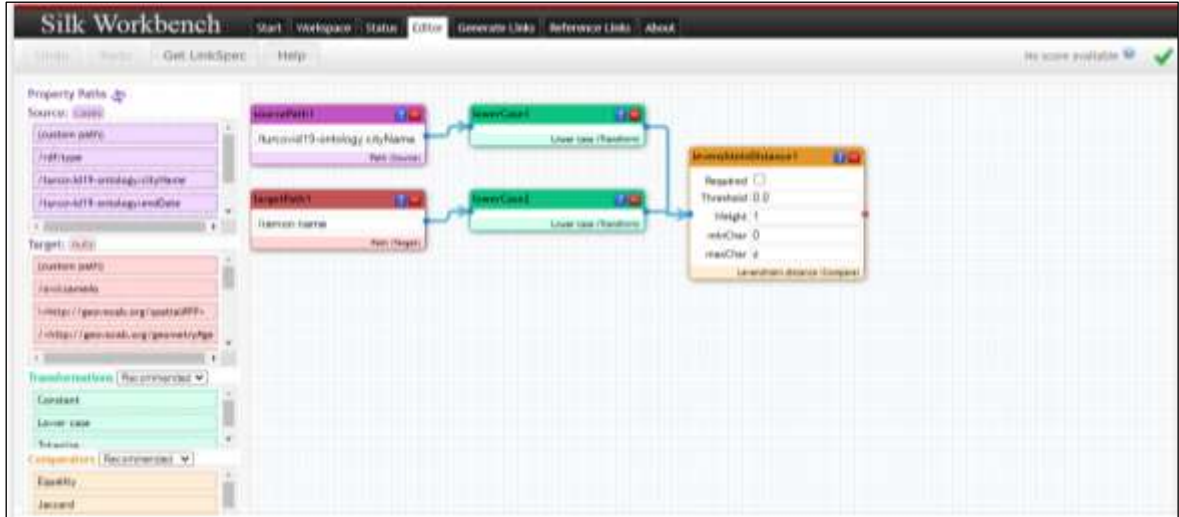
Path operatörü ile etiket özelliği (label property) gibi her varlığın belirli bir özellik yolunun tüm değerlerinin alınması sağlanmaktadır. Amaç, diğer operatörler için girdi olarak kullanılan veri kümesindeki değerlere erişim sağlanmaktadır. Şekil 49’da görüldüğü gibi sources olarak tanımlanan NumberOfProvincialLevelCases sınıfındaki turcovid-19:cityName özelliği ve target olarak tanımlanan NUTSRegion sınıfındaki ramon:name özelliği seçilmiştir. Böylece birbirine bağlanacak veri setlerin özellik yolları gösterilmiştir.



Şekil 49. Path operatörünün gösterimi

Transformation Operator ile değerlerin dönüştürülmesi, normalleştirilmesi veya birleştirilmesi sağlanmaktadır. lowerCase, numReduce, upperCase gibi fonksiyonlar vardır. Çalışma kapsamında bir dizeyi küçük harfe dönüştüren lowerCase fonksiyonu kullanılmıştır.

Comparison operator ile bir mesafe ölçümü ve bir mesafe eşiği uygulayarak iki özellik veya dönüşüm operatörü tarafından döndürülen değerlere dayalı olarak iki varlık arasındaki benzerlikler değerlendirilmektedir. Mesafe ölçülerine örnek olarak Levenshtein, Jaccard veya Geographic Distance verilebilir. Çalışma kapsamında Levenshtein Distance karşılaştırma operatörü kullanılmıştır. Levenshtein Distance, bir sözcüğü diğerine dönüştürmek için tek bir karakterin eklenmesi, silinmesi veya değiştirilmesi gibi düzenleme işlemlerini kapsayan minimum düzenleme sayısını hesaplayarak benzerlikleri bulmaya yardımcı olmaktadır (Bryl vd., 2014). Şekil 50’de Number of Provincial Level Cases sınıfında yer alan il isimleri ile NUTSRegion sınıfında yer alan il isimleri arasındaki benzerlikler Levenshtein Distance karşılaştırma operatörü ile bulunması istenmiştir.



Şekil 50. Levenshtein Distance karşılaştırma operatöründe bağlantı tanımı

2.6.3. Generating Links

Bağlantı kuralları tanımlandıktan sonra bağlantı sonuçları Şekil 51’de gösterildiği gibi Start butonuna tıklanarak sağlanmaktadır. İllere göre olan haftalık vaka sayılarındaki il isimleri ile NUTS veri setinde yer alan il isimleri arasındaki bağlantı sonuçları Şekil 51’de gösterilmiştir.

| Source | Target | Score | Correct |
|---------------------------|-----------------------------------|--------|---------|
| turcovid19-ontology:TR611 | http://nuts.geovocab.org/id/TR611 | 100.0% | ✓ |
| turcovid19-ontology:TR611 | http://nuts.geovocab.org/id/TR61 | 100.0% | ✓ |
| turcovid19-ontology:TR621 | http://nuts.geovocab.org/id/TR62 | 100.0% | ✓ |

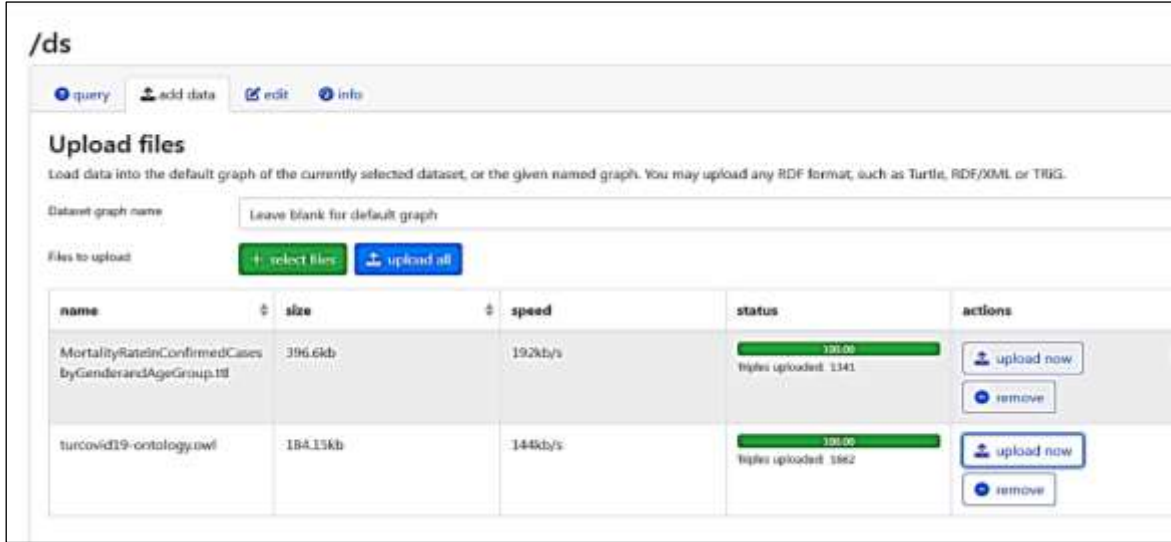
Şekil 51. Eşleştirmelerin belirlenmesi

Yapılan bu çalışma kapsamında, TURCOVID19 verileri, Covid-on-the-Web, NUTS (GeoVocab) datasetleri ve WHOnto, DBpedia, Wikidata ontolojileri ile ilişkilendirilmiştir. Linklerin oluşturulması ile TURCOVID19 verilerinin webde veri kullanıcıları tarafından bulunabilirliği sağlanmaktadır ve veri setinde yer almayan bilgilerle zenginleştirilmektedir. Bu işlem adımı ile Bağlantılı Veri İlkeleri'nden 4. Maddenin gereği yerine getirilmiş olur.

2.7. Bağlantılı Verinin Yayınlanması

COVID-19 verilerinin, RDF dosyaları üzerinde sorgulamaların gerçekleştirilmesi için veri tabanlarında depolanması gerekir. TURCOVID19 verilerinden elde edilen RDF üçlülerini depolamak için Triple Store'lara ihtiyaç vardır. Uygulamada verilerin yayınlanması ve sorgulanması için Fuseki kullanılmıştır. Fuseki üzerinde sırasıyla Şekil 52'de dosya yükleme, Şekil 53'de SPARQL sorgu dili ile sorgulama ve Şekil 54'de RDF üçlüsü sorgulama sonucu gösterilmiştir.

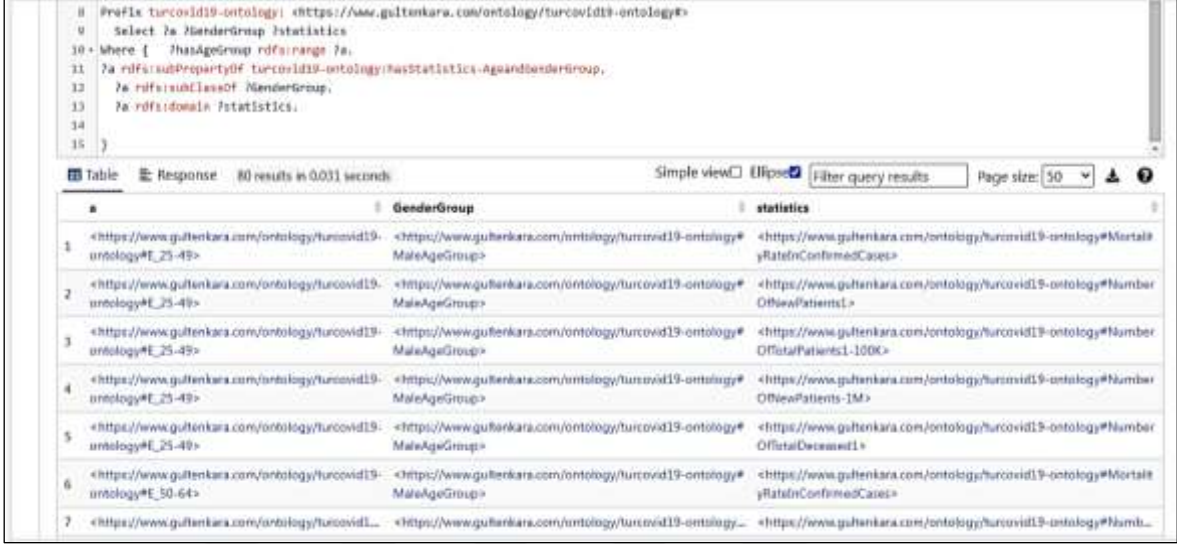
Şekil 52'de verilen RDF dosyaları ve TURCOVID19 Ontolojisi select kaynaklar files yolu ile seçilir ve Apache Jena Fuseki sunucusuna yüklenir.



Şekil 52. Dosya yükleme

Seçilen dosyalar Apache Jena Fuseki'ye yükledikten sonra Şekil 53'de verilen SPARQL sorgusu doğrudan çalıştırılır. Böylece, Protégé Tool ve Jena Fuseki kullanmadan

sadece SPARQL sorgusunun sonucu gösterilir, aynı zamanda sorgular düzenlenebilir ve silinebilir.



```

8 Prefix turcovid19-ontology: <https://www.gulenkara.com/ontology/turcovid19-ontology#>
9 Select ?a ?benderGrup ?statistics
10 Where {
11 ?a rdfs:subPropertyOf turcovid19-ontology:hasStatistics-AgeandbenderGrup.
12 ?a rdfs:subClassOf ?benderGrup.
13 ?a rdfs:domain ?statistics.
14 }
15 }
  
```

| # | GenderGroup | statistics |
|---|--|--|
| 1 | <https://www.gulenkara.com/ontology/turcovid19-ontology#E_25-49> MaleAgeGroup | <https://www.gulenkara.com/ontology/turcovid19-ontology#MortalityRateInConfirmedCases> |
| 2 | <https://www.gulenkara.com/ontology/turcovid19-ontology#E_25-49> MaleAgeGroup | <https://www.gulenkara.com/ontology/turcovid19-ontology#NumberOfPatients1> |
| 3 | <https://www.gulenkara.com/ontology/turcovid19-ontology#E_25-49> MaleAgeGroup | <https://www.gulenkara.com/ontology/turcovid19-ontology#NumberOfPatients1-100K> |
| 4 | <https://www.gulenkara.com/ontology/turcovid19-ontology#E_25-49> MaleAgeGroup | <https://www.gulenkara.com/ontology/turcovid19-ontology#NumberOfPatients1M> |
| 5 | <https://www.gulenkara.com/ontology/turcovid19-ontology#E_25-49> MaleAgeGroup | <https://www.gulenkara.com/ontology/turcovid19-ontology#NumberOfFatalDecases1> |
| 6 | <https://www.gulenkara.com/ontology/turcovid19-ontology#E_50-64> MaleAgeGroup | <https://www.gulenkara.com/ontology/turcovid19-ontology#MortalityRateInConfirmedCases> |
| 7 | <https://www.gulenkara.com/ontology/turcovid19-ontology#E_65-74> MaleAgeGroup | <https://www.gulenkara.com/ontology/turcovid19-ontology#NumberOfPatients1> |

Şekil 53. SPARQL sorgu dili ile sorgulama

Apache Jena Fuseki’de birkaç SPARQL sorgusu örneği vardır. Şekil 54’de yüklenen .ttl ve .owl uzantılı dosyaların subject (özne), predicate (yüklem) ve object (nesne) üçlüleri olacak şekilde sorgu sonucunda gösterilmiştir.



```

1 SELECT ?subject ?predicate ?object
2 WHERE {
3 ?subject ?predicate ?object
4 }
5
  
```

| # | subject | predicate | object |
|----|---|---|---|
| 51 | <https://www.gulenkara.com/ontology/turcovid19-ontology#numberOfDailyFirstDose> | <http://www.w3.org/2000/01/rdf-schema#comment> | Görüş 1. Doz |
| 52 | <https://www.gulenkara.com/ontology/turcovid19-ontology#numberOfDailyFirstDose> | <http://www.w3.org/1999/02/22-rdf-syntax-ns#type> | <http://www.w3.org/2002/07/owl#DatatypeProperty> |
| 53 | <https://www.gulenkara.com/ontology/turcovid19-ontology#numberOfDailyFirstDose> | <http://www.w3.org/2000/01/rdf-schema#range> | <http://www.w3.org/2001/XMLSchema#integer> |
| 54 | <https://www.w3.org/ns/rdf-test#Male> | <http://www.w3.org/1999/02/22-rdf-syntax-ns#type> | <https://www.gulenkara.com/ontology/turcovid19-ontology#MaleAgeGroup> |

Şekil 54. RDF üçlüsü sorgulama sonucu

3. BULGULAR VE İRDELEMELER

Birçok kuruluş büyük hacimli heterojen sağlık hizmetleri açık veri kümeleri yayınlamaktadır. Günümüzde, birçok COVID-19 veri kümeleri, insan tarafından okunabilen tescilli formatlarda mevcuttur ve erişilebilir durumdadır. COVID-19 hastalığının dünya üzerindeki yayılımının artmasından dolayı da sağlık problemlerinde sıklıkla konumsal analizlere yer verilmektedir. TURCOVID19 platformu üzerinden yayınlanan açık COVID-19 verilerinin semantik tanımlarının oluşturulması, sorgulanması ve bağlantılı veri olarak yayınlanması hedeflenmiştir.

3.1. COVID-19 Verilerinin Sorgulanması ve Karşılaştırılması

Anlamsal Web, verinin anlamını sunan bilginin kullanılmasını hedefleyen ve bilginin kullanıldığı sistemlerin birbirlerine bağlandığı günümüzdeki verinin genişletilmesi olarak tanımlanabilmektedir. Çalışmada COVID-19 alanı içinde tanımlanan ontolojilerin ve RDF verilerinin üçlüler olarak Fuseki veritabanına yüklenmesi, saklanması ve sorgulanması üzerine değerlendirmeler yapılmıştır.

Elektronik tablolar karmaşık veri manipülasyonuna ihtiyaç duymayan tek veya az sayıda kullanıcı için uygun bir seçenektir. Fakat pandemi döneminde COVID-19'a ait veri hacmi gün geçtikçe büyümektedir. Elektronik tablolar üzerinden yapılan sorgular ile de excell tablosu uyumluluk sorunları, veri boyutunun büyük olması ile karmaşıklığa ve yavaşlamaya neden olması, birden fazla dosyanın birleştirilmesiyle oluşan veri tekrarı, veri formatlarının farklı olması gibi dezavantajlar ortaya çıkmıştır.

Şekil 55'de verilen SPARQL sorgusu ile günlük 1. Doz aşılarının sayıları küçükten büyüğe doğru sıralanmıştır. Ek olarak, aynı tarihte yapılan 2. Doz aşısı sayıları da sorguya eklenmiş olup hızlı bir şekilde kıyaslama yapılmaktadır.

```

13 Select ?city ?dailyfirstdose ?date ?dailyseconddose
14 Where {
15   ?a turcovid19-ontology:vaccined_on ?date.
16   ?a turcovid19-ontology:numberOfDailyFirstDose ?dailyfirstdose.
17   ?a turcovid19-ontology:cityName ?city.
18   ?a turcovid19-ontology:numberOfDailySecondDose ?dailyseconddose
19   FILTER ( ?city="Ardahan" )
20 }
21 ORDER BY ( ?dailyfirstdose )
22

```

Table Response 105 results in 0.045 seconds Simple view Ellipsis Filter query results Page size: 50

| city | dailyfirstdose | date | dailyseconddose |
|------------|----------------|---------------------|-----------------|
| 1 Ardahan | 0 | 20.02.2021 23:12:58 | 1 |
| 2 Ardahan | 0 | 21.02.2021 23:13:59 | 3 |
| 3 Ardahan | 0 | 20.03.2021 23:13:11 | 0 |
| 4 Ardahan | 0 | 27.03.2021 23:12:37 | 0 |
| 5 Ardahan | 0 | 13.03.2021 23:15:43 | 0 |
| 6 Ardahan | 0 | 06.03.2021 23:15:59 | 0 |
| 99 Ardahan | 8 | 28.03.2021 23:12:43 | 70 |
| 7 Ardahan | 10 | 21.03.2021 23:13:28 | 72 |
| 24 Ardahan | 14 | 18.04.2021 23:40:56 | 1 |
| 36 Ardahan | 18 | 03.04.2021 23:15:42 | 4 |
| 41 Ardahan | 19 | 14.02.2021 23:13:31 | 6 |

Şekil 55. Ardahan iline ait günlük 1. ve 2. doz sayılarını gösteren SPARQL sorgusu

Aynı sorgu işlemi verilere ait Excel dosyası üzerinde yapılmıştır. Şekil 56'da sadece istenilen değişkenlere ulaşılamadığı ve bu nedenle karmaşıklığa neden olduğu mevcut verilerden çıkarsamanın zorlaştığı görülmektedir.

| A | B | C | D | E | F | G | H | I | J | K | L |
|---------|---------------|---------------|---------------|---------------|----------------------|----------------------|----------------------|----------------------|---|---|---|
| Tarih | Toplam 1. Doz | Toplam 2. Doz | Toplam 3. Doz | Toplam 4. Doz | Günlük Toplam 1. Doz | Günlük Toplam 2. Doz | Günlük Toplam 3. Doz | Günlük Toplam 4. Doz | | | |
| Ardahan | 7.256 | 5.207 | 1.945 | | | | | | | | |
| Ardahan | 7.253 | 5.207 | 1.945 | | | | | | | | |
| Ardahan | 1.996 | 1.996 | | | | | | | | | |
| Ardahan | 1.996 | 1.996 | | | | | | | | | |
| Ardahan | 2.820 | 2.820 | | | | | | | | | |
| Ardahan | 2.820 | 2.820 | | | | | | | | | |
| Ardahan | 11.944 | 9.381 | 2.957 | | | | | | | | |
| Ardahan | 13.040 | 9.951 | 3.989 | | | | | | | | |
| Ardahan | 15.697 | 10.217 | 5.480 | | | | | | | | |
| Ardahan | 18.112 | 10.573 | 7.536 | | | | | | | | |
| Ardahan | 77.583 | 46.716 | 25.347 | 5.521 | | | | | | | |
| Ardahan | 18.190 | 10.591 | 7.689 | | | | | | | | |
| Ardahan | 15.779 | 10.227 | 5.552 | | | | | | | | |
| Ardahan | 24.907 | 15.196 | 9.791 | | | | | | | | |
| Ardahan | 77.690 | 46.732 | 25.428 | 5.520 | | | | | | | |
| Ardahan | 20.614 | 11.897 | 8.717 | | | | | | | | |
| Ardahan | 90.403 | 50.052 | 32.499 | 7.542 | | | | | | | |
| Ardahan | 3.860 | 3.451 | 499 | | | | | | | | |
| Ardahan | 15.697 | 10.217 | 5.480 | | | | | | | | |
| Ardahan | 13.123 | 9.971 | 5.152 | | | | | | | | |

Şekil 56. Ardahan iline ait günlük 1. ve 2. doz sayılarını gösteren Excel sorgu sonucu

Apache Jena Fuseki ile bütün anlamsal veriler tek bir veritabanında saklanmaktadır. COVID-19 RDF verileri ile geliştirilen TURCOVID19 ontolojisi veritabanında saklanır ve hem ontolojinin hem de ontoloji yardımıyla anlamsal verilerin SPARQL sorguları ile sorgulanması gerçekleşir. Ayrıca aynı anda birden fazla kullanıcının hızlı ve güvenli bir şekilde verilere erişmesine ve verilerin sorgulanmasına olanak sağlar.

Şekil 50’de verilen SPARQL sorgusu Ege Bölgesine ait istatistikleri vermektedir. Sorguda bölgenin adı, bölgeye ait 16.08.2020 tarihinde açıklanan istatistiklerin vaka kayıtları listelenmektedir.

```
Select ?istatistik ?vakakaydı ?tarih ?ad
Where {
  ?a rdfs:subPropertyOf turcovid19-ontology:hasStatistics-Region.
  ?a rdfs:subClassOf turcovid19-ontology:Region.
  ?a rdfs:domain ?istatistik.
  ?x rdfs:domain ?b.
  ?b rdfs:comment ?ad.
  ?x rdfs:range ?istatistik.
  ?c turcovid19-ontology:TR3 ?vakakaydı.
  ?c rdf:type ?istatistik.
  ?c codo:covid-19caseOn ?tarih.
  FILTER (?tarih="16.08.2020")
  FILTER (?a=turcovid19-ontology:TR3)
  FILTER (?ad="Ege Aegean Region (TR3)")
}
```

Herhangi bir İBBS-1 bölgesine ait yeni hasta sayısı, yeni hastane yatışları, yeni taburcu edilenler, toplam hasta sayısı, 100.000 nüfusta hasta sayısı, son 7 günlük hasta sayısı gibi veriler Excel formatta farklı dosyalarda yer almaktadır. Şekil 57’de verilen örnek sorguda farklı dosyalarda yer alan verilere SPARQL sorgusu ile aynı anda hızlı bir şekilde ulaşılmıştır.

Aynı sorgu işlemi Apache Jena Fuseki üzerinde denenmiştir. 02.04.2022 ile 05.06.2022 tarihleri arasında toplam ve günlük olarak hasta ve ölüm sayılarına ulaşım hızlı bir şekilde sağlanmıştır.

```

4 PREFIX tur:covid19-ontology: <https://www.isibang.ac.in/ns/covid19-ontology/#>
5 SELECT ?mydate ?newcases ?totalcases ?newdeaths ?totaldeaths
6 WHERE {?node codo:covid-19case?n ?mydate.
7   ?node rdfs:cases ?totalcases.
8   ?node rdfs:dailyIncreaseCases ?newcases.
9   ?node rdfs:deceased ?totaldeaths.
10  ?node tur:covid19-ontology:numberOfDailyDeceased ?newdeaths.
11  FILTER (?mydate > "4/1/2022" ^^
12  ?mydate <= "5/16/2022").
13
14  } ORDER BY ?mydate
15

```

| mydate | newcases | totalcases | newdeaths | totaldeaths |
|--------------|----------|------------|-----------|-------------|
| 1 4/10/2022 | 5,609 | 14,918,974 | 32 | 98,409 |
| 2 4/11/2022 | 6,893 | 14,965,867 | 28 | 98,437 |
| 3 4/12/2022 | 6,635 | 14,972,502 | 25 | 98,462 |
| 4 4/13/2022 | 5,529 | 14,978,031 | 31 | 98,493 |
| 5 4/14/2022 | 5,127 | 14,983,158 | 21 | 98,514 |
| 6 4/15/2022 | 4,425 | 14,987,583 | 18 | 98,532 |
| 7 4/16/2022 | 4,086 | 14,991,669 | 13 | 98,551 |
| 8 4/17/2022 | 3,268 | 14,994,937 | 17 | 98,568 |
| 9 4/18/2022 | 4,542 | 14,999,479 | 13 | 98,587 |
| 10 4/19/2022 | 4,217 | 15,003,696 | 21 | 98,610 |

Şekil 59. SPARQL sorgusu sonucu

İlk satırda sorgu içinde kullanılan Prefix, ön ek (namespace) tanımını içermektedir. Sorgu içinde birden çok fazla kısaltma kullanıldığı için, her biri sorgunun en üstünde bulunan bölüme yeni bir satır olarak eklenmiştir. Sorguda kullanılan Prefixler aşağıda verilmiştir. 5. satır içinde geçen SELECT sorgu türünü ve ?mydate, ?newcases, ?totalcases, ?newdeaths ve ?totaldeaths ifadeleri, aranan çizge düğümlerini göstermektedir. Aranan çizge düğümleri birer değişkendir ve soru işareti ile başladıkları sürece farklı kelimeler ile kullanılabilirler. 7. ve 10. satırlar arasında kalan kısım, aranan değişkenlerin sağlanması gereken çizge üçlülerini göstermektedir. Hepsini sağlayan değerler cevap olarak dönecektir. 11. satırda verilen FILTER ile istenilen tarihler arasında sorgu sonucunda filtreleme işlemi yapılmıştır. 14.satırda verilen ORDER BY ile sorgu sonucu filtrelenen tarihlere göre sıralanmıştır. Şekil 59'da verilen SPARQL sorgusu aşağıda verilmiştir.

Prefix xsd: <http://www.w3.org/2001/XMLSchema#>

Prefix rdfs: <http://www.w3.org/2000/01/rdf-schema#>

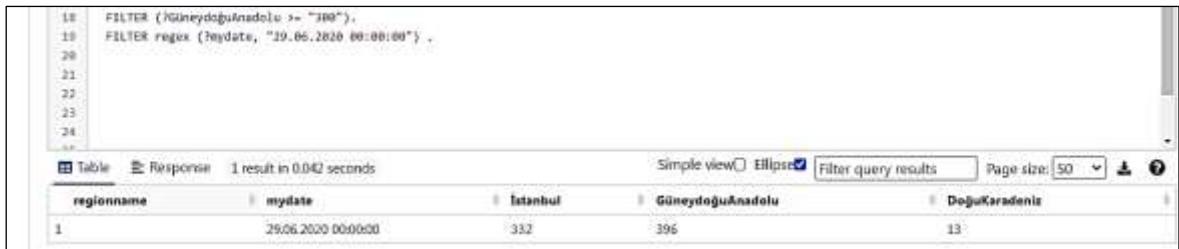
Prefix codo:<http://www.isibang.ac.in/ns/codo#>

Prefix turcovid19-ontology: <https://www.gultenkara.com/ontology/turcovid19-ontology#>

```
SELECT ?mydate ?newcases ?totalcases ?newdeaths ?totaldeaths
WHERE {
?node codo:covid-19caseOn ?mydate.
?node codo:cases ?totalcases.
?node codo:dailyIncreasedCases ?newcases.
?node codo:deceased ?totaldeaths.
?node turcovid19-ontology:numberOfDailyDeceased ?newdeaths.
FILTER (?mydate > "4/1/2022" &&
?mydate <= "5/16/2022").
} ORDER BY ?mydate
```

SPARQL sorguları, sorguları cevaplamak için geliştirilmiş olan ve genel olarak tüm üçlü veritabanlarında bulunan bir birim tarafından cevaplanmaktadır. Bu birim SPARQL Uç Birimi (SPARQL Endpoint) olarak adlandırılır. Uç birimler farklı şekillerde sorgulanabilmektedir. Bu yöntemler arasında, HTTP protokolü kullanılarak web üzerinden sorgulama, Uygulama Programlama Arayüzü (API) üzerinden gönderilen sorgulama ve Komut Satırı kullanılarak yapılan sorgulamalar sayılabilir. Birçok üçlü veritabanı, aynı anda bu yöntemlerin birden fazlasını sağlayabilmektedir.

Şekil 60'da verilen SPARQL sorgusu Güneydoğu Anadolu Bölgesi'nde yeni hasta sayısının 300'den fazla olduğu tarihi ve o tarihte diğer bölgelerin yeni hasta sayılarını göstermektedir. Hem Excel sorgusu hem de Fuseki üzerinden yapılan SPARQL Sorgusu ile aynı sonuca ulaşılmaktadır.



| regionname | mydate | İstanbul | GüneydoğuAnadolu | DoğuKaradeniz |
|------------|---------------------|----------|------------------|---------------|
| 1 | 19.06.2020 00:00:00 | 332 | 396 | 13 |

Şekil 60. Bölgelere göre yeni hasta sayısı SPARQL sorgusu sonucu

Şekil 61'de verildiği gibi Excel üzerinden yapılan sorguda da aynı sonuca ulaşılması, yeni hasta verilerinin tek bir Excel tablosunda yer almasından kaynaklıdır. Bu nedenle,

birden fazla dosyanın birleştirilmesi ile kaynaklanan Excel tablosu uyumluluk sorunları ile karşılaşmamıştır. Ancak Fuseki de farklı dosyalardaki veriler aynı anda sorgulanmaktadır. Excel’de yaşanan problemler, Fuseki üzerinde yapılan SPARQL sorguları sonucunda karşılaşılmadığı tespit edilmiştir.

| Tarih | İstanbul | Batı Marmara | Ege | Doğu Marmara | Batı Anadolu | Akdeniz | Orta Anadolu | Batı Karadeniz | Doğu Karadeniz | Kuzeydoğu Anadolu | Orta Doğu Anadolu | Güneydoğu Anadolu |
|---------------------|----------|--------------|-----|--------------|--------------|---------|--------------|----------------|----------------|-------------------|-------------------|-------------------|
| 29.06.2020 00:00:00 | 332 | 5 | 89 | 131 | 228 | 69 | 24 | 34 | 13 | 9 | 49 | 396 |

Şekil 61. Excel’de bölgelere göre yeni hasta sayısı sorgu sonucu

3.2. Triple Store/RDF Store ile İlişkisel Veritabanlarının Karşılaştırılması

RDF, orijinal olarak veri kaynaklarının meta verilerini tanımlamak için geliştirilmiş, bilgilerin tanımlanması için bir metodoloji olarak geliştirilmiştir. Günümüzde ise, sıklıkla Semantik Web ile ilişkili olarak ve aynı zamanda diğer uygulamalarda verilerin web üzerinde yayınlanması için standart olarak kullanılmaktadır. RDF modeli, bilgiyi özne-yüklem-nesne şeklinde üçlü olarak temsil etmektedir. Bu şekilde temsil edilen RDF üçlülerini depolayabilen, işleyebilen ve sorgulayabilen veri tabanı yönetim sistemleri RDF Store ya da Triple Store olarak adlandırılmaktadır.

RDF depoları, yüklemi özne ve nesne arasındaki bir ilişki olarak yorumlayarak, Graph DBMS'nin bir alt sınıfı olarak değerlendirilebilir. Ancak, RDF depoları, genel Graph DBMS ötesine geçen özel yöntemler sunmaktadır. Örneğin, RDF verileri için SQL benzeri bir sorgu dili olan SPARQL, RDF depoları tarafından desteklenir. Verilerin yapısal olarak saklanması ve sorgulanmasını sağlamak amacı ile özelleştirilmiş veritabanı sistemleri olarak Triple Store’lar kullanılmaktadır. En yaygın örnekleri MarkLogic, Virtuoso, Amazon Neptune, Apache Jena – TDB, GraphDB, AllegroGraph, RDFox, Sesame, Bigdata(R), YARS2 ve Oracle Spatial and Graph’dır. İlişkisel veritabanları, verilerin boyutunun büyük olması halinde ve ilişkisel veritabanları çizge verilerini depolamak amacı ile özelleştirilmediği için, sorguların cevaplanmasının çok uzun zaman almasına sebep olmaktadır. Ayrıca OWL kullanılarak üçlü veritabanları, ilişkisel veritabanlarının sağlayamadığı çıkarsama işlemlerini de yapabilmektedir.

Veri kümelerinin birbirlerinden bağımsız olarak oluşturulmaları sadece sözdizimsel düzeyde birlikte çalışılabilirlik problemini ortaya çıkarmaktadır. İlişkisel veri tabanları içinde bulundurduğu verilerin anlamı hakkında bilgi saklamamaktadır. Veriler arasında anlamsal ilişkiler tanımlı değildir ve anlam farklılıkları oluşmaktadır. Bu nedenle anlamsal çıkarsamayı destekleyen bir yapıya sahip olmadıkları tespit edilmiştir. İlişkisel veritabanlarının aksine Triple Store'lar, ontoloji adı verilen isteğe bağlı şema modellerini desteklemektedir. Ontolojiler, verilerin resmi bir açıklamasına izin verir. Hem nesne sınıflarını hem de ilişki özelliklerini ve bunların hiyerarşik sırasını belirtmektedirler. Eksiklikler dikkate alınarak veri kümeleri arasında anlamsal ilişkilerin tanımlanması için TURCOVID19 ontolojisi geliştirilmiştir. Ontoloji, TURCOVID19 platformu tarafından COVID-19 alanında paylaşılan bilgilerin anlamsal birlikte çalışabilirlik kapsamında yapısal olarak bilgi gösterimini sağlayacaktır.

Mevcut veritabanı sorgulama dillerinin çıkarsama ve sorgulama yeteneklerinin olmadığı tespit edilmiştir. Ontoloji, sınıflandırma ve hiyerarşik yapısı sayesinde kendi içerisinde de çıkarımları sağlayacak şekilde tasarlanmıştır. Böylelikle ontolojide tutulan COVID-19 açık verilerden çıkarsama ile yeni bilgiler elde edilmektedir.

Triple Store'lar ayrıca yapılandırılmamış verilerden bilgi çıkarsama ve içeriği zenginleştirme gibi birçok metin analizi tekniği de kolaylaştırmaktadır. Sağlık hizmetlerinde Webde bulunan bilgiler dağınık ve farklı veritabanlarında farklı formatlarda bulunmaktadır. Triple Store'lar sayesinde varlıkların kullanıldığı bağlamın yanı sıra anlamı da öğrenerek varlıkları sınıflandırabilir ve aralarındaki belirsizliği de giderilmektedir. Triple Store olarak kullanılan Fuseki ile RDF verileri ve ontolojiler de sorgulanabilmektedir. Ayrıca, anlamsal ilişkileri bulabilmek için geleneksel ilişkisel ve anlamsal veri, ontoloji yardımıyla sorgulanabilmektedir.

SPARQL Server ile;

- Ontolojiler ve anlamsal veriler saklanır,
- Anlamsal veri sorgulanır,
- Ontoloji yardımı ile sorgular yapılır,
- Kullanıcı tanımlı veya desteklenen çıkarsamanın kullanılmasıyla anlamsal veri üzerinde sorgulamanın gücü artar.

Çıkarsama, veri ve kurallara dayanarak mantıksal sonuçlar çıkartmayı sağlamaktadır. Şekil 62'de verilen SPARQL sorgusu cinsiyet ve yaş gruplarına ait ölüm oranlarını vermektedir. RDF verileri, ontoloji yardımıyla sorgulanabilmektedir. Şekil 55'te görüldüğü

gibi ?a değişkeni 50-64 yaş grubuna sahip ontolojideki FemaleAgeGroup sınıfının her bir tarihe göre ölüm oranları sorgu sonucunda gösterilmiştir.



```

1 Prefix codo: <http://www.isihang.ac.in/ns/codo#>
2 Prefix turcovid19-ontology: <https://www.gulbenkara.com/ontology/turcovid19-ontology#>
3 Select ?a ?GenderGroup ?statistics ?a ?date
4 Where {
5   ?a rdfs:subPropertyOf turcovid19-ontology:hasStatistics-AgeandGenderGroup.
6   ?a rdfs:subClassOf ?GenderGroup.
7   ?a rdfs:domain ?statistics.
8   ?a rdfs:type turcovid19-ontology:MortalityRateInConfirmedCases.
9   ?a turcovid19-ontology:K_50-64 ?a.
10  ?a codo:co-50-64case00 ?date.
11 }
12 #filter (?a=<https://www.gulbenkara.com/ontology/turcovid19-ontology#K_50-64>)
13 }

```

| | GenderGroup | statistics | date |
|---|--|---|-----------------|
| 1 | <https://www.gulbenkara.com/ontology/turcovid19-ontology#FemaleAgeGroup> | <https://www.gulbenkara.com/ontology/turcovid19-ontology#MortalityRateInConfirmedCases> | 1.79 05.07.2020 |
| 2 | <https://www.gulbenkara.com/ontology/turcovid19-ontology#FemaleAgeGroup> | <https://www.gulbenkara.com/ontology/turcovid19-ontology#MortalityRateInConfirmedCases> | 1.81 25.10.2020 |
| 3 | <https://www.gulbenkara.com/ontology/turcovid19-ontology#FemaleAgeGroup> | <https://www.gulbenkara.com/ontology/turcovid19-ontology#MortalityRateInConfirmedCases> | 1.73 02.08.2020 |

Şekil 62. Yaş ve cinsiyet gruplarına ait ölüm oranları sorgu sonucu

3.3. TURCOVID19 Verileri İçin Belirlenen Ontolojilerin İrdelenmesi

Web’de bulunan COVID-19 veri kümelerinin entegrasyonunu, araştırılmasını ve analizini kolaylaştırmak için anlamsal olarak yapılandırılmış, birbirine bağlı ve bilgisayarlar tarafından okunabilir şekilde biçimlendirilmeleri gerekmektedir. Bu nedenle heterojen biçimlerde sağlanan COVID-19 verileri temsili bir veri modeli ile ortak bir şemada ilişkilendirilmeleri sağlanmalıdır. Açık veri kaynaklarında bulunan COVID-19 verilerinin semantik tanımlarının oluşturulması ile SWT kullanılarak bütün COVID-19 verilerinin sorgu sonucunda döndürülecektir. Bu nedenle TURCOVID19 platformundan indirilen COVID-19 verilerinin semantik tanımlarının oluşturulması gerekir. Semantik tanımların oluşturulması için temel gereksinim ise mevcut COVID-19 ile ilgili ontolojilerin belirlenmesidir. Ontoloji geliştirme süreci için de, bir yandan mevcut verilerle, diğer yandan da ontoloji tarafından cevaplanabilecek ilgi çekici sorularla yönlendirilen bir yaklaşım izlenerek TURCOVID19 ontolojisinin geliştirilmesidir.

CIDO, COVID-19 ve diğer hastalıkların karşılaştırmalı analizini semptomlar, ilaçlar, klinik araştırmaları vb. olarak temsil etmektedir. En ilgili çalışmalardan biri, CIDO’nun geliştirilmesini başlatan CODO (Dutta ve DeBellis, 2016), FAIR (Wilkinson 2016) ilkelerine dayalı olarak COVID-19 vakaları ve coğrafi konuma göre vakaların istatistikleri

olarak bir bilgi grafiği (knowledge graph) modeli olarak oluşturulmuştur. Web'de bulunan daha birçok COVID-19 veri setini entegre ederek üretilen verilerin semantik olarak sorgulanmasını sağlamaktadır. Bu nedenle, çalışmanın amacına uygun yeniden uygulanabilir bir kelime hazinesine sahip olduğu ve mevcut CODO'nun genişletilmesine uygun olduğu tespit edilmiştir. CODO'nun genişletilmesiyle oluşturulan TURCOVID19 ontolojisinde 282 classes, 3271 axioms, 108 object properties, 142 data properties ve 57 individuals bulunmaktadır. Ayrıca, İBBS bölgelerine ait konumsal ilişkileri tanımlamak amacı ile GeoSPARQL Ontolojisi kullanılmıştır. Konumsal verilerin geometri tiplerini ve topolojik ilişkileri tanımlamak için GeoSPARQL Ontolojisi kullanılmaktadır. COVID-19'a ait istatistiklerin konumsal bazlı olarak yayınlanması ile konuma göre çıkarsamaların yapılması sağlanmaktadır. Bu nedenle TURCOVID19 Ontolojisi topolojik çıkarsama imkânı veren GeoSPARQL Ontolojisi ile ilişkilendirilmiştir. COVID-19'a ait bilgilerin konumsal bazlı analizi, salgının izlenmesini ve veri keşfini kolaylaştırmaktadır.

Doğu Karadeniz Bölgesi'ne ait istatistikler ve bölgenin kapsadığı illeri gösteren SPARQL sorgu satırları aşağıda verilmiştir. Sorgu, bölgeye ait istatistikler, her bir istatistiğe göre vaka kaydı, vakaların bulunduğu bölgenin kapsadığı alt bölge ve illeri listeler.

```
Select ?statistics ?date ?cases ?regionname ?subregionname ?provincename
Where {
  ?a rdfs:comment ?provincename.
  ?a rdfs:subClassOf ?subregion.
  ?subregion rdfs:subClassOf ?region.
  ?subregion rdfs:comment ?subregionname.
  ?region rdfs:comment ?regionname.
  ?x rdfs:subPropertyOf codo:countryWiseStatistics .
  ?x rdfs:domain ?region.
  ?x rdfs:range ?statistics.
  ?c turcovid19-ontology:TR9 ?cases.
  ?c rdf:type ?statistics.
  ?c codo:covid-19caseOn ?date.
  FILTER (?regionname="Doğu Karadeniz East Black Sea Region (TR9)")
}
```

Şekil 63’de görüldüğü gibi İBBS-1 bölgelerine ait ülkemizde yayınlanan hangi COVID-19 vaka istatistiklerinin olduğu ve bu bölgelerin hangi alt bölge ve illeri kapsadığı sonucuna hızlıca ulaşılmaktadır.

| statistics | date | cases | regionname | subregionname | provincename |
|--|-----------|-------|----------------------------------|-------------------------------------|---------------------------|
| <https://www.guibenkara.com/ontology/turcovid19-ontology#NumberOfTotalPatients-100K> | 16.08.2.. | 113.7 | Doğu Karadeniz East Black Sea... | Trabzon alt bölgesi Trabzon Subr... | Trabzon Province (TR9... |
| <https://www.guibenkara.com/ontology/turcovid19-ontology#NumberOfTotalPatients-100K> | 16.08.2.. | 113.7 | Doğu Karadeniz East Black Sea... | Trabzon alt bölgesi Trabzon Subr... | Rize Province (TR904) |
| <https://www.guibenkara.com/ontology/turcovid19-ontology#NumberOfTotalPatients-100K> | 16.08.2.. | 113.7 | Doğu Karadeniz East Black Sea... | Trabzon alt bölgesi Trabzon Subr... | Ordu Province (TR902) |
| <https://www.guibenkara.com/ontology/turcovid19-ontology#NumberOfTotalPatients-100K> | 16.08.2.. | 113.7 | Doğu Karadeniz East Black Sea... | Trabzon alt bölgesi Trabzon Subr... | Gümüşhane Province (...) |
| <https://www.guibenkara.com/ontology/turcovid19-ontology#NumberOfTotalPatients-100K> | 16.08.2.. | 113.7 | Doğu Karadeniz East Black Sea... | Trabzon alt bölgesi Trabzon Subr... | Giresun Province (TR9... |
| <https://www.guibenkara.com/ontology/turcovid19-ontology#NumberOfTotalPatients-100K> | 16.08.2.. | 113.7 | Doğu Karadeniz East Black Sea... | Trabzon alt bölgesi Trabzon Subr... | Artvin Province (TR905) |
| <https://www.guibenkara.com/ontology/turcovid19-ontology#NumberOfTotalPatients> | 16.08.2.. | 3.058 | Doğu Karadeniz East Black Sea... | Trabzon alt bölgesi Trabzon Subr... | Trabzon Province (TR9... |
| <https://www.guibenkara.com/ontology/turcovid19-ontology#NumberOfTotalPatients> | 16.08.2.. | 3.058 | Doğu Karadeniz East Black Sea... | Trabzon alt bölgesi Trabzon Subr... | Rize Province (TR904) |
| <https://www.guibenkara.com/ontology/turcovid19-ontology#NumberOfTotalPatients> | 16.08.2.. | 3.058 | Doğu Karadeniz East Black Sea... | Trabzon alt bölgesi Trabzon Subr... | Ordu Province (TR902) |
| <https://www.guibenkara.com/ontology/turcovid19-ontology#NumberOfTotalPatients> | 16.08.2.. | 3.058 | Doğu Karadeniz East Black Sea... | Trabzon alt bölgesi Trabzon Subr... | Gümüşhane Province (...) |
| <https://www.guibenkara.com/ontology/turcovid19-ontology#NumberOfTotalPatients> | 16.08.2.. | 3.058 | Doğu Karadeniz East Black Sea... | Trabzon alt bölgesi Trabzon Subr... | Giresun Province (TR90... |

Şekil 63. Doğu Karadeniz Bölgesi’ne ait istatistikler ve kapsadığı alt bölge ve illeri gösteren sorgu sonucu

Ontolojilerin sağlık alanında kullanımının artmasıyla daha akıllı sağlık hizmetleri verilebilir olduğu tespit edilmiştir. COVID-19 verilerinin eşleştirilebilir, paylaşılabilir olması ve yapısal olarak saklanması sonucunda bu veriler kullanılarak çıkarsamalar gerçekleştirilecek ve sağlık uygulamalarına katkı sağlayacaktır.

3.4. Mevcut COVID-19 Bağlantılı Verilerin İrdelenmesi

LOD Cloud’da Covid-on-the-Web Dataset, Naturopathy Knowledge Graph (Ontology and Dataset) - RDF distribution of the naturopathy dataset ve COvid-19 statistics Ontology mevcuttur.

Covid-on-the-Web Dataset, RDF dökümü ve SPARQL endpoint olarak DBpedia ve Wikidata veri setlerine bağlantılar kurularak yayınlanmıştır. Çoğunlukla biyomedikal ontolojilere odaklanmaktadır. Bu nedenle, tanımlanan adlandırılmış varlıklar genler, proteinler, ilaçlar, hastalıklar, fenotipler ve yayınlarla ilgilidir. Fakat veri setinin çoğunluğu yapılandırılmamış veriler içermektedir.

Naturopathy Knowledge Graph (Ontology and Dataset), gıdalar, besinler, renkler, ruh haller, hastalıklar, diyetler ve alerjiler arasındaki ilişkileri tanımlamaktadır. Gıdalar, bağışıklık sistemini güçlendirmek ve faydalı olarak tanımlamak için genişletilmiştir. COVID-19 pandemi krizi nedeniyle bağışıklık sistemini gıda ile güçlendirmektedir. DBpedia'ya bağlantı kurularak yayınlanmıştır.

COvid-19 statistics Ontology (whonto), LOD Cloud üzerinde sadece ontoloji olarak yayınlanmıştır. Ontoloji, COVID-19 verilerini Web'de grafik olarak yayınlamak için oluşturulmuş anlamsal modeldir. Coğrafi konum (bölge, eyalet, ülke ve kıta) içindeki vakaların yeni vakalar, toplam aktif, iyileşen, ölen ve bir günde yeni ölen olarak günlük temsil etmektedir. Farklı veri setlerine bağlantı kurulmamıştır.

TURCOVID19 ontolojisinde ise hasta, vefat, iyileşen, uygulanan test sayıları, yoğun bakım yatışı, entübe hasta, aşı, hastane yatışları, taburcu edilenler kategorilerindeki verilerin coğrafi konuma göre istatistikleri ve cinsiyet-yaş grubuna göre de ölüm ve hasta sayıları istatistikleri olarak günlük ve kümülatif şekilde desteklenmesi amaçlanmıştır. WHOnto ve Covid-on-the-Web Veri setlerine bağlantılar kurularak genişletilmiştir. COVID-19 veri setinin çoğunluğu yapılandırılmamış verilerden oluştuğu belirtilmiştir.

NUTS (GeoVocab), NUTS (İstatistik için Bölgesel Birimlerin İsimlendirilmesi), Avrupa Birliği'nin Eurostat ofisi tarafından tanımlanan bir sınıflandırmadır. Burada, GeoData için NeoGeo RDF sözlüğünün örnek bir kullanımı olarak coğrafi bilgilerle birlikte NUTS bölgeleri sunulmuştur.

Tez kapsamında yapılan çalışmada ise açık veriler ontoloji ile ilişkilendirilip yapılandırılmış RDF formatına dönüştürülmüştür. Elde edilen RDF COVID-19 veri seti ile mevcut veri kaynakları arasında SILK ile kurulan bağlantılar ile Linked Open Data (LOD) bulutunda yayınlanacaktır. Ülke karşılaştırmalarına yönelik içeriklerde diğer ülkelere göre farklı kavramlar kullanılmaktadır. Kurulan bağlantılar, Türkiye'ye ait sağlık hizmetleri hakkında, veri setinde bulunmayan bilgileri sağlamak için kullanım senaryoları sağlayacaktır.

3.5. Farklı Veri Setleri ile Oluşturulan Linklerin İrdelenmesi

TURCOVID19 veri seti ilgili LOD veri kaynaklarıyla owl:sameAs ilişkileri kurularak eşleştirilmiştir ve zenginleştirilmiştir. COVID-19 veri setlerindeki veriler

doğrudan bağlantılıdır. Bu nedenle, veri kümelerindeki veriler Şekil 64’de görüldüğü gibi NUTS veri setinde yer alan bölge birimleri isimleri ile eşleştirilerek ilişkilendirilmiştir. Levenshtein Distance karşılaştırma operatörü, iki veri setinde yer alan kavramsal benzerlikleri bulmaktadır. Şekilde görüldüğü üzere bağlantı sonuçları %100 doğrulukla bulunmuştur.



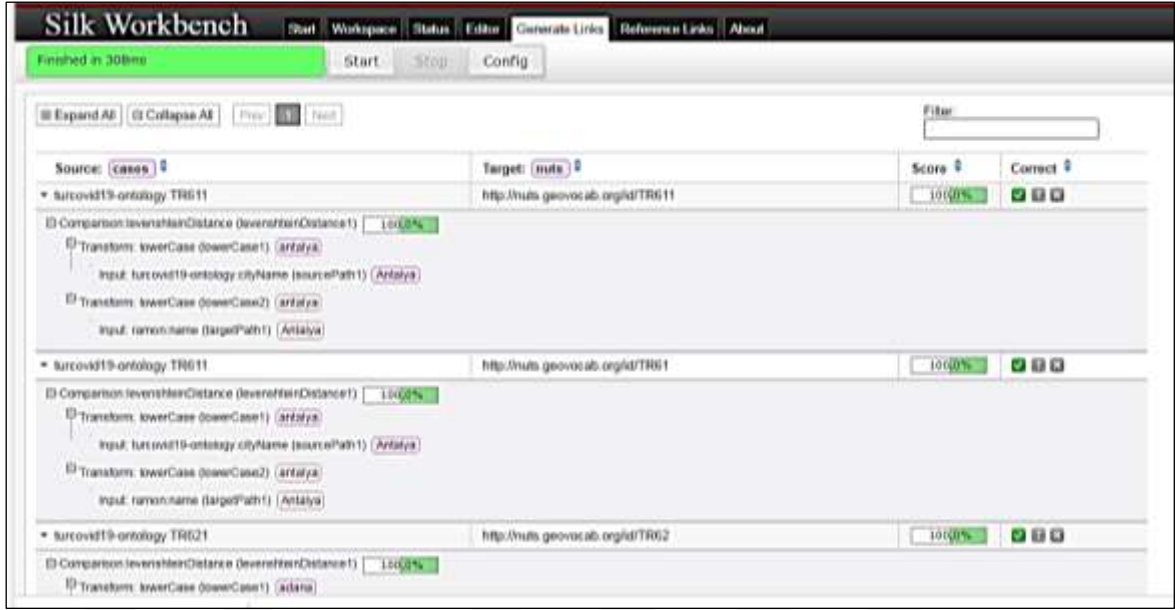
Şekil 64. Haftalık vaka sayıları verilerindeki il isimleri ile NUTS bölge birimleri isimleri arasındaki bağlantı sonuçları

İllere göre haftalık vaka sayıları verilerinde koordinat bilgileri yer almamaktadır. Şekil 65’de görüldüğü owl:sameAs ilişkisi ile oluşturulan linklerle kullanıcılar ek olarak koordinat ve daha fazla bilgiye ulaşacaktır.

| Feature | | | | | | | | | | | | | | | | | | | | | |
|---------------------------------|---|-----------|--------------|---------------------------------|--|----------|------------------|---------|------------------|----------|--------------------|---------|--------------------|----------|--------------------|---------|--------------------|----------|--------------------|---------|------------------|
| rdfs:type | ramon:NUTSRegion_spatial:Feature | | | | | | | | | | | | | | | | | | | | |
| rdfs:label | TR332 - Afyon | | | | | | | | | | | | | | | | | | | | |
| ramon:name | Afyon | | | | | | | | | | | | | | | | | | | | |
| ramon:level | 3 | | | | | | | | | | | | | | | | | | | | |
| ramon:code | TR332 | | | | | | | | | | | | | | | | | | | | |
| ngeo:geometry | nuts:TR332_geometry | | | | | | | | | | | | | | | | | | | | |
| spatial:PP | TR33 | | | | | | | | | | | | | | | | | | | | |
| spatial:EC | http://nuts.geovocab.org/id/TR4 | | | | | | | | | | | | | | | | | | | | |
| spatial:EC | http://nuts.geovocab.org/id/TR6 | | | | | | | | | | | | | | | | | | | | |
| spatial:EC | http://nuts.geovocab.org/id/TR5 | | | | | | | | | | | | | | | | | | | | |
| spatial:EC | http://nuts.geovocab.org/id/TR41 | | | | | | | | | | | | | | | | | | | | |
| spatial:EC | http://nuts.geovocab.org/id/TR52 | | | | | | | | | | | | | | | | | | | | |
| spatial:EC | http://nuts.geovocab.org/id/TR32 | | | | | | | | | | | | | | | | | | | | |
| spatial:EC | http://nuts.geovocab.org/id/TR61 | | | | | | | | | | | | | | | | | | | | |
| spatial:EC | http://nuts.geovocab.org/id/TR412 | | | | | | | | | | | | | | | | | | | | |
| spatial:EC | http://nuts.geovocab.org/id/TR333 | | | | | | | | | | | | | | | | | | | | |
| spatial:EC | http://nuts.geovocab.org/id/TR521 | | | | | | | | | | | | | | | | | | | | |
| spatial:EC | http://nuts.geovocab.org/id/TR334 | | | | | | | | | | | | | | | | | | | | |
| spatial:EC | http://nuts.geovocab.org/id/TR612 | | | | | | | | | | | | | | | | | | | | |
| spatial:EC | http://nuts.geovocab.org/id/TR322 | | | | | | | | | | | | | | | | | | | | |
| spatial:EC | http://nuts.geovocab.org/id/TR613 | | | | | | | | | | | | | | | | | | | | |
| owl:sameAs | http://estatwrap.ontologycentral.com/dic/geo#TR332 | | | | | | | | | | | | | | | | | | | | |
| owl:sameAs | http://rdfdata.eionet.europa.eu/ramon/nuts2008/TR332 | | | | | | | | | | | | | | | | | | | | |
| owl:sameAs | http://ec.europa.eu/eurostat/ramon/rdfdata/nuts2008/TR332 | | | | | | | | | | | | | | | | | | | | |
| owl:sameAs | http://nuts.psi.enacting.org/id/TR332 | | | | | | | | | | | | | | | | | | | | |
| Geometry | | | | | | | | | | | | | | | | | | | | | |
| rdfs:type | ngeo:Polygon | | | | | | | | | | | | | | | | | | | | |
| ngeo:exterior | <table border="1"> <thead> <tr> <th>rdfs:type</th> <th>ngeo:posList</th> </tr> </thead> <tbody> <tr> <td>ngeo:LinearRing</td> <td></td> </tr> <tr> <td>geo:long</td> <td>31.2333493548673</td> </tr> <tr> <td>geo:lat</td> <td>38.4099727699639</td> </tr> <tr> <td>geo:long</td> <td>31.129906476990328</td> </tr> <tr> <td>geo:lat</td> <td>38.490236959035997</td> </tr> <tr> <td>geo:long</td> <td>30.986740882972185</td> </tr> <tr> <td>geo:lat</td> <td>38.466414593197499</td> </tr> <tr> <td>geo:long</td> <td>30.426787947621733</td> </tr> <tr> <td>geo:lat</td> <td>38.1538943794055</td> </tr> </tbody> </table> | rdfs:type | ngeo:posList | ngeo:LinearRing | | geo:long | 31.2333493548673 | geo:lat | 38.4099727699639 | geo:long | 31.129906476990328 | geo:lat | 38.490236959035997 | geo:long | 30.986740882972185 | geo:lat | 38.466414593197499 | geo:long | 30.426787947621733 | geo:lat | 38.1538943794055 |
| rdfs:type | ngeo:posList | | | | | | | | | | | | | | | | | | | | |
| ngeo:LinearRing | | | | | | | | | | | | | | | | | | | | | |
| geo:long | 31.2333493548673 | | | | | | | | | | | | | | | | | | | | |
| geo:lat | 38.4099727699639 | | | | | | | | | | | | | | | | | | | | |
| geo:long | 31.129906476990328 | | | | | | | | | | | | | | | | | | | | |
| geo:lat | 38.490236959035997 | | | | | | | | | | | | | | | | | | | | |
| geo:long | 30.986740882972185 | | | | | | | | | | | | | | | | | | | | |
| geo:lat | 38.466414593197499 | | | | | | | | | | | | | | | | | | | | |
| geo:long | 30.426787947621733 | | | | | | | | | | | | | | | | | | | | |
| geo:lat | 38.1538943794055 | | | | | | | | | | | | | | | | | | | | |

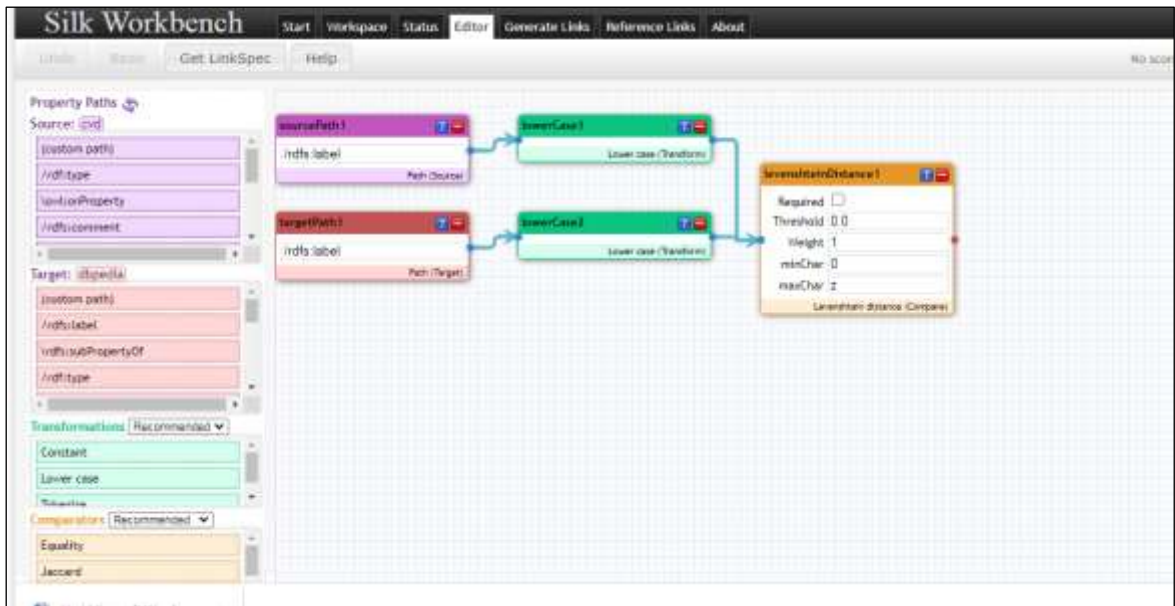
Şekil 65. owl:sameAs ilişkisi ile farklı veri setlerindeki bilgilerin gösterimi

Oluşturulan bağlantılarda şekilde görüldüğü üzere NUTS veri setindeki Antalya alt bölgesi ve Antalya ilinin her ikisi de syntactic (sözdizimsel) eşleştirmeden dolayı TURCOVID19 veri setinde yer alan Antalya ili eşleşmiştir.



Şekil 66. Silk ile belirlenen eşleştirme sonuçları

COVID-19 veri setlerindeki veriler webde bağlantılı veri olarak yer almadığından dolayı TURCOVID19 ontolojisindeki sınıflar ile DBpedia ve Wikidata ontolojilerin de yer alan Covid-19 ve Pandemic sınıflarıyla eşleştirilmiştir. Şekilde görüldüğü gibi property path kısmında rdfs:label ve rdfs:comment özellikleri ile eşleştirme yapılmıştır.



Şekil 67. Link editör ile bağlantıların tanımlanması

| Source: turcovid_ontology | Target: dbpedia | Score | Correct |
|--|------------------|-------|---------|
| turcovid19-ontology-NumberOfTotalDeceased1 | dbpedia:Pandemic | 100% | ✓ |
| turcovid19-ontology-NumberOfTotalDeceased1 | dbpedia:حکومت | 100% | ✓ |
| turcovid19-ontology-NumberOfNewPatients | dbpedia:Pandemic | 100% | ✓ |
| turcovid19-ontology-NumberOfNewPatients | dbpedia:حکومت | 100% | ✓ |
| turcovid19-ontology-NumberOfTotalDeceased | dbpedia:Pandemic | 100% | ✓ |
| turcovid19-ontology-NumberOfTotalDeceased | dbpedia:حکومت | 100% | ✓ |
| turcovid19-ontology-NumberOfPatientsInTheLastDays | dbpedia:Pandemic | 100% | ✓ |
| turcovid19-ontology-NumberOfPatientsInTheLastDays | dbpedia:حکومت | 100% | ✓ |
| turcovid19-ontology-NumberOfPatientsDischargedFromHospital | dbpedia:Pandemic | 100% | ✓ |
| turcovid19-ontology-NumberOfPatientsDischargedFromHospital | dbpedia:حکومت | 100% | ✓ |
| turcovid19-ontology-NumberOfNewPatients1 | dbpedia:Pandemic | 100% | ✓ |
| turcovid19-ontology-NumberOfNewPatients1 | dbpedia:حکومت | 100% | ✓ |
| turcovid19-ontology-NumberOfPatientsInTheLastDays-100K | dbpedia:Pandemic | 100% | ✓ |
| turcovid19-ontology-NumberOfPatientsInTheLastDays-100K | dbpedia:حکومت | 100% | ✓ |
| turcovid19-ontology-NumberOfNewHospitalAdmissions | dbpedia:Pandemic | 100% | ✓ |
| turcovid19-ontology-NumberOfNewHospitalAdmissions | dbpedia:حکومت | 100% | ✓ |

Şekil 68. Dbpedia Pandemic Sınıfı ile TURCOVID19 sınıflarının eşleştirilmesi

DBpedia

About: Pandemic

An Entity of Type: Class, from Named Graph: <http://dbpedia.org/resource/class>, within Data Space: dbpedia.org

Global epidemic of infectious disease

| Property | Value |
|------------------------------------|---|
| rdf:type | • owl:Class |
| rdfs:comment | • Global epidemic of infectious disease (en) |
| rdfs:isDefinedBy | • http://dbpedia.org/ontology/ |
| rdfs:label | • Pandemic (en) |
| rdfs:subClassOf | • owl:Thing |
| skos:isDescribedBy | • data/definitions.ttl |
| owl:wasDerivedFrom | • http://mappings.dbpedia.org/index.php/OntologyClass/Pandemic |
| skos:isDefinedBy | • http://dbpedia.org/ontology/ |

Şekil 69. Dbpedia Pandemic Sınıfı ile TURCOVID19 sınıflarının eşleştirilmesi

| Source: onto | Target: DBpedia2 | Score | Correct |
|--|--|--------|---------|
| turcovid19-ontology NumberOfTotalDeceased1 | http://www.wikidata.org/entity/Q84263196 | 100,0% | ✓ |
| turcovid19-ontology NumberOfNewPatient | http://www.wikidata.org/entity/Q84263196 | 100,0% | ✓ |
| turcovid19-ontology NumberOfTotalDeceased | http://www.wikidata.org/entity/Q84263196 | 100,0% | ✓ |
| turcovid19-ontology NumberOfPatientsInTheLast7Days | http://www.wikidata.org/entity/Q84263196 | 100,0% | ✓ |
| turcovid19-ontology NumberOfPatientsDischargedFromHospital | http://www.wikidata.org/entity/Q84263196 | 100,0% | ✓ |
| turcovid19-ontology NumberOfNewPatients1 | http://www.wikidata.org/entity/Q84263196 | 100,0% | ✓ |
| turcovid19-ontology NumberOfPatientsInTheLast7Days-100K | http://www.wikidata.org/entity/Q84263196 | 100,0% | ✓ |
| turcovid19-ontology NumberOfNewHospitalAdmissions | http://www.wikidata.org/entity/Q84263196 | 100,0% | ✓ |
| turcovid19-ontology PivotalTable-Total | http://www.wikidata.org/entity/Q84263196 | 100,0% | ✓ |
| turcovid19-ontology NumberOfTotalPatients-100K | http://www.wikidata.org/entity/Q84263196 | 100,0% | ✓ |
| turcovid19-ontology MortalityRateInConfirmedCases | http://www.wikidata.org/entity/Q84263196 | 100,0% | ✓ |
| turcovid19-ontology NumberOfNewPatients-1M | http://www.wikidata.org/entity/Q84263196 | 100,0% | ✓ |
| turcovid19-ontology NumberOfTotalDeceased-100K | http://www.wikidata.org/entity/Q84263196 | 100,0% | ✓ |
| turcovid19-ontology NumberOfProvincialLevelCases | http://www.wikidata.org/entity/Q84263196 | 100,0% | ✓ |
| turcovid19-ontology PivotalTable | http://www.wikidata.org/entity/Q84263196 | 100,0% | ✓ |

Şekil 70. Wikidata COVID-19 (Q84263196) sınıfı ile TURCOVID19 sınıflarının eşleştirilmesi

| Language | Label | Description | Also shown as |
|----------|----------|---|--|
| English | COVID-19 | respiratory syndrome and infectious disease in humans, caused by SARS coronavirus 2 | 2019-nCoV acute respiratory disease, coronavirus disease 2019, COVID-19, COVID 19, 2019 novel coronavirus pneumonia, Coronavirus disease 2019, nCoV2019, nCoVD 19, nCoV-19, COVID-2019, seafood market pneumonia, Wuhan pneumonia, 2019-NCP, WuRS, severe acute respiratory syndrome type 2, SARS-CoV-2 infection, 2019 novel coronavirus respiratory syndrome, Wuhan respiratory syndrome, CD-19, Covid-19, COVID, Novel Coronavirus Pneumonia, Severe Acute Respiratory Syndrome Coronavirus 2, SARS-CoV-2 |

Şekil 71. Wikidata COVID-19 (Q84263196) sınıfı

TURCOVID19 ontolojisinde yer alan datapropertyler ile Covid-on-the web ve Dbpedia'da yer data propertyler property pathda gösterilen özelliklere gösterilmiştir. Aynı doğrultuda kavramsal eşleştirme operatörü olan Levenshtein Distance seçilmiştir. Bağlantı oluşturma sonuçları %100 doğrulukla şekilde gösterilmiştir.

| Source | Target | Score | Correct |
|--------------------------------|--------------------|--------|---------|
| turcovid19-ontology endDate | dbo endTime | 100.0% | ✓ |
| turcovid19-ontology startDate | dbo startDate | 100.0% | ✓ |
| codo area | dbo رايون | 100.0% | ✓ |
| codo date | dbo date | 100.0% | ✓ |
| https://schema.org/postalCode | dbo postalCode | 100.0% | ✓ |
| http://xmlns.com/foaf/0.1/name | dbo book | 100.0% | ✓ |
| http://xmlns.com/foaf/0.1/name | dbo name | 100.0% | ✓ |
| codo deceased | dbo pandemicDeaths | 100.0% | ✓ |
| codo deceased | dbo pandemicDeaths | 100.0% | ✓ |
| codo countryCode | dbo countryCode | 100.0% | ✓ |
| https://schema.org/address | dbo address | 100.0% | ✓ |
| codo age | dbo age | 100.0% | ✓ |

Şekil 72. Dbpedia ve TURCOVID19 ontojilerindeki data propertylerin eşleştirilmesi

Covid-on-the-web datasetinde yer alan data propertyler ile TURCOVID19 ontolojisinde yer alan datapropertylerde yer alan rdf:type özelliği ile semantik olarak değil syntactic (sözdizimsel) olarak eşleşmiştir.

| Source | Target | Score | Correct |
|---|--|--------|---------|
| ontology | covidontheweb | 100.0% | ✓ |
| http://www.gulenkara.com/ontology/turcovid19-ontology#hasStatistics-TR3 | http://ns.inria.fr/covid19/property#DeathsEphad | 100.0% | ✓ |
| Comparison levenshteinDistance (levenshteinDistance1) | | 100.0% | |
| Transform: lowerCase (lowerCase1) | http://www.w3.org/2002/07/owl#DatatypeProperty | | |
| Input: rdf:type (sourcePath1) | http://www.w3.org/2002/07/owl#DatatypeProperty | | |
| Transform: lowerCase (lowerCase2) | http://www.w3.org/2002/07/owl#DatatypeProperty | | |
| Input: rdf:type (targetPath1) | http://www.w3.org/2002/07/owl#DatatypeProperty | | |
| http://www.gulenkara.com/ontology/turcovid19-ontology#hasStatistics-TR3 | http://ns.inria.fr/covid19/property#confidence | 100.0% | ✓ |
| Comparison levenshteinDistance (levenshteinDistance1) | | 100.0% | |
| Transform: lowerCase (lowerCase1) | http://www.w3.org/2002/07/owl#DatatypeProperty | | |
| Input: rdf:type (sourcePath1) | http://www.w3.org/2002/07/owl#DatatypeProperty | | |
| Transform: lowerCase (lowerCase2) | http://www.w3.org/2002/07/owl#DatatypeProperty | | |
| Input: rdf:type (targetPath1) | http://www.w3.org/2002/07/owl#DatatypeProperty | | |
| atankara.com/ontology/turcovid19-ontology#numberOfTotalHrdDoses | http://ns.inria.fr/covid19/property#CasesIntensiveCare | 100.0% | ✓ |
| atankara.com/ontology/turcovid19-ontology#numberOfTotalHrdDoses | http://ns.inria.fr/covid19/property#DeathsEphad | 100.0% | ✓ |
| atankara.com/ontology/turcovid19-ontology#numberOfTotalHrdDoses | http://ns.inria.fr/covid19/property#confidence | 100.0% | ✓ |

Şekil 73. Covid-on-the-web ve TURCOVID19 ontojisindeki datapropertylerin eşleştirilmesi

Karma ile oluşturulan RDF verilerindeki web üzerindeki veri kaynağına işaret eden URI referansı RDF üçlüsündeki nesnesi (object) veya yüklemi (predicate) şeklindedir.


```

+ → 🌐 localhost:8080/pulinh/0D/WSP1W514735M
https://www.gultenkara.com/ontology/Adana<https://www.gultenkara.com/ontology/turcovid19-ontology#cityName> "Adana" .
https://www.gultenkara.com/ontology/Adana<https://www.gultenkara.com/ontology/turcovid19-ontology#startDate> "14.02.2021" .
https://www.gultenkara.com/ontology/Adana<https://www.gultenkara.com/ontology/turcovid19-ontology#startDate> "06.02.2021" .
https://www.gultenkara.com/ontology/Adana<http://www.w3.org/1999/02/22-rdf-syntax-ns#type> <https://www.gultenkara.com/ontology/turcovid19-ontology#MemberOfProvincialLevelCases> .
https://www.gultenkara.com/ontology/Adana<https://www.gultenkara.com/ontology/turcovid19-ontology#MemberOfWeeklyCases-1000> "49,58" .

https://www.gultenkara.com/ontology/AdiYaman<https://www.gultenkara.com/ontology/turcovid19-ontology#startDate> "14.02.2021" .
https://www.gultenkara.com/ontology/AdiYaman<https://www.gultenkara.com/ontology/turcovid19-ontology#MemberOfWeeklyCases-1000> "100,55" .
https://www.gultenkara.com/ontology/AdiYaman<http://www.w3.org/1999/02/22-rdf-syntax-ns#type> <https://www.gultenkara.com/ontology/turcovid19-ontology#MemberOfProvincialLevelCases> .

https://www.gultenkara.com/ontology/AdiYaman<https://www.gultenkara.com/ontology/turcovid19-ontology#startDate> "06.02.2021" .
https://www.gultenkara.com/ontology/AdiYaman<https://www.gultenkara.com/ontology/turcovid19-ontology#cityName> "AdiYaman" .

https://www.gultenkara.com/ontology/afyon<http://www.w3.org/1999/02/22-rdf-syntax-ns#type> <https://www.gultenkara.com/ontology/turcovid19-ontology#MemberOfProvincialLevelCases> .
https://www.gultenkara.com/ontology/afyon<https://www.gultenkara.com/ontology/turcovid19-ontology#cityName> "Afyon" .
https://www.gultenkara.com/ontology/afyon<https://www.gultenkara.com/ontology/turcovid19-ontology#startDate> "14.02.2021" .
https://www.gultenkara.com/ontology/afyon<https://www.gultenkara.com/ontology/turcovid19-ontology#startDate> "06.02.2021" .
https://www.gultenkara.com/ontology/afyon<https://www.gultenkara.com/ontology/turcovid19-ontology#MemberOfWeeklyCases-1000> "21,21" .

https://www.gultenkara.com/ontology/Alayunt<https://www.gultenkara.com/ontology/turcovid19-ontology#MemberOfWeeklyCases-1000> "27,50" .
https://www.gultenkara.com/ontology/Alayunt<http://www.w3.org/1999/02/22-rdf-syntax-ns#type> <https://www.gultenkara.com/ontology/turcovid19-ontology#MemberOfProvincialLevelCases> .
https://www.gultenkara.com/ontology/Alayunt<https://www.gultenkara.com/ontology/turcovid19-ontology#cityName> "Alayunt" .
https://www.gultenkara.com/ontology/Alayunt<https://www.gultenkara.com/ontology/turcovid19-ontology#startDate> "14.02.2021" .
https://www.gultenkara.com/ontology/Alayunt<https://www.gultenkara.com/ontology/turcovid19-ontology#startDate> "06.02.2021" .

https://www.gultenkara.com/ontology/Akaray<https://www.gultenkara.com/ontology/turcovid19-ontology#cityName> "Akaray" .
https://www.gultenkara.com/ontology/Akaray<https://www.gultenkara.com/ontology/turcovid19-ontology#startDate> "14.02.2021" .
https://www.gultenkara.com/ontology/Akaray<http://www.w3.org/1999/02/22-rdf-syntax-ns#type> <https://www.gultenkara.com/ontology/turcovid19-ontology#MemberOfProvincialLevelCases> .
https://www.gultenkara.com/ontology/Akaray<https://www.gultenkara.com/ontology/turcovid19-ontology#startDate> "06.02.2021" .
https://www.gultenkara.com/ontology/Akaray<https://www.gultenkara.com/ontology/turcovid19-ontology#MemberOfWeeklyCases-1000> "92,81" .

https://www.gultenkara.com/ontology/Asay<https://www.gultenkara.com/ontology/turcovid19-ontology#cityName> "Asay" .
https://www.gultenkara.com/ontology/Asay<https://www.gultenkara.com/ontology/turcovid19-ontology#startDate> "14.02.2021" .
https://www.gultenkara.com/ontology/Asay<http://www.w3.org/1999/02/22-rdf-syntax-ns#type> <https://www.gultenkara.com/ontology/turcovid19-ontology#MemberOfProvincialLevelCases> .
https://www.gultenkara.com/ontology/Asay<https://www.gultenkara.com/ontology/turcovid19-ontology#startDate> "06.02.2021" .
https://www.gultenkara.com/ontology/Asay<https://www.gultenkara.com/ontology/turcovid19-ontology#MemberOfWeeklyCases-1000> "36,66" .

https://www.gultenkara.com/ontology/Askara<https://www.gultenkara.com/ontology/turcovid19-ontology#cityName> "Askara" .
https://www.gultenkara.com/ontology/Askara<https://www.gultenkara.com/ontology/turcovid19-ontology#startDate> "06.02.2021" .

```

Şekil 76. NumberOfProvincialLevelCases.ttl dosyasının gösterimi

Yapılan işlemlerden sonra veri seti küresel veri uzayına bağlanarak veri setinde yer alan kavramlar hakkında ek bilgilerin bulunması sağlanmış olacaktır (Ulutaş Karakol vd., 2018).

3.5. TURCOVID19 Bağlantılı Verilerinin Sorgulanması

Bağlantılı Açık Veri Bulutu ve kelime hazineleri veri kaynaklarındaki kavram tanımlarını oluşturmak amacıyla geliştirilmiştir (Ulutaş Karakol vd., 2018). Bağlantılı veri kaynaklarında bulunan URI'ler farklı veri kaynaklarında bulunan aynı kavramı temsil eden URI'ler ile ilişkilendirilmesi için owl:sameAs ilişkisi kurulmuştur.

TURCOVID19 ontolojisinde COVID-19 alanına ait sınıflar oluşturulmuştur ve DBpedia ontolojisinde bulunan Pandemic sınıfı ile owl:sameAs ilişkisi SILK yazılımı kullanılarak kurulmuştur. Aşağıda verilen sorguda 16.05.2022 tarihinde Türkiye geneli toplam ölüm-vaka ve yeni ölüm-vaka sayılarının gösterilmesi ve dünya geneli hakkında bilgi edinmek için de farklı gösterge tablolarına ulaşılmak istenmiştir.

PREFIX owl: <http://www.w3.org/2002/07/owl#>

Prefix codo:<<http://www.isibang.ac.in/ns/codo#>>

Prefix turcovid19-ontology: <<https://www.gultenkara.com/ontology/turcovid19-ontology#>>

Prefix dbo: <http://dbpedia.org/ontology/>

Prefix dbr: http://dbpedia.org/resource/

```

SELECT ?stat ?mydate ?newcases ?totalcases ?newdeaths ?totaldeaths ?pandemic
?resources

WHERE {
?node rdf:type ?stat.
?node codo:covid-19caseOn ?mydate.
?node codo:cases ?totalcases.
?node codo:dailyIncreasedCases ?newcases.
?node codo:deceased ?totaldeaths.
?node turcovid19-ontology:numberOfDailyDeceased ?newdeaths.
?stat owl:sameAs ?pandemic.
?pandemic dbo:wikiPageExternalLink ?resources.
FILTER (?mydate="5/16/2022")
FILTER (?resources=<https://www.who.int/en/>)
}

```

Şekil 77’de gösterilen sorgu sonucunda 16.05.2022 tarihinde Türkiye geneli toplam ölüm-vaka ve yeni ölüm-vaka verilerine ve dünya genelinde bu veriler hakkında bilgi edinmek için de WHO Coronavirus (COVID-19) Dashboard sayfasına ulaşılmıştır.



Example Queries

Selection of triples Selection of classes

Prefixes: [rdf](#) [rdfs](#) [owl](#) [xsd](#)

SPARQL Endpoint: /dbpedia/ Content Type (SELECT): JSON Content Type (GRAPH): Turtle

```

10 Prefix dbo: <http://dbpedia.org/ontology/>
11 Prefix dbr: <http://dbpedia.org/resource/>
12 SELECT ?stat ?mydate ?newcases ?totalcases ?newdeaths ?totaldeaths ?pandemic ?resources
13 WHERE {
14   ?node rdf:type ?stat.
15   ?node codo:covid-19caseOn ?mydate.
16   ?node codo:cases ?totalcases.
17   ?node codo:dailyIncreasedCases ?newcases.
18   ?node codo:deceased ?totaldeaths.
19   ?node turcovid19-ontology:numberOfDailyDeceased ?newdeaths.
20   ?stat owl:sameAs ?pandemic.
21   ?pandemic dbo:wikiPageExternalLink ?resources.
22   FILTER (?mydate="5/16/2022")
23   FILTER (?resources=<https://www.who.int/en/>)
24 }

```

Table Response 1 result in 0.073 seconds

Simple view Ellipse Filter query results Page size: 50

| stat | mydate | newcases | totalcases | newdeaths | totaldeaths | pandemic | resources |
|------------------------------------|-----------|----------|------------|-----------|-------------|-------------------------|---------------------------|
| <https://www.guftenkara.com/ontolo | 5/16/2022 | 1,254 | 11,055,376 | 4 | 88,904 | <http://dbpedia.org/res | <https://www.who.int/en/> |

Showing 1 to 1 of 1 entries

Şekil 77. Türkiye ve dünya geneli COVID-19 verileri SPARQL sorgusu



Şekil 78. WHO Coronavirus (COVID-19) Dashboard sayfası

Aşağıda verilen sorguda tarihlere göre yeni hastane yatış sayılarına ve pandemi hakkında daha fazla bilgi edinmek için ilgili sayfa ve dokümana ulaşılmak istenmiştir.

PREFIX owl: <<http://www.w3.org/2002/07/owl#>>

Prefix rdf: <<http://www.w3.org/1999/02/22-rdf-syntax-ns#>>

Prefix codo:<<http://www.isibang.ac.in/ns/codo#>>

Prefix turcovid19-ontology: <<https://www.gultenkara.com/ontology/turcovid19-ontology#>>

Prefix dbo: <<http://dbpedia.org/ontology/>>

Prefix dbp: <<http://dbpedia.org/property/>>

Prefix georss: <<http://www.georss.org/georss/>>

Select ?date ?HospitalAdmissions ?page ?documents

Where {

?b turcovid19-ontology:numberOfNewHospitalAdmissions ?HospitalAdmissions.

?b codo:covid-19caseOn ?date.

?b rdf:type ?stat.

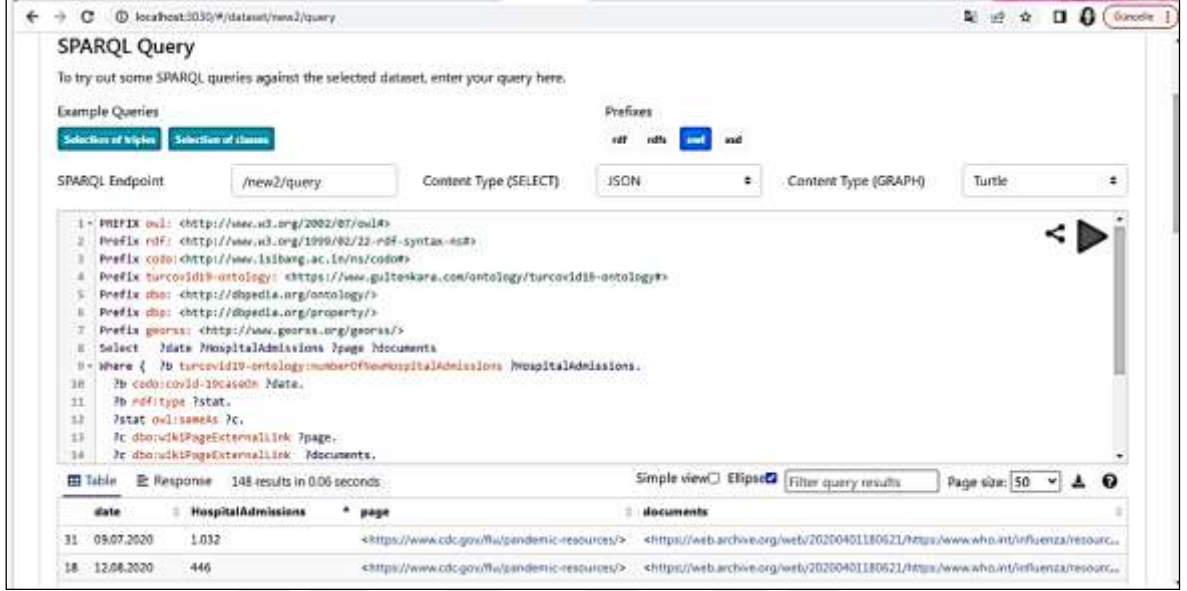
?stat owl:sameAs ?c.

?c dbo:wikiPageExternalLink ?page.

?c dbo:wikiPageExternalLink ?documents.

}

Şekil 79’da görüldüğü sorgu sonucunda her bir tarihe göre yeni hastane yatış sayıları ve pandemi ile ilgili yayın yapan sayfa ve dökümana hızlıca ulaşım sağlanmaktadır.



The screenshot shows a SPARQL Query interface. The query is as follows:

```

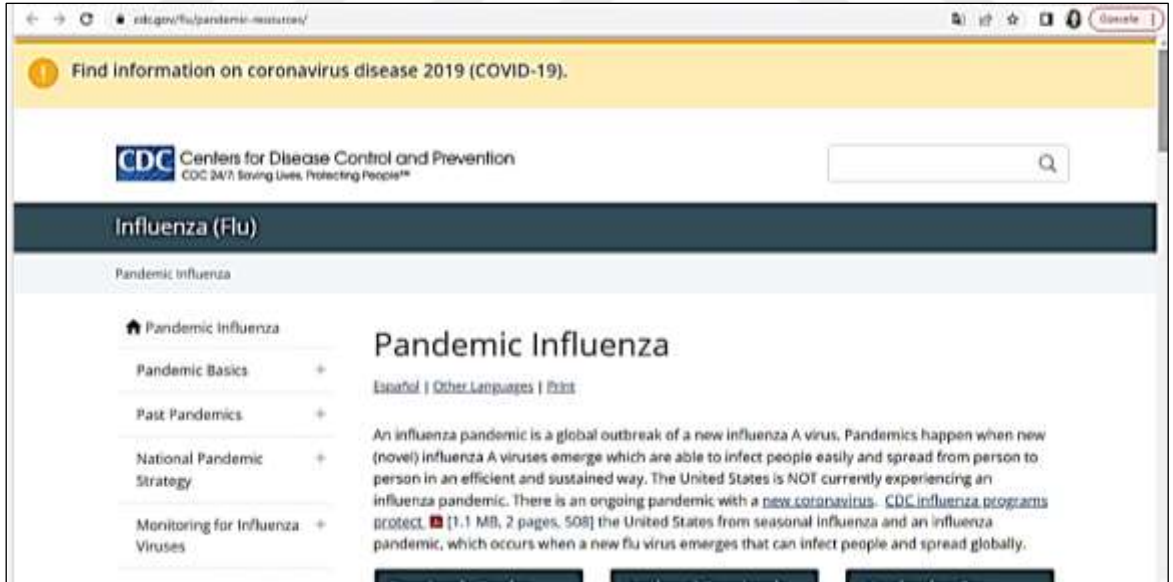
1 PREFIX owl: <http://www.w3.org/2002/07/owl#>
2 PREFIX rdf: <http://www.w3.org/1999/02/22-rdf-syntax-ns#>
3 PREFIX covid: <http://www.isibang.ac.in/ns/covid#>
4 PREFIX turcovid19-ontology: <https://www.guilekara.com/ontology/turcovid19-ontology#>
5 PREFIX dpo: <http://dbpedia.org/ontology/>
6 PREFIX dbp: <http://dbpedia.org/property/>
7 PREFIX georss: <http://www.georss.org/georss/>
8 SELECT ?date ?hospitalAdmissions ?page ?documents
9 WHERE {
10   ?d turcovid19-ontology:numberOfNewHospitalAdmissions ?hospitalAdmissions.
11   ?d covid:covid-19caseOn ?date.
12   ?d rdfs:type ?stat.
13   ?stat owl:sameAs ?c.
14   ?c dbp:wikiPageExternalLink ?page.
15   ?c dbp:wikiPageExternalLink ?documents.

```

The results table shows the following data:

| # | date | HospitalAdmissions | page | documents |
|----|------------|--------------------|---|---|
| 31 | 09.07.2020 | 1.032 | <https://www.cdc.gov/flu/pandemic-resources/> | <https://web.archive.org/web/20200401180621/https://www.who.int/fluenza/resourc...> |
| 18 | 12.06.2020 | 446 | <https://www.cdc.gov/flu/pandemic-resources/> | <https://web.archive.org/web/20200401180621/https://www.who.int/fluenza/resourc...> |

Şekil 79. COVID-19 bağlantılı verilerinin sorgulanması



The screenshot shows the CDC website page for Pandemic Influenza. The page title is "Pandemic Influenza" and the subtitle is "Pandemic Influenza". The page content includes a search bar, a navigation menu, and a main text block. The main text block contains the following information:

Pandemic Influenza

An influenza pandemic is a global outbreak of a new influenza A virus. Pandemics happen when new (novel) influenza A viruses emerge which are able to infect people easily and spread from person to person in an efficient and sustained way. The United States is NOT currently experiencing an influenza pandemic. There is an ongoing pandemic with a [new coronavirus](#). [CDC influenza programs protect](#) [1.1 MB, 2 pages, 508] the United States from seasonal influenza and an influenza pandemic, which occurs when a new flu virus emerges that can infect people and spread globally.

Şekil 80. Centres for Disease Control and Prevention Web Sitesi

| PHASE | DESCRIPTION | MAIN ACTIONS | | | | |
|---------|--|--|--|--|--|--|
| | | PLANNING AND COORDINATION | DETECTION, MONITORING AND REPORTING | COMMUNICATION | REDUCING THE SPREAD OF DISEASE | CONTINUITY OF HEALTH-CARE SYSTEMS |
| PHASE 1 | No animal influenza virus circulating among animals has been reported to cause infection in humans. | | | | | |
| PHASE 2 | An animal influenza virus circulating in domesticated or wild animals is known to have caused infection in humans and is therefore considered a specific potential pandemic threat. | Develop, exercise, and periodically update national influenza pandemic preparedness and response plans. | Develop robust national surveillance systems in collaboration with national animal health authorities, and other relevant sectors. | Complete communications planning and initiate communications activities to communicate real and potential risks. | Promote beneficial behaviours in individuals for and countries. Plan for use of pharmaceuticals and vaccines. | Prepare the health system to scale up. |
| PHASE 3 | An animal or human/animal influenza reassortant virus has caused sporadic cases or small clusters of disease in people, but has not resulted in human-to-human transmission sufficient to sustain community-level outbreaks. | | | | | |
| PHASE 4 | Human to human transmission of an animal or human/animal influenza reassortant virus able to sustain community-level outbreaks has been verified. | Effect and coordinate rapid response mechanisms. Activities to collaborate with WHO to limit or delay the spread of infection. | Increase surveillance. Monitor development of variants. Share findings with WHO and the international community. | Promote and continuously re-evaluate interventions to prevent and reduce population and individual risk. | Implement rapid pandemic containment operations and other activities, collaboratively with WHO and the international community as necessary. | Activate contingency plans. |

Şekil 81. Pandemi ile ilgili webde bulunan belgeler

COVID-19 veri setlerindeki veriler doğrudan bağlantı değildir. Bu nedenle verilerde yer alan il adları DBpedia'daki iller ile ilişkilendirilmiştir. Aşağıda yer alan sorguda 12.02.2022 tarihinde Trabzon iline ait 100.000 kişide haftalık vaka sayısı ve Trabzon ilinin yüzölçümü, koordinatlarına ve il hakkında genel bilgiler edinmek için Wikipedia sayfasına ve konum bilgileri için Geonames sayfasına ulaşılmak istenmiştir.

PREFIX owl: <<http://www.w3.org/2002/07/owl#>>

Prefix turcovid19-ontology: <<https://www.gultenkara.com/ontology/turcovid19-ontology#>>

Prefix dbp: <<http://dbpedia.org/property/>>

Prefix georss: <<http://www.georss.org/georss/>>

Select ?date ?weeklycases ?city ?point ?area ?wikipedia ?geonames

Where { ?b turcovid19-ontology:numberOfWeeklyCases-100K ?weeklycases.

?b turcovid19-ontology:cityName ?city.

?b turcovid19-ontology:startDate ?date.

?b owl:sameAs ?c.

?c foaf:isPrimaryTopicOf ?wikipedia.

?c dbp:areaTotalKm ?area .

?c georss:point ?point.

?c owl:sameAs ?geonames

FILTER (?geonames= <http://sws.geonames.org/738648/>)
}

Aşağıda verilen Şekil 82’de sorgu sonucunda Trabzon iline ait vaka sayılarına ve Trabzon ilinin yüzölçümü, nüfüsü ve konum bilgilerine RDF linkleri sayesinde hızlıca ulaşılmıştır.

SPARQL Query

To try out some SPARQL queries against the selected dataset, enter your query here.

Example Queries
[Selection of triples](#) [Selection of classes](#)

Prefixed
 rdf rdfs owl xsd

SWARQL Endpoint: /dbpedia2/sparql Content Type (SELECT): JSON Content Type (GRAPH): Turtle

```

10 PREFIX turcovid19-ontology: <https://www.galtestara.com/ontology/turcovid19-ontology#>
11 PREFIX dbpi: <http://dbpedia.org/property/>
12 PREFIX geonax: <http://www.geonax.org/geonax/>
13 SELECT ?date ?weeklycases ?city ?point ?area ?wikipedia ?geonames
14 WHERE {
15   ?tr turcovid19-ontology:summarizeWeeklyCases-1998 ?weeklycases.
16   ?tr turcovid19-ontology:cityName ?city.
17   ?tr owl:sameAs ?c.
18   ?c foaf:isPrimaryTopicOf ?wikipedia.
19   ?c dbpi:areaTotalIn ?area .
20   ?c geonax:point ?point.
21   ?c owl:sameAs ?geonames.
22   FILTER ( ?geonames= <http://sws.geonames.org/738648/> )
23 }
  
```

Table Response 1 result in 0.006 seconds

| date | weeklycases | city | point | area | wikipedia | geonames |
|------------|-------------|---------|-------------------------------------|------------|---|-----------------------------------|
| 12.02.2020 | 149185 | Trabzon | 40.78055555555556 39.81222222222222 | 6485777777 | <http://en.wikipedia.org/wiki/Trabzon_Province> | <http://sws.geonames.org/738648/> |

Showing 1 to 1 of 1 entries

Şekil 82. Trabzon iline ait Bağlantılı Verilerin sorgulanması

WIKIPEDIA
The Free Encyclopedia

Article Tab

Read Edit View history Search Wikipedia

Trabzon Province

From Wikipedia, the free encyclopedia

Coordinates: 42°40′07″N 39°42′44″E﻿ / ﻿42.66861°N 39.71222°E﻿ / 42.66861; 39.71222

Trabzon Province (Turkish: Trabzon ili) is a province of Turkey on the Black Sea coast. Located in a strategically important region, Trabzon is one of the oldest trade port cities in Anatolia. Neighbouring provinces are Giresun to the west, Gümüşhanev to the southwest, Bayburt to the southeast and Rize to the east. İsmail Ustaoglu was appointed the Governor of the province in October 2018.^[d]

The capital of the province is Trabzon.

Contents [hide]

- Districts
- History
 - 2.1 In Turkey
 - 2.2 Archaeology
- Attractions
- Demographics
- See also
- References

Trabzon Province
Trabzon ili
Province of Turkey

Bunla bağlantılı

Şekil 83. Trabzon iline ait genel bilgiler



Şekil 84. Trabzon iline ait nüfus ve konum bilgileri

TURCOVID-19 verileri bölge birimlerine (İBBS) göre yayınlanmıştır. Veriler LOD veri kaynağı olan NUTS veri setinde yer alan bölge birimleri isimleri ile eşleştirilerek ilişkilendirilmiştir. Aşağıda verilen sorguda illere ait vaka sayıları ve illerin NUTS (İBBS) kodları, sınıflandırma içinde hangi düzeyde oldukları ve NeoGeo kelime hazinesi ile geometri tanımlarına ulaşılmak istenmiştir.

Prefix owl: <<http://www.w3.org/2002/07/owl#>>

Prefix codo:<<http://www.isibang.ac.in/ns/codo#>>

Prefix turcovid19-ontology: <<https://www.gultenkara.com/ontology/turcovid19-ontology#>>

Prefix ramon: <<http://rdfdata.eionet.europa.eu/ramon/ontology/>>

Prefix ngeo: <<http://geovocab.org/geometry#>>

Select ?baş ?bit ?weeklycases ?city ?nuts ?code ?level ?geometry

Where { ?b turcovid19-ontology:numberOfWeeklyCases-100K ?haftalıkvaka.

?b turcovid19-ontology:cityName ?city.

?b turcovid19-ontology:startDate ?baş.

?b turcovid19-ontology:endDate ?bit.

?b owl:sameAs ?nuts.

?nuts ramon:code ?code.

?nuts ramon:level ?level.

?nuts ngeo:geometry ?geometry

FILTER (?baş="15.05.2021")

}

4. SONUÇLAR VE ÖNERİLER

Pandemi sürecinde, birkaç kuruluş büyük hacimli heterojen açık veri kümeleri yayınlamaktadır. Birçoğu kendi özel formatlarında yayınlamıştır. Yapılan çalışma ile farklı formatlarda alınan COVID-19 verilerinin semantik tanımları oluşturularak entegre edilmiştir ve kullanıcılar SPARQL sorgu dilini kullanarak tek tip olarak erişmesini ve sorgulamasını sağlamak için LOD Cloud'da yapılandırılmış RDF formatında kullanıma sunmaktadır.

SWT kullanarak verilerin LOD'a dönüştürülmesinin sağlayabileceği avantajlar ortaya çıkmıştır. Bu avantajlar arasında, heterojen verilerin semantik tanımların oluşturulması ve birbirine bağlanması, veri entegrasyonun maliyetinin düşürülmesi, basitleştirilmiş açıklama ve paylaşım, verilerin HTTP aracılığıyla erişim için gösterilmesi ile kolay bir şekilde yeniden kullanımı ve ek bilgilerin çıkarılması yer almaktadır.

Tez çalışmasında COVID-19 alanı içinde tanımlanan ontolojiler ile RDF verileri Fuseki veritabanında üçlüler olarak yüklenmiştir. Sorgulamalar yapılarak sonuçlar değerlendirilmiştir. Elde edilen sonuçlara göre yükleme ve çıkarsama sürelerinin veri setinin büyüklüğüne bakılmaksızın değişmiyor olması Fuseki üzerinde çalışılabileceğini göstermektedir.

Trabzon İl Sağlık Müdürlüğü aracılığı ile Sağlık Bakanlığında tez kapsamında Trabzon'da vakaların başladığı tarihten itibaren hasta bilgileri (cinsiyeti, yaşı, bildirilen vakanın semptomatik hale geldiği tarih, rapor edilen vakanın COVID-19 virüsüne yakalandığının onaylı tarihi, vakanın bildirildiği konum bilgileri, vakanın seyahat geçmişi, vakaların ölüm veya taburculuk tarihi, vakanın rapor edildiği idari birim düzeyi, rapor edilen vakanın bildirilen kronik bilgileri) ve il düzeyinde günlük hasta sayısı, günlük vefat sayısı, günlük test sayısı, günlük iyileşen hasta sayısı verileri dilekçe yolu ile talep edilmiştir. Trabzon İl Sağlık Müdürlüğü'nde ilgili birimde görev yapan kişiler ile de sürekli iletişim halinde olunmuştur. Veri talebinin cevaplanması için CİMER'e de yazılmıştır. Fakat 1 yılı geçen süre içerisinde veri talebinin bakanlık tarafından incelenmekte olduğu, istenilen verilerin paylaşılmasının uygun olup olmadığı yönünde herhangi bir bilgi verilmemiştir. Bu süreçte ayrıca Sağlık Bakanlığı sitesinde akademik çalışmalar için hızlı bir şekilde veri talebinin gerçekleştirileceği yönünde yapılan duyuru

kapsamında başvuru formu doldurulmuştur. Hızlı bir şekilde geri dönüş yapılmıştır ve veri talebimiz için Sağlık Bilgi Sistemleri Genel Müdürlüğü Veri Madenciliği, Analiz, Büyük Veri ve Raporlama Birimi'ne yönlendirme yapılmıştır. Ancak bu girişimde olumsuz sonuçlanmıştır. Bakanlık tarafından istenilen verilerin paylaşılması ile aşağıda verilmiş olan ontolojinin amacına uygun yetkinlik sorularına yanıtlar bulunacaktı.

- Trabzon iline ait belirli bir zamana kadar COVID-19 hastalarından iyileşenlerin ve ölenlerin sayısı
- COVID-19 hastalarını ve varsa temaslıları kimlerdir?
- Aile ilişkileri içinde olan COVID-19 hastaları?
- Maksimum COVID-19 hasta sayısının birincil nedenleri?
- Şiddetli COVID-19'un semptomları?
- COVID-19 teşhisi konmuş ve henüz test edilmemiş biriyle akraba olan tüm kişiler?
- Salgın sırasında yerel politikalar hangi seyahat kısıtlamalarını uyguluyor?
- SARS-CoV-2'nin varyantları nelerdir ve nasıl tedavi edilmeli veya önlenmelidir?
- Dünya çapındaki bu görünmez savaşta erzak, sıcak olaylar ve ön saflardaki sağlık çalışanlarının durumu nedir?

Web üzerindeki veri setleri ile RDF formatındaki verilerin ilişkilendirilmesi için uygun veri setlerinin seçilmesi gerekir. Bağlantılı Açık Veri Bulutu içerisinde yer alan veri setleri incelendiğinde 3 tane COVID-19 veti bulunmaktadır. Bu nedenle Bağlantılı Açık Veri Bulutu'nda bulunan COVID-19 veri setlerinin genişletilmesi gerekir. Ayrıca uygun veri setlerinin belirlenmesi ve eşleştirilmesi, kullanıcıların veri setlerinin içeriği hakkında bilgi sahibi olmaları gerekir ve bu durum bağlantılı veri yaklaşımındaki engellerden birisidir.

Üretilen TURCOVID19 veri setlerinin LOD Cloud'da yer alan veri setleri ile arasındaki bağlantılar Silk bağlantı editörü'nde otomatik olarak bulunmuştur. Bulunan bağlantı sonuçları yüksek doğrulukta oluşmuştur. Böylece TURCOVID19 veri seti global web kaynakları ile zenginleştirilmiştir. Farklı kaynaklar arasında kurulan RDF linkleri ile SPARQL sorguları oluşturularak TURCOVID19 veri setinde yer alan kavramlar hakkında birçok bilgiye hızlıca erişim sağlandığı tespit edilmiştir.

Çalışmada kullanılan veri setindeki kavramların Türkçe karakterler içermesi bağlantı kurma aşamasında problem oluşturmuştur. Problemi gidermek için veri seti Türkçe karakterlerden ayıklandıktan sonra kullanılmıştır (Karakol ve Cömert, 2021). SILK ile

yapılan eşleştirmeler syntactic (sözdizimsel) olarak gerçekleşmektedir. Semantik düzeyde eşleştirmeler için SILK'in geliştirilmesi gerekmektedir. Ek olarak Türkçe karakterler tanıyan bağlantı kurma araçları geliştirilmelidir.

Kullanıcı tarafından sağlanan bağlantı özelliklerini ve koşullarını kullanarak mevcut farklı Web veri kaynakları içindeki varlıklar arasında zaman açısından verimli RDF bağlantı keşfi oluşturmak için araçlar sağlanmaktadır. Oluşturulan RDF bağlantıları, LOD bulutundaki orijinal RDF veri kümesiyle birlikte yayınlanması gelecek çalışmalarda yapılması hedeflenmektedir.

COVID-19 verilerinin bağlantılı açık veri olarak yayınlandığı bu çalışmada oluşturulan bağlantılı veri Fuseki ile yayınlanmıştır. Konum tabanlı görselleştirme, sağlık verilerinde büyük önem kazanmaktadır. Bununla birlikte, geleneksel veritabanı sistemlerinde konum tabanlı görselleştirme, verilerin analizi için Coğrafi Bilgi Sistemleri (CBS) ortamına aktarılmasını gerektirir. Gelecek çalışmalarda COVID-19 verileri için topolojik ilişkilerin kullanılması ve SPARQL sorgu sonucunun harita üzerinde gösterilerek sunulması hedeflenmektedir.

5. KAYNAKLAR

- Ali, S., Zada, I., Mehmood, Z., Ullah, A., Ali, H. ve Ullah, M., 2022. Publishing and Interlinking COVID-19 Data Using Linked Open Data Principles: Toward Effective Healthcare Planning and Decision-Making, *Mathematical Problems in Engineering*.
- Angelis, S. ve Kotis, K., 2020. Generating and Exploiting Semantically Enriched: Integrated, Linked and Open Museum Data, Metadata and Semantic Research, 14th International Conference, December, Spain, Revised Selected Papers, 1355, 367-379.
- Ayers, D. ve Völkel, M., 2008. Cool URI, Semantic Web W3C Interest Group Note.
- Babcock, S., Beverley, J., Cowell, L. G. ve Smith, B., 2021. The Infectious Disease Ontology in the age of COVID-19, Journal of biomedical semantics, 12, 1, 1-20.
- Battle, R. ve Kolas, D., 2012. Enabling the Geospatial Semantic Web with Parliament and Geosparql, Semantic Web, 3, 4, 355-370.
- Battle, R. ve Kolas, D., 2011. Geosparql: Enabling a Geospatial Semantic Web, Semantic Web Journal, 3, 4, 355-370.
- Becker, C. ve Bizer, C., 2009. Exploring the Geospatial Semantic Web with Dbpedia Mobile, Web Semantics, 7, 4, 278-286.
- Berners-Lee, T., Linked Data, <https://www.w3.org/DesignIssues/LinkedData.html>, 21 Ocak 2021.
- Berners-Lee, T., Chen, Y., Chilton, L., Connolly, D., Dhanaraj, R., Hollenbach, J., Lerer, A. ve Sheets, D., 2006. Tabulator: Exploring and Analyzing Linked Data on the Semantic Web, 3rd International Semantic Web User Interaction Workshop, November.
- Berners-Lee, T., Fielding, R. ve Masinter, L., 1998. RFC2396: Uniform resource identifiers (URI): generic syntax.
- Berners-Lee, T., Fielding, R. ve Masinter, L., 2005. Uniform resource identifier (URI): Generic syntax (No. rfc3986).
- Berners-Lee, T., Hendler, J. ve Lassila, O., 2001. The Semantic Web, Scientific American, 284, 5, 34-43.
- Bikakis, N., Tsinaraki, C., Gioldasis, N., Stavrakantonakis, I. ve Christodoulakis, S., 2013. The XML and semantic web worlds: technologies, interoperability and integration: a survey of the state of the art, In Semantic hyper/multimedia adaptation, Anagnostopoulos I., Bieliková M., Mylonas P., Tsapatsoulis N., Springer, Berlin, Heidelberg.

- Bizer, C., Heath, T. ve Berners-Lee, T., 2011. Linked data: The story so far, in Semantic services, interoperability and web applications: emerging concepts, IGI global, 205-227.
- Bryl, V., Bizer, C., Isele, R., Verlic, M., Hong, S. G., Jang, S. ve Choi, K. S., 2014. Interlinking and knowledge fusion, In Linked Open Data--Creating Knowledge Out of Interlinked Data, Springer, Cham, 70-89.
- Cardoso, J., 2007. Semantic Web Services: Theory, Tools and Applications, IGI Global, 372.
- de León, A., Wisniewki, F., Villazón-Terrazas, B. ve Corcho, O., 2012. Map4rdf-faceted browser for geospatial datasets, In Proceedings of the First Workshop on Using Open Data.
- Devedžić, V., 2006. Semantic web and education, 12. Springer Science and Business Media.
- Demirli, C. ve Kütük, Ö. F., 2010. Anlamsal web (web 3.0) ve ontolojilerine genel bir bakış, İstanbul Ticaret Üniversitesi Fen Bilimleri Dergisi, 9,18, 97-71.
- DuCharme, B., 2013. Learning SPARQL: querying and updating with SPARQL 1.1., 2nd Edition, 386, O'Reilly Media, Inc.
- Dunwoodie, B., What is an Ontology? and Why We Need Them, <https://www.cmswire.com/cms/knowledge-management/what-is-an-ontology-and-why-we-need-them-001479.php> , 22 Ocak 2021
- Dutta, B. ve DeBellis, M., 2020. CODO: an ontology for collection and analysis of COVID-19 data, arXiv preprint arXiv:2009.01210.
- Dong, E., Du, H. ve Gardner, L., 2020. An interactive web-based dashboard to track COVID-19 in real time, The Lancet Infectious Diseases, 20,5, 533-534.
- Ege, B., 2011. Yeni Bilgi Modelleme ve Programlama Felsefesiyle Semantik Web, Bilim ve Teknik dergisi.
- Ehimwenma, K. E., Crowther, P., Beer, M. ve Al-Sharji, S., 2020. An SQL Domain Ontology Learning for Analyzing Hierarchies of Structures in Pre-Learning Assessment Agents, SN Computer Science, 1, 6, 1-19.
- Emmons, I., Parliament User Guide, Raytheon BBN Technologies, <https://usermanual.wiki/Pdf/ParliamentUserGuide.1438980087/help>, 23 Ocak 2021.
- Fielding, R., Gettys, J., Mogul, J., Frystyk, H., Masinter, L., Leach, P. ve Berners-Lee, T., Hypertext transfer protocol–HTTP/1.1, <https://www.ietf.org/rfc/rfc2616.txt>., 23 Ocak 2021.
- Kara, G. ve Cömert, Ç., 2016. Konumsal Verilerin Bağlantılı Veri Olarak Yayınlanması, 6. Uzaktan Algılama-CBS Sempozyumu, 5-7.

- Karakol, D. U. ve Cömert, Ç., 2021. Meteorolojik Verilerin Bağlantılı Veriye Dönüştürülmesi: Karadeniz Bölgesi Örneği, Turkish Journal of Remote Sensing and GIS, 2, 1, 47-58.
- Kepeklian, G., Bihanic, L. ve Troncy, R., 2014. Datalift: A platform for integrating big and linked data, in Proceedings of the International Conference on Big Data from Space, November, Italy, 370-373.
- Gandon, F., Krummenacher, R., Han, S. K. ve Toma, I., 2011. The resource description framework and its schema.
- Geroimenko, V. ve Chen, C., 2006. Visualizing the Semantic Web: XML-based internet and information visualization, Springer Science & Business Media, 2nd Edition, 248, Springer-Verlag London.
- Giri, K., 2011. Role of ontology in Semantic Web, DESIDOC Journal of Library & Information Technology, 31,2, 116-120.
- Gültepe, Y. ve Ünalır, M., 2014. Ulusal Sağlık Veri Sözlüğü'nün Ontoloji Tabanlı Hale Getirilmesi ve Tıbbi Bilişim Standartları ile Zenginleştirilmesi. Gazi Üniversitesi Mühendislik Mimarlık Fakültesi Dergisi, 29, 3.
- Hale, T., Angrist, N., Goldszmidt, R., Kira, B., Petherick, A., Phillips, T. ve Tatlow, H., 2021. A global panel database of pandemic policies (Oxford COVID-19 Government Response Tracker), Nature Human Behaviour, 5, 4, 529-538.
- Hart, G. ve Dolbear, C., 2013. Linked data: a geographic perspective, 289, Taylor & Francis.
- Hayes, J., A graph model for RDF, Master's Thesis, Darmstadt University of Technology/University of Chile.
- He, Y., Yu, H., Ong, E., Wang, Y., Liu, Y., Huffman, A., Huang, H., John, B., Lin, A. Y., William, D. D., Arabandi, S., Xie, J., Hur, J., Yang, X., Chen, L., Omenn, G. S., Athey, B. ve Smith, B., 2021. CIDO: The Community-Based Coronavirus Infectious Disease Ontology, Proceedings of the 11th International Conference on Biomedical Ontologies (ICBO) and 10th Workshop on Ontologies and Data in Life Sciences (ODLS).
- Heath, T. ve Bizer, C., 2011. Linked data: Evolving the web into a global data space. Synthesis lectures on the semantic web: theory and technology, 1, 1, 1-136.
- Hebeler, J., Fisher, M., Blace, R. ve Perez-Lopez, A., 2011. Semantic Web Programming, First Edition, 652, John Wiley ve Sons.
- Hitzler, P., Krotzsch, M. ve Rudolph, S., 2009. Foundations of Semantic Web Technologies. First Edition, 456, CRC press, New York.
- Horridge, M., Knublauch, H., Rector, A., Stevens, R. ve Wroe, C., A practical guide to building OWL ontologies using the Protégé-OWL plugin and CO-ODE tools edition 1.0. University of Manchester.

- Hyvönen, E., 2012. Publishing and using cultural heritage linked data on the Semantic Web, *Synthesis Lectures on the Semantic Web: Theory and Technology*, 159, Morgan and Claypool Publishers.
- Jacob, E. K., 2003. Ontologies and the Semantic Web, *Bulletin of the American Society for Information Science and Technology*, 29, 4, 19-22.
- Kalyanpur, A., Parsia, B., Sirin, E., Grau, B. C. ve Hendler, J., 2006. Swoop: A web ontology editing browser, *Journal of Web Semantics*, 4, 2, 144-153.
- Kara, A., 2014. İdari Birimler İçin Bir Yer Adları Dizin Ontolojisi Geliştirilmesi ve İstanbul Örneğinde Bağlı Veri Yaklaşımıyla Sunumu.
- Kara, G., Akyazı, İ. ve Cömert, Ç., 2020. Konumsal Verilerin Bağlantılı Veri Olarak Yayınlanması: Trabzon Örneği, *Erciyes Üniversitesi Fen Bilimleri Enstitüsü Fen Bilimleri Dergisi*, 36, 2, 228-237.
- Kara, G. ve Cömert, Ç., 2011. Ulusal Konumsal Veri Altyapısı İçin Semantik Veri Tanımlama, *Jeodezi ve Jeoinformasyon Dergisi*, 104.1, 85-91.
- Kara, G., Karakol, D. U., Yılmaz, C. ve Cömert, Ç., 2019. Konumsal Verilerin Bağlantılı Açık Veri Olarak Yayınlanması, TMMOB Harita ve Kadastro Mühendisleri Odası 17. Türkiye Harita Bilimsel ve Teknik Kurultayı, Nisan, Ankara.
- Karalar, H. ve Özdemir, S., 2013. Anlamsal Web Temelli Öğretimde Yönlendirmenin Kazanıma ve Kalıcılığa Etkisi, *Uluslararası Türk Eğitim Bilimleri Dergisi*, 1, 1-16.
- Khusro, S., Khan, A., Mahfooz, S., Rauf, A. ve Ali, S., 2013. Navigating and Browsing Linked Open Data: State-of-the-Art, *Life Science Journal*, 10, 11, 349-358.
- Kyzirakos, K., Savva, D., Vlachopoulos, I., Vasileiou, A., Karalis, N., Koubarakis, M. ve Manegold, S., 2018. GeoTriples: Transforming geospatial data into RDF graphs using R2RML and RML mappings, *Journal of Web Semantics*, 52, 16-32.
- Lonsdale, D., Embley, D. W., Ding, Y., Xu, L. ve Hepp, M., 2010. Reusing ontologies and language components for ontology generation, *Data & Knowledge Engineering*, 69, 4, 318-330.
- Martinez-Cruz, C., Blanco, I. J. ve Vila, M. A., 2012. Ontologies versus relational databases: are they so different? A comparison, *Artificial Intelligence Review*, 38, 4, 271-290.
- Maier, R., Hädrich, T. ve Peinl, R., 2009. Enterprise knowledge infrastructures, 2nd Edition, 445, Springer-Verlag Berlin Heidelberg.
- McGuinness, D. L. ve Van Harmelen, F., OWL web ontology language overview, W3C recommendation, <https://www.w3.org/TR/owl-features/>, 24 Ocak 2021.

- Michel, F., Gandon, F., Ah-Kane, V., Bobasheva, A., Cabrio, E., Corby, O. ve Winckler, M., 2020. Covid-on-the-Web: Knowledge graph and services to advance COVID-19 research, In *International Semantic Web Conference*, Springer, Cham, 294-310.
- Micsik, A., Tóth, Z. ve Turbucz, S., 2013. LODmilla: Shared visualization of linked open data, In *International Conference on Theory and Practice of Digital Libraries*, September, Malta, 89-100.
- Miller, E. (1998), 1998. An introduction to the resource description framework, *Bulletin of the American Society for Information Science and Technology*, 25, 1, 15-19.
- Ngomo, A. C. N., Auer, S., Lehmann, J. ve Zaveri, A., 2014. Introduction to linked data and its lifecycle on the web, In *Reasoning Web International Summer School*, September, Greece, 1-99.
- Nikolaou, C., Dogani, K., Bereta, K., Garbis, G., Karpathiotakis, M., Kyzirakos, K. ve Koubarakis, M., 2015. Sextant: Visualizing time-evolving linked geospatial data, *Journal of Web Semantics*, 35, 35-52.
- Nikolaou, C., Dogani, K., Kyzirakos, K. ve Koubarakis, M., 2013. Sextant: browsing and mapping the ocean of linked geospatial data, In *Extended Semantic Web Conference*, May, France, 209-213.
- Osman, T., 2022. GeoSPARQL-Jena: Implementation and Benchmarking of a GeoSPARQL Graphstore, In *European Conference on Knowledge Management*, August, 23, 2, 885-894.
- Perry, M. ve Herring, J., OGC GeoSPARQL-a geographic query language for RDF data. Version 1.0, OGC 11-052r4, Open Geospatial Consortium, https://portal.ogc.org/files/?artifact_id=47664, 2 Nisan 2021.
- Pietriga, E., Gözükan, H., Appert, C., Destandau, M., Čebirić, Š., Goasdoué, F. ve Manolescu, I., 2018. Browsing Linked Data Catalogs with LODAtlas, In *International Semantic Web Conference*, October, USA, 137-153.
- Potnis, A. ve Durbha, S. S., 2016, October. Exploring Visualization of Geospatial Ontologies using Cesium, In *VOILA@ ISWC*, 143-150.
- Qundus, J. A., Schäfermeier, R., Karam, N., Peikert, S. ve Paschke, A., 2021. ROC: An Ontology for Country Responses towards COVID-19, arXiv preprint arXiv:2104.07345.
- Saavedra, J., Vilches-Blázquez, L. M. ve Boada, A., 2014. Cadastral data integration through Linked Data, *Proceedings of the AGILE 2014 International Conference on Geographic Information Science*, June, Spain, 81-85.
- Sauermann, L., Cyganiak, R. ve Völkel, M., Cool URIs for the Semantic Web, W3C Interest Group Note, <https://www.w3.org/TR/cooluris/>, 11 Şubat 2021.

- Sayers, C. ve Karp, A. H., Computing the digest of an RDF graph, <https://www.semanticscholar.org/paper/Computing-the-digest-of-an-RDF-graph-Sayers-Karp/e49756a0bf7bcf6ce4b033c4f5261b283d0be394>, 12 Ocak 2021.
- Segaran, T., Evans, C. ve Taylor, J., 2009. Programming the semantic web: build flexible applications with graph data, O'Reilly Media, Inc.
- Stadler, C., Lehmann, J., Höffner, K. ve Auer, S., 2012. Linkedgeodata: A core for a web of spatial open data, Semantic Web, 3, 4, 333-354.
- Scharffe, F., Atemezing, G., Troncy, R., Gandon, F., Villata, S., Bucher, B., Hamdi, F., Bihanic, L., Képéklian, G., Cotton, F., Euzenat, J., Fan, Z., Vandenbussche, P. ve Vatan, B., 2012. Enabling linked data publication with the Datalift platform, Workshops at the Twenty-Sixth AAAI Conference on Artificial Intelligence-Semantic Cities, France, 25-30.
- Sikos, L., 2015. Mastering structured data on the Semantic Web: From HTML5 microdata to linked open data, First Edition, 256, Apress.
- Sir, M., Bradac, Z. ve Fiedler, P., 2015. Ontology versus database, IFAC-PapersOnLine, 48, 4, 220-225.
- Turki, H., Hadj Taieb, M. A., Shafee, T., Lubiana, T., Jemielniak, D., Aouicha, M. B. ve Mietchen, D. 2021. Representing COVID-19 information in collaborative knowledge graphs: the case of Wikidata. Semantic Web, 1-32.
- Turner, E. K., Hinze, A. ve Jones, S., A review of user interface adaption in current semantic web browsers, <https://researchcommons.waikato.ac.nz/handle/10289/5044>, 14 Şubat 2021.
- Uçar, A., Arslan, Ş., Manap, H., Gürkan, T., Çalışkan, M., Dayıoğlu, A. ve Ören, S., 2020. Türkiye’de Covid-19 pandemisinin monitörizasyonu için interaktif ve gerçek zamanlı bir web uygulaması: TURCOVID19, Anatolian Clinic the Journal of Medical Sciences, 25(Special Issue on COVID 19), 154-155.
- Ulutaş Karakol, D., Kara, G., Yılmaz, C. ve Cömert, Ç., 2018. Semantic Linking Spatial Rdf Data to the Web Data Sources. International Archives of the Photogrammetry, Remote Sensing & Spatial Information Sciences.
- URL-1, <https://www.ecdc.europa.eu/en/about-ecdc>, 22 Ocak 2021.
- URL-2, <https://ilostat.ilo.org/topics/covid-19/>, 24 Ocak 2021.
- URL-3, <https://ilostat.ilo.org/about/data-collection-and-production/>, 24 Ocak 2021.
- URL-4, <https://worldhealthorg.shinyapps.io/covid/> , 25 Ocak 2021.
- URL-5, <https://covid19.saglik.gov.tr/> , 26 Ocak 2021
- URL-6, <https://disease-ontology.org/about/>, 11 Mart 2021.

- URL-7, <http://purl.org/whonto/onto>, 15 Nisan 2021.
- URL-8, <https://bioportal.bioontology.org/ontologies/COVID19>, 16 Nisan 2021
- URL-9, <https://github.com/oeg-upm/drugs4covid19-kg>, 16 Nisan 2021
- URL-10, <http://covid19.squirrel.link/ontology/>, 17 Nisan 2021
- URL-11, <https://bioportal.bioontology.org/ontologies/COVIDCRFRAPID>, 17 Nisan 2021
- URL-12, <https://knowledge-graph-hub.github.io/kg-covid-19/>, 17 Nisan 2021
- URL-13, <https://zenodo.org/record/3765375#.YhuSjNVBzIX>, 18 Nisan 2021
- URL-14, https://tr.wikipedia.org/wiki/Anlamsal_a%C4%9F, 19 Nisan 2021
- URL-15, <https://prateekvjoshi.com/2014/02/22/url-vs-uri-vs-urn/>, 20 Nisan 2021.
- URL-16, <https://aysedemirel.medium.com/url-urn-uri-12c21ad0dc20>, 21 Nisan 2021.
- URL-17, <https://tr.wikipedia.org/wiki/XML>, 22 Nisan 2021.
- URL-18, <https://www.w3.org/MarkUp/html3/HTMLandSGML.html>, 23 Nisan 2021.
- URL-19, https://www.dijitalders.com/icerik/1335/php_ve_yerlesik_dom_siniflari.html, 24 Nisan 2021.
- URL-20, <https://docs.stardog.com/tutorials/rdf-graph-data-model>, 25 Nisan 2021.
- URL-21, <https://www.techtarget.com/searchapparchitecture/definition/Resource-Description-Framework-RDF>, 26 Nisan 2021.
- URL-22, <https://dvc.w3.org/hg/rdf/raw-file/default/rdf-primer/index.html#section-rdfs-a>, 16 Nisan 2021.
- URL-23, [https://www.wikiwand.com/en/Turtle_\(syntax\)#/google_vignette](https://www.wikiwand.com/en/Turtle_(syntax)#/google_vignette), 27 Nisan 2021
- URL-24, <https://www.wikiwand.com/en/N-Triples>, 28 Nisan 2021.
- URL-25, <https://www.json.org/json-en.html>, 29 Nisan 2021.
- URL-26, <https://book.validatingrdf.com/bookHtml008.html>, 29 Nisan 2021.
- URL-27, <https://www.ionos.co.uk/digitalguide/websites/website-creation/tutorial-rdfa-markup-with-schemaorg/>, 30 Nisan 2021.
- URL-28, <https://www.w3.org/TR/rdfa-lite/>, 18 Nisan 2021, 1 Mayıs 2021.
- URL-29, https://www.w3.org/TR/?tag=data#w3c_all, 2 Mayıs 2021.
- URL-30, <https://www.erkyazilim.com.tr/icerik/324/ontoloji>, 3 Mayıs 2021.

- URL-31, http://www-inf.it-sudparis.eu/~gaaloulw/KM/courses/2-Course_KM_RDF.pdf, 4 Mayıs 2021
- URL-32, https://en.wikipedia.org/wiki/Simple_Knowledge_Organization_System, 5 Mayıs 2021
- URL-33, <https://www.w3.org/TR/skos-primer/>, 6 Mayıs 2021.
- URL-34, <https://www.w3.org/TR/daml+oil-reference/>, 7 Mayıs 2021
- URL-35, <https://tr.wikipedia.org/wiki/OWL>, 8 Mayıs 2021
- URL-36, <https://www.w3.org/TR/owl-primer/>, 9 Mayıs 2021.
- URL-37, <https://www.w3.org/TR/owl2-profiles/>, 10 Mayıs 2021.
- URL-38, <https://www.w3.org/TR/rdf-sparql-query/> 11 Mayıs 2021.
- URL-39, <https://www.w3.org/standards/semanticweb/query>, 12 Mart 2020.
- URL-40, <https://www.w3.org/TR/sparql11-update/>, 12 Mayıs 2021.
- URL-41, <https://dbpedia.org/sparql>, 13 Mayıs 2021.
- URL-42, <https://data.ordnancesurvey.co.uk/datasets/os-linked-data/explorer/sparql>, 14 Mayıs 2021.
- URL-43, <https://pitt.libguides.com/metadatadiscovery/linked-data>, 15 Mayıs 2021
- URL-44, <https://ontola.io/what-is-linked-data/>, 16 Mayıs 2021.
- URL-45, https://www.bulut-kbs.gov.tr/Giris/content/web/dosya/WebCBS_HTML_V_1.0/veri_cesitleri.htm, 17 Mayıs 2021
- URL-46, https://webdosya.csb.gov.tr/db/cbs/icerikler/metaveri_rehber_dokumani-20210421111339.PDF, 17 Mayıs 2021.
- URL-47, <https://setav.org/assets/uploads/2020/04/P266.pdf>, 18 Mayıs 2021.
- URL-48, <https://covid19data.link/> , 21 Nisan 2021
- URL-48, <https://github.com/fhamdi/GeomRDF>, 19 Mayıs 2021.
- URL-49, <https://github.com/GeoKnow/TripleGeo>, 20 Mayıs 2021.
- URL-50, <https://openrefine.org/documentation.html>, 21 Mayıs 2021.
- URL-51, <https://openrefine.org/>, 22 Mayıs 2021.
- URL52, <https://github.com/usc-isi-i2/Web-Karma>, 23 Mayıs 2021.
- URL-53, <https://virtuoso.openlinksw.com/> , 24 Mayıs 2021.

- URL-54, <https://jena.apache.org/documentation/fuseki2/>, 25 Mayıs 2021.
- URL-55, [https://www.wikiwand.com/en/Turtle_\(syntax\)#/google_vignette](https://www.wikiwand.com/en/Turtle_(syntax)#/google_vignette), 26 Mayıs 2021.
- URL-56, <http://d2rq.org/d2r-server>, 27 Mayıs 2021.
- URL-57, <https://github.com/4store/4store>, 28 Mayıs 2021.
- URL-58, <https://www.swirrl.com/>, 29 Mayıs 2021.
- URL-59, <http://silkframework.org/>, 31 Mayıs 2021.
- URL-60, <https://aksw.org/Projects/LIMES.html>, 1 Haziran 2021.
- URL-61, <http://dblab.cs.toronto.edu/project/xcurator/>, 2 Haziran 2021.
- URL-62, <https://datos.gob.es/es/node/7358>, 3 Haziran 2021.
- URL-63, <https://ckan.org/>, 4 Haziran 2021.
- URL-64, <https://datahub.io/>, 5 Haziran 2021.
- URL-65, <https://lod-cloud.net/>, 6 Haziran 2021.
- URL-66, <https://www.w3.org/2001/sw/wiki/Strabon>, 7 Haziran 2021.
- URL-67, <https://rdfshape.herokuapp.com/>, 8 Haziran 2021.
- URL-68, <http://vapour.sourceforge.net/>, 9 Haziran 2021.
- URL-69, <https://www.w3.org/RDF/Validator/>, 10 Haziran 2021.
- URL-70, <https://turcovid19.com/etkinlikler/turkiye-toplam-vaka-sayisi-analizi/>, 17 Haziran 2021.
- URL-71, <http://tuikapp.tuik.gov.tr/Bolgesel/sorguSayfa.do?target=tablo>, 18 Haziran 2021.
- URL-72, <https://turcovid19.com/tribbs1/>, 19 Haziran 2021.
- Valle, E. D. ve Ceri, S., 2011. Querying the Semantic Web, In Handbook of Semantic Web Technologies, Domingue, J., Fensel, D. ve Hendler, J. A., First Edition, 1035, Springer-Verlag Berlin Heidelberg.
- Valsecchi, F. ve Ronchetti, M., 2014. Spacetime: a two dimensions search and visualisation engine based on linked data, In Eight International Conference on Advances in Semantic Processing (SEMAPRO 2014), August, Italy, 8-12.
- Valsecchi, F., Abrate, M., Bacciu, C., Tesconi, M. ve Marchetti, A., 2015. Dbpedia atlas: Mapping the uncharted lands of linked data, In Proceedings of the Workshop on Linked Data on the Web, LDOW 2015, co-located with the 24th International World Wide Web Conference (WWW 2015), January, Italy.

- Wang, H., Du, H., Qi, G., Chen, H., Hu, W. ve Chen, Z., 2022. Construction of a Linked Data Set of COVID-19 Knowledge Graphs: Development and Applications. JMIR Medical Informatics, 10, 5, e37215.
- Wood, D., Zaidman, M., Ruth, L. ve Hausenblas, M., 2014. Linked Data, First Edition, 336, Manning Publications Co.
- Xue, X., Wang, Y. ve Hao, W., 2015. Optimizing Ontology Alignments by using NSGA-II, International Arab Journal of Information Technology (IAJIT), 12, 2, 176-182.
- Yu, L., 2011. A developer's guide to the semantic Web, First Edition, 608, Springer – Verlag Berlin Heidelberg.



ÖZGEÇMİŞ

2012 yılında Tonya Çok Programlı Anadolu Lisesi'ni bitirmiştir. 2015-2016 eğitim-öğretim yılında Farabi Değişim Programı ile Karadeniz Teknik Üniversitesi Harita Mühendisliği Bölümü'nde okumuştur. 2016 yılında Aksaray Üniversitesi Harita Mühendisliği Bölümü'nü bitirmiştir. 2018 yılında Karadeniz Teknik Üniversitesi Fen Bilimleri Enstitüsü, Harita Mühendisliği Anabilim Dalında yüksek lisans programına başladı.

