

PAPER DETAILS

TITLE: İdari Sınırlar Verisinin Bağlantılı Açık Veri Olarak Yayınlanması

AUTHORS: Gülten KARA, Çetin CÖMERT

PAGES: 171-185

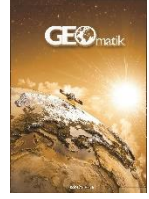
ORIGINAL PDF URL: <https://dergipark.org.tr/tr/download/article-file/1087311>



GEOMATİK

<https://dergipark.org.tr/tr/pub/geomatik>

e-ISSN 2564-6761



İdari Sınır Verilerinin Bağlantılı Açık Veri Olarak Yayımlanması: Trabzon Sınırları Örneği

Gülten Kara*¹, Çetin Cömert¹

¹Karadeniz Teknik Üniversitesi, Mühendislik Fakültesi, Harita Mühendisliği Bölümü, Trabzon, Türkiye

Anahtar Kelimeler

Bağlantılı veri
Açık veri
Bağlantılı açık veri
semantik web

ÖZ

Bağlantılı Veri, Verinin Webi olarak bilinen Semantik Web' in temel yapı taşlarından biridir. "Bağlantılı Açık Veri" den söz edebilmek için "Açık Veri" ve "Bağlantılı Veri" kavramlarının anlaşılması gerekir. Veriler herkes tarafından serbest bir şekilde kullanılıp paylaşıyorsa bu verilere Açık Veri denir. Bağlantılı veri, RDF verisi arasındaki linkleri kurmak için en iyi uygulamaları sunar. Diğer bir deyişle, Bağlantılı veri, web üzerinde bilgisayarlar tarafından okunabilir bağlantılı verilerin paylaşımı için tasarım ilkeleri gerektirir. Son yıllarda "Açık Veri Yaklaşımı" ile birlikte web üzerindeki konumsal verilerin boyutu artmıştır. "Bağlantılı Açık Veri Yaklaşımı" ile daha da büyük artış göstereceği öngörülmektedir. Konumsal alanda bağlantılı açık verilerin yayınlanması, mevcut veri ve bilgilerden yeni bilgiler çıkarılmasını ve farklı konumsal veri sağlayıcıları arasında veri paylaşımını semantik olarak gerçekleştirilmesini sağlayacaktır. Belirli bir alandaki verilerin bağlantılı veri olarak yayınlanması için gerekli veri setlerinin bulunması, veri setlerinin yeniden kullanılabilirliğinin önündeki en büyük engeldir. Konumsal verilerin daha anlamlı hale getirilmesi, link kurulacak veri setlerinin bulunması ve eldeki konumsal verilerle veri setleri arasındaki ilişkilerin belirlenmesi ile gerçekleştirilir. Açıklanan bu bilgiler bağlamında çalışmanın ana amacı, konumsal verilerin bağlantılı açık veri" olarak yayınlanması için gerekli metodolojinin belirlenmesidir. Bu kapsamda Trabzon idari sınırlar veri seti bağlantılı açık veri olarak yayınlanmıştır.

Publishing Administrative Units Data as Linked Open Data

Keywords

Linked data
Open data
Linked open data
Semantic web

ABSTRACT

Linked data is one of the basic building blocks of the Semantic Web, known as the Web of Data. In order to talk about Linked Open Data, it is necessary to understand the concepts of Open Data and Linked Data. If the data is freely used and shared by everyone, the data is called Open Data. Everyone can use Open Data without linking other datasets. Linked Data requires many design principles for sharing data that can be read by computers on the web. In recent years, the size of spatial data on the web has increased with the Open Data Approach. It is foreseen that it will show an even more significant increase with the Linked Open Data Approach. The publication of Linked Open Data in a spatial domain will allow the extraction of new information from existing data and the data sharing between different spatial data providers to be performed semantically. The Linked Open Data Approach provides an alternative way to merge and share data on the web. The finding the datasets that are required to publish data in a specific area as linked data is the biggest obstacle to the reuse of datasets. In order to make the spatial data more meaningful, it is necessary to find the datasets to be linked and to determine the relationships between the spatial data and datasets. The focus of the study is to determine the methodology required for the publication of spatial data as Linked Open Data. In this context, the administrative boundary dataset was published as Linked Open Data.

1. GİRİŞ

Bağlantılı veri, web üzerinde yapılandırılmış verileri ilişkilendirmek ve yayınlamak için en iyi uygulamalar anlamına gelir (Bizer, Heath & Berners-Lee, 2009; Heath & Bizer 2011). Bağlantılı Veri Teknolojileri web üzerinde yapılandırılmış veriyi yayınlamak için RDF (Resource Description Framework) dilini ve HTTP (HyperText Transfer Protocol) kullanır. Son yıllarda birçok alanda kullanılan verilerin Bağlantılı Veri İlkeleri' ne göre yayınlanmasının yanı sıra çok sayıda kitle kaynaklı girişim gerçekleştirilmiştir. Bunların başında DBpedia¹ gelir. DBpedia, Wikipedia dan yapısal bilginin çıkarılmasıyla oluşturulan bağlantılı veri setlerinden biridir. Bir diğer bağlantılı veri girişimi ise; GeoNames² yer adları dizinidir. GeoNames, 8 milyon yerin bağlantılı verisini yayınlamış olan açık lisanslı konumsal veri tabanıdır. Bu veri tabanı için GeoNames Ontolojisi³ oluşturulmuştur. Dünyanın büyük bir bölümünü kapsayan OpenStreetMap⁴, gönüllülerin işbirliğiyle konumsal bilgi tabanı oluşturmak amacıyla oluşturulan açık bir ortamdır. OpenStreetMap verisi bağlantılı veri olarak yayınlanmıştır. Özellikle bağlantılı veri yaklaşımı ile web üzerinde verileri bağlantılı veri olarak yayınlamak için veri setlerinin geliştirildiği Bağlantılı Açık Veri Bulutu⁵ (Linked Open Data Cloud) şüphesiz en önemli girişimlerden birisidir.

W3C RDB2RDF⁶ (Relational Databases to RDF) Çalışma Grubu herhangi bir ortak model kullanmaksızın ilişkisel veri tabanındaki tablolardan otomatik olarak RDF üretmek için Doğrudan Eşleştirmeyi (Direct Mapping) ve kelime hazinesi kavramlarını veri tabanı şemasına ilişkilendirmek için R2RML⁷ (RDB to RDF Mapping Language) önerir. R2RML W3C Tavsiyesi ilişkisel tabloların RDF ye eşleştirilmesini, RDF veri modelinin görüntülenmesini ve sorgulanmasını sağlamak için bir RDF notasyonu tanımlar (W3C, 2012). İlişkisel veri tabanlarını RDF ye dönüştürmek için çok sayıda yazılım aracı mevcuttur. Bunlardan bazıları, Triplify (Auer, Dietzold, Lehmann, Hellmann, & Aumüller, 2009), D2R Server⁸, Virtuoso RDFizer Middleware⁹, OpenRefine¹⁰ RDF genişletmesi, vb. dir. Bu yazılım araçlarının yanı sıra farklı dosya formatlarındaki verileri RDF' ye dönüştürmek için çok sayıda yazılım geliştirilmiştir¹¹. Ancak belirtilen yazılım araçlarının hiçbirisinde konumsal veri depolama, sorgulama veya dönüştürme desteği yoktur.

Bağlantılı veri en genel tanımıyla farklı formatlardaki verilerin URI (Uniform Resource Identifier) kullanılarak web üzerindeki veri setleri ile arasında link kurmaktır. Link kurmak hem kaynak veriyi hem de web üzerinde link kurulacak veri

setlerini zenginleştirir. Verinin Web' i üzerinde veri ile ilişkili olan veri setlerinin bulunarak ilişkilendirilmesi anlamına gelen link kurmak, bağlantılı verinin en temel sorunudur. Web üzerinde link kurulacak veri setlerinin kullanıcı tarafından bilinmesi gerekir. Bu bağlamda web üzerinde bağlantılı verilerin yayınlanması için gerekli veri setlerinin oluşturulması amacıyla Bağlantılı Açık Veri Bulutu oluşturulmuştur. Bağlantılı Açık Veri Bulutu üzerinde farklı veri setleri bağlantılı veri olarak yayınlanmıştır ve veri setleri birbiri ile ilişkilendirilmiştir. Mart 2019 itibarıyla Bağlantılı Açık Veri Bulutu diyagramında bağlantılı veri olarak yayınlanan veri setlerinin sayısı 1239 a ulaşmıştır. Bu veri setleri arasında 16.147 link kurulmuştur.

Bağlantılı verilerin yayınlanmasında temel gereksinimlerden biri web üzerinde uygun veri setlerinin bulunmasıdır. Bu amaçla bağlantılı veri tarayıcıları, ontoloji arama motorları, Semantik Web arama motorları vb. kullanılabilir. Verinin Web' i üzerinde linkleri bulmak ve eşleştirmek amacıyla SILK, LİMES, vb. geliştirilen araçlar mevcuttur. Web veri kaynakları arasında veri linklerini bulmak ve sağlamak için geliştirilen SILK ile kullanıcılar, veri kaynaklarındaki veriler ile web üzerindeki bağlantılı veri setleri arasında linkler kurar. Böylece hem veri kaynağındaki veriler web üzerindeki veri setleri ile hem de web üzerindeki veri setleri farklı veri kaynaklarında yer alan verilerle zenginleştirilerek yayınlanmış olur. Bağlantılı verilerin yayınlanmasından sonra bağlantılı veriler; yeniden referanslandırılabilir URI'lerle (dereferenceable URI-http protokolü kullanarak bir URI aranabiliyorsa ve veri kaynağının tanımı alınabiliyorsa bu URI yeniden referanslandırılabilir URI olarak adlandırılır) mevcut veri setleri, RDF Endpoint ve RDF Dump olmak üzere üç şekilde sunulur. Bağlantılı veriyi sorgulamak için SPARQL¹², tanımlamak için RDF standartları kullanılır. Bağlantılı verinin amacı, veriyi web üzerinde işlemek için RDF modeline dayanarak web üzerindeki bağlantılı veri setleri ile aralarındaki linklerin kurularak yayınlanması ve bu verilerin farklı Endpoint (uç nokta) aracılığıyla SPARQL ile sorgulanmasını sağlamaktır.

Kitle kaynaklı konumsal verilerin gerçekleştirilen girişimlerle web üzerinde sunulması, açık konumsal veri anlayışının yaygınlaşmasına öncülük etmiştir. Buna paralel olarak konumsal gösterim ve Bağlantılı Veri Yaklaşımına göre farklı veri kaynaklarının semantik tanımları kullanılarak çıkarsama gerçekleştirilmesi Semantik Web uygulamaları için son derece önemlidir.

¹ <http://wiki.dbpedia.org>

² <http://www.geonames.org>

³ <http://www.geonames.org/ontology/documentation.html>

⁴ <https://www.openstreetmap.org/#map=7/39.031/35.252>

⁵ <https://lod-cloud.net>

⁶ <https://www.w3.org/TR/2012/REC-rdb-direct-mapping-20120927/>

⁷ <https://www.w3.org/TR/r2rml/>

⁸ <http://d2rq.org/d2r-server>

⁹ <http://docs.openlinksw.com/virtuoso/virtuososponger.html>

¹⁰ <http://openrefine.org>

¹¹ <https://www.w3.org/wiki/ConverterToRdf>

¹² <https://www.w3.org/TR/rdf-sparql-query/>

Konumsal verilerin Semantik Web uygulamalarında kullanılması için web üzerinde veya konumsal veri tabanlarında tutulan verilerin web üzerindeki uygun bağlantılı veri setleri ile aralarında link kurulması gerekir. RDF ve OWL (Web Ontology Language)¹³ gibi semantik web dilleri ve SPARQL gibi sorgulama dilleri sayesinde konumsal verilerin paylaşımı ve yayınlanmasını kolaylaştırmak amacıyla konumsal verileri bağlantılı veri olarak yayınlamak için çok sayıda proje girişim ve akademik çalışma yapılmıştır. Önemli girişimlerden birisi, OpenStreetMap projesiyle elde edilen verinin bağlantılı veriye dönüştürülmesiyle oluşturulan LinkedGeoData¹⁴ dır. LinkedGeoData, OpenStreetMap projesiyle toplanan bilgileri kullanır. Linked Sensor Data, sensör verilerinin REST (Representational State Transfer) ve Bağlantılı Veri İlkeleri' nin birlikte kullanımıyla İngiltere Kanal Kıyı Gözlemevi (Channel Coastal Observatory) tarafından elde edilen verilerin bağlantılı veri olarak yayınlanmasını amaçlar (Page, De Roure, Martinez, Sadler, & Kit, 2009). GeoLinkedData (es)¹⁵ İspanya'da konumsal veri ile Semantik Web' i zenginleştirmek için açık bir girişimdir. Konumsal verilerin en temel kaynağını teşkil eden kadastral verileri dikkate alan Saavedra, Vilches-Blázquez, & Boada, (2014), ISO 19152 Arazi İdaresi Temel Modeli' ni (Land Administration Domain Model - LADM) (ISO, 2011) kullanarak kadastral alan için bir model geliştirmiştir. Bu bağlamda Bağlantılı Veri

İlkeleri' ne uygun olarak kadastral verilerin Bağlantılı Açık Veri Bulutu' nda ilgili veri setleri ile ilişkilendirilmesi ve bağlantılı veri olarak yayınlanması için çalışmalar gerçekleştirmiştir. Açık veri kapsamında ise; ilk girişim 9 Mayıs 2013'te ABD Başkanı Barack Obama tarafından imzalanan ve yürürlüğe giren yürütme kararı ile devlet verilerinin açık olması ve bilgisayarlar tarafından anlaşılabilir olması açıkça ifade edilmiştir. Hemen ardından 16 Mayıs 2013'te ise ABD Açık Veri Projesi (Project Open Data) başlatılmıştır. Bu proje, sağlık, enerji, iklim, eğitim, finans, kamu güvenliği, küresel gelişim gibi çok sayıda sektörü kapsamaktadır. 14 Ocak 2019 tarihinde ise ABD Açık Devlet Veri Kanunu (Open Government Data Act) yürürlüğe girmiştir. Açık veri bağlamında Tayvan, Avustralya, Büyük Britanya gibi ülkeler Küresel Açık Veri İndeksi (Global Open Data Index)¹⁶ nde ilk sıralarda yer almaktadır. Açık konumsal verilerin bağlantılı veri olarak yayınlanmasında ise Ordnance Survey, Büyük Britanya için Ordnance Survey Linked Data, Code Point Open Linked Data, Open Names Linked Data, 50K Gazetteer Linked Data ve Boundary Line Linked Data olmak üzere veri setlerini bağlantılı veri olarak yayınlamıştır. Ordnance Survey, konumsal verileri bağlantılı veri olarak yayınlayan ilk harita kurumudur. Yayınlanan bağlantılı veriler için gerekli ontolojiler Ordnance Survey tarafından Tablo 1'de görüldüğü gibi geliştirilmiştir ve kullanıcıların erişimine sunulmuştur.

Tablo 1. Ordnance Survey Ontolojileri (URL-1).

Ontology	Description	Imports
Postcode Ontology	An ontology describing the postcode geography in Great Britain.	Core, The administrative geography and civil voting area ontology
The administrative geography and civil voting area ontology	An ontology describing the administrative and voting area geography of Great Britain.	Geometry Ontology Spatial Relations Ontology
Geometry Ontology	An ontology to describe abstract geometries.	-
Spatial Relations Ontology	An ontology describing basic spatial relations.	GeoSPARQL
50k Gazetteer	An ontology describing the 50K Gazetteer.	-

Herhangi bir konumsal veri setini web üzerindeki veri setleri ile ilişkilendirmek, tüm veri sağlayıcılarına belirli bir formatta ve standartta dönüşüm yapmaksızın verilerini küresel veri uzayında yayınlama imkânı verir. Bütün bu girişimlere rağmen bağlantılı veri olarak yayınlanan konumsal veri, Konumsal Veri Altyapısı ve açık veri platformları ile kıyaslandığında çok sınırlı düzeydedir. Çünkü geleneksel konumsal veri formatlarından RDF formatına dönüşüm için sınırlı sayıda yazılım aracı mevcuttur. Ayrıca link kurmak için Bağlantılı Açık Veri Bulutu' ndaki konumsal veri

kaynaklarının yeterli sayıya ulaştırılması ve konumsal verilerin bağlantılı veriye dönüşümü için daha fazla yazılım aracının geliştirilmesi gerekir. Web üzerinde açık bağlantılı verileri paylaşmak için bir veri yönetimi platformu olan Datahub' da Ordnance Survey (OS), National Geographical Institute of Spain, U.S. Geological Survey ve daha birçok önemli konumsal veri sağlayıcıları tarafından yayınlanan konumsal bağlantılı veriler yer almaktadır. Bunun yanı sıra Bağlantılı Açık Veri Bulutu, çeşitli bağlantılı veri tarayıcıları veri setlerinin bulunmasında kullanılabilir. OGC

¹³ <https://www.w3.org/OWL/>

¹⁴ <http://linkedgeo.org/About>

¹⁵ <http://geo.linkeddata.es>

¹⁶ <https://index.okfn.org/place/>

GeoSPARQL (OGC, 2012) standardı için GeoSPARQL Ontolojisi¹⁷ geliştirilmiştir. GeoSPARQL Ontolojisi konumsal verilerin geometri tiplerini ve topolojik ilişkileri tanımlamak için kullanılabilir durumdadır. OGC GeoSPARQL standardına göre topolojik çıkarsama imkânı veren ontolojilerin geliştirilmesi için konumsal verileri tanımlayan ontolojilerin GeoSPARQL Ontolojisi ile ilişkilendirilmesi gerekir. İlgili literatürdeki çalışmaların esin kaynağı olduğu bu çalışmada Trabzon idari sınır veri setinin bağlantılı veri olarak yayınlanması için metodoloji önerilmiştir ve gereksinimler belirlenmiştir.

2. YÖNTEM

Bağlantılı veri kavramı, yapılandırılmış verilerin web üzerinde RDF veri modeli kullanılarak yayınlanması anlamına gelir. Bu yaklaşım, yapılandırılmış verilerin web üzerinde yayınlanması ve farklı kaynaklardan gelen verilerin arasında linklerin kurulması temeline dayanır. Bağlantılı veri yaklaşımının dört temel ilkesi vardır (Berners-Lee, 2006):

1. “Şey” leri (things) isimlendirmek için URI’ ler kullanılmalıdır.
2. Kullanıcıların bu isimleri arayabilmeleri için http URI’ leri kullanılmalıdır.
3. Kullanıcılar bir URI’ yi incelediğinde RDF, SPARQL gibi standartları kullanarak yararlı bilgiler elde etmelidir.
4. Veri hakkında daha fazla bilgiye ulaşmak için veriler ve ilgili bağlantılı veri setleri arasında linkler kurulmalıdır.

Bağlantılı Veri İlkeleri’ nin birinci ilkesi; sadece web dokümanlarının değil tüm gerçek dünya nesneleri ve soyut kavramların da URI referansları kullanılarak tanımlanmasını gerektirir. Bu ilkenin amacı, gerçek dünyadaki herhangi bir nesneyi veya kavramı içerecek şekilde webin kapsamını çevrimiçi kaynaklardan genişletmek olarak görülebilir. Geleneksel Web’ in genel erişim mekanizması HTTP protokolleridir. İkinci ilkesi ise, URI’ lerle tanımlanan gerçek dünya nesnelerine, kavramlara ve ilişkilere erişebilmek için HTTP protokolünün kullanımını gerektirir. Bu ilke, bir HTTP istemcisinin HTTP ile bir URI’ ye bakıp, bu URI ile tanımlanmış veri kaynağındaki verilere erişebilmesini sağlar. Farklı veri kaynaklarında farklı formatlarda veriler sunulabilir. Bir veri kaynağındaki veriler link kurulabilecek harici veri setleri ile RDF linkleri ile sınırsız sayıda ilişkilendirilebilir.

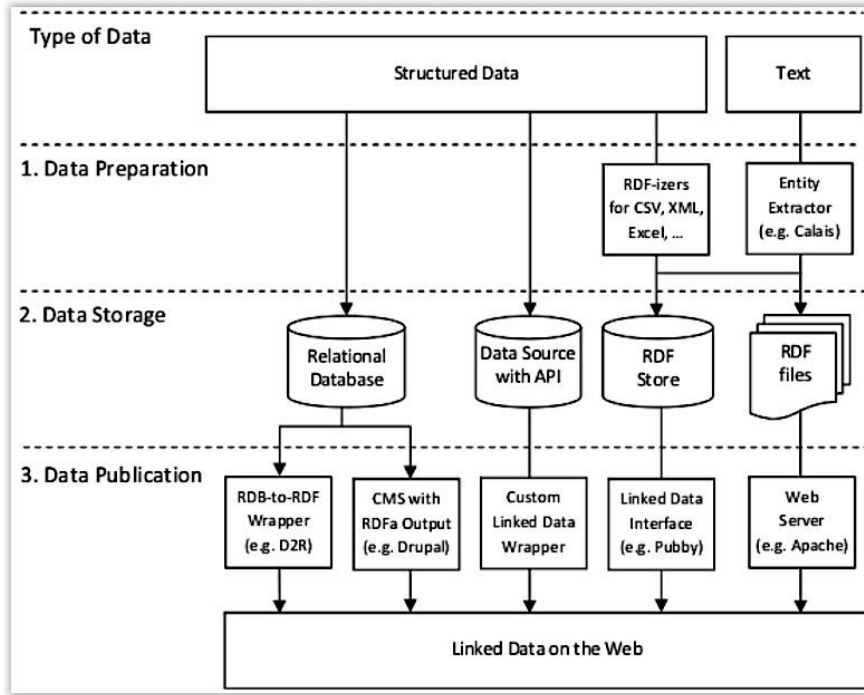
Bağlantılı veri yaklaşımının üçüncü ilkesi ise; yapılandırılmış verilerin web üzerinde yayınlanması için standart bir veri modelinin kullanılmasını öngörür. Bu veri modeli webin içeriğinin tanımlanmasında kullanılmak amacıyla geliştirilen basit bir çizge veri modeli olan RDF dir. RDF dokümanlarını sorgulamak için SPARQL dilinin kullanılmasını gerektirir.

Bağlantılı Veri İlkeleri’ nden dördüncüsü ise, veri kaynağındaki veriler ile web üzerindeki bağlantılı veri setleri arasında hiper linkler kullanılarak link kurulmasını ve bu linklerin ilişki tipinin tanımlanmasını gerektirir. Geleneksel Web’ de sadece web dokümanlarını ilişkilendiren hiper linkler, Semantik Web ve bağlantılı veri yaklaşımında; veri veya bilginin ilgili olduğu düşünülen her şeyin ilişkilendirilmesinde kullanılır. Bağlantılı veri kapsamında bu linkler *RDF linki* olarak adlandırılır. RDF linkleri RDF üçlüsü olarak ifade edilir geleneksel Web’ de kullanılan hiper linklerin aksine tipli linklerdir (Heath ve Bizer 2011). Bağlantılı veri kapsamında kullanılan RDF linkleri bu yönüyle Geleneksel Web’ de kullanılan hiper linklerden ayrılır. Bu bağlamda, iki farklı veri kaynağındaki veriler arasında kurulan linkler *rdf:type*, *owl:sameAs* gibi ilişkinin türünü belirlemek için kullanılır. Belirtilen RDF linkleri (predicate) sayesinde herhangi bir veri kaynağındaki veriyi ifade eden URI (subject), bağlantılı veri setindeki veriyi ifade eden URI’ ye (object) işaret eder. Böylece kaynak veri seti ve hedef veri seti arasında RDF linki kurulmuş olur.

Semantik Web’ in temel yapı taşlarından biri olan bağlantılı veri, herhangi bir veri kaynağındaki verilerin web üzerinde URI’ ler ile tanımlanmasını, RDF veri modeliyle temsil edilmesini ve kavramların anlamlarını RDF linkleriyle genişletmek üzerine kurulmuştur.

Konumsal verilerin bağlantılı veri olarak yayınlanmasında izlenmesi gereken işlem adımları Heath & Bizer (2011) tarafından önerilmiştir. İlişkisel veri tabanlarında depolanan veriler bağlantılı veri olarak RDF paketleyicileri (RDF wrappers) sayesinde kolay bir şekilde yayınlanabilir. Bu araçlar kullanıcıların Bağlantılı Veri İlkeleri’ ne göre web üzerinde sunulan ilişkisel veri tabanlarından RDF çizgelerine eşleştirme kurallarını tanımlamalarına izin verir. Girdi verilerinin CSV, XML, Excel, vb. olması durumunda RDF dosyasını üreten veya doğrudan bir RDF deposuna yükleyen bir dönüşüm gerekir. Girdi verisinin RDF olması durumunda ise Bağlantılı Veri İlkeleri uygulanır ve web sunucusu kullanılarak web üzerinde yayınlanır. Ya da bir bağlantılı veri ara yüzüne sahip bir RDF deposuna yüklenir. Uygulama senaryosu kapsamında ESRI Shapefile formatındaki verilerin RDF dönüşümünün gerçekleştirilmesi için Şekil 1’de belirtilen işlem adımları esas alınarak verilerin RDF dönüşümü gerçekleştirilmiştir. Bu çalışmada, Semantik Web Teknolojileri kullanılarak Trabzon idari sınır verisi RDF formatına dönüştürülmüştür ve web üzerindeki bağlantılı veri setleri ile ilişkilendirerek idari sınır bağlantılı verileri üretilmiştir.

¹⁷ http://schemas.opengis.net/geosparql/1.0/geosparql_vocab_all.rdf



Şekil 1. Bağlantılı Veri Yayınlama Yöntemleri ve İş Akışı (Heath & Bizer 2011).

2.1. Verilerin RDF Formatına Dönüştürülmesi

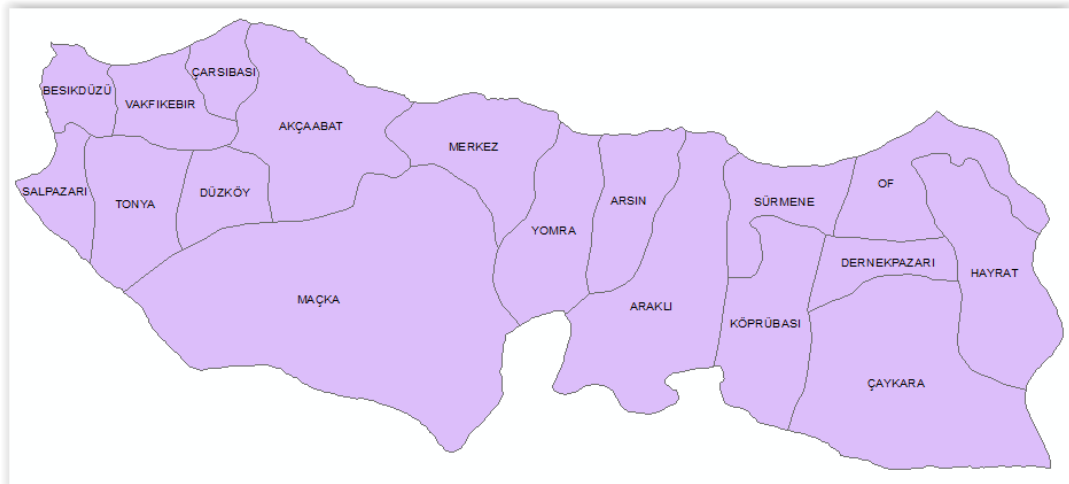
Bağlantılı verilerin yayınlanması için öncelikle verilerin RDF formatına dönüşümünün gerçekleştirilmesi gerekir. Veriler farklı formatlarda depolanmaktadır. Farklı veri formatlarının RDF dönüşümü için çok sayıda yazılım geliştirilmiştir. Bunlardan bazıları; EasyRDF, Csv2JSON, RDF123, Karma, Open Refine (spreadsheets, tabular data), W3C R2RML (veri tabanı), GATE, Zemanta, OpenCalais, DBoedia Spotlight (metin), vb. dir. RDF veri modelinin kullanımıyla farklı formattaki verilerden çıkarılan RDF verisi, özne (subject), nesne (object) ve yüklem (predicate) şeklinde ifade edilir. RDF veri modeli, web kaynaklarının tanımlanması için kavramsal bir çerçeve sunarak, farklı formattaki verileri herhangi bir alandan (domain) bağımsız olarak tanımlanmasını sağlamaktadır.

RDF üçlüsü Şekil 2’de belirtilen “Özne” ve “Nesne” düğüm, “Yüklem” iki düğümü birleştiren ilişki türü olmak üzere üç bileşenden oluşur. Düğümler veri seti içerisindeki kavramları veya varlıkları gösterir. İki düğüm arasındaki ilişki çizgi ile gösterilir. Bir düğüm URI ile boş düğüm (bir URI veya literal ile ifade edilmeyen) ile veya literal ile ifade edilebilir.



Şekil 2. RDF üçlüsü (W3C, 2004).

Uygulama senaryosu için Trabzon il sınırları içerisinde yer alan ilçe sınırları seçilmiştir (Şekil 3). Trabzon ilçe sınırları verisinin RDF formatına dönüştürülmesi gerekir.

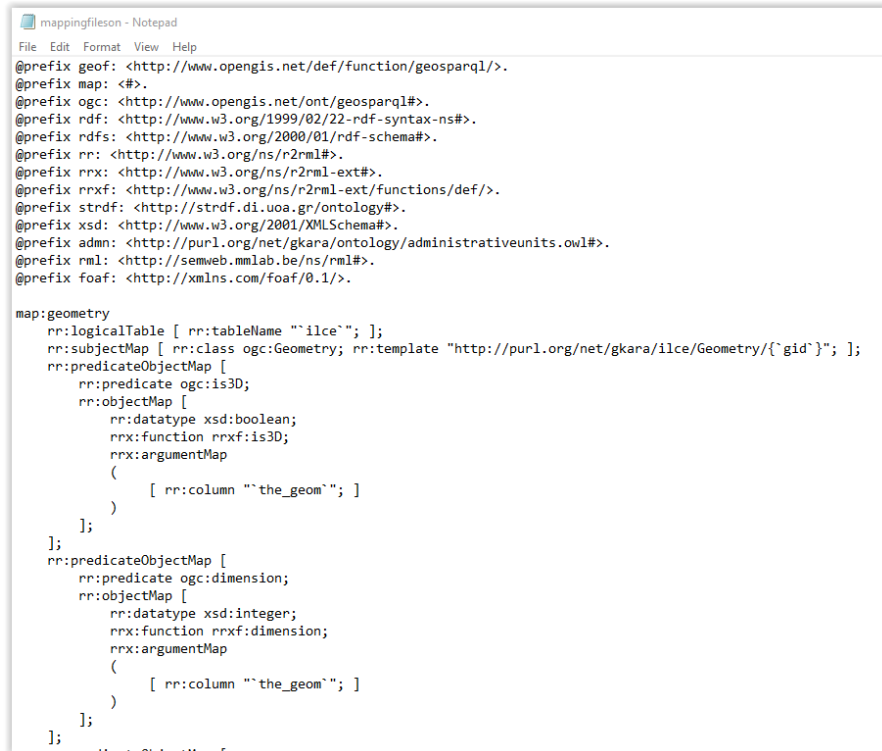


Şekil 3. Çalışma bölgesi: Trabzon ili ilçe sınırları

Konumsal verilerin RDF formatına dönüştürülmesi için mevcut yazılımlar incelendiğinde Geometry2RDF, shp2GeoSPARQL, GeomRDF, TripleGeo, GeoTriples gibi yazılımlar geliştirilmiştir. Uygulama senaryosu kapsamında konumsal verilerin RDF dönüşümünün gerçekleştirilmesi için konumsal özniteliklerin yanında konumsal olmayan özniteliklerin tamamını RDF formatına dönüşümünü desteklediği için GeoTriples yazılımı kullanılmıştır. Genel olarak isimlendirildiğinde veri çıkarma (data extraction) işleminin gerçekleştirildiği bu aşamada, HTTP URI'lerinin oluşturulması, kelime hazinesi (vocabulary) seçimi (GeoTriples - stSPARQL ve GeoSPARQL desteği) ve ilişkilendirilmesi işlemleri gerçekleştirilir. Bu aşamada gerçekleştirilen işlemler ile birlikte Bağlantılı Veri İlkeleri' nin 1. ve 2. maddesinin gereği yerine getirilmiş olur.

Uygulama senaryosu kapsamında GeoTriples yazılımı kullanılarak Trabzon idari sınırlar veri setinin ESRI shape formatından RDF formatına dönüştürülmesi gerekir. Bu dönüşüm eşleştirme dosyasının (mapping file) oluşturulması ile

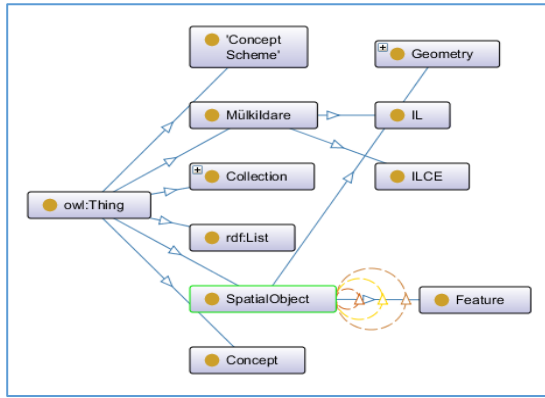
gerçekleştirilir. Eşleştirme dosyasında idari sınırlar ontolojisi ile ilişkilendirme bu aşamada gerçekleştirilir. Bunun için oluşturulan eşleştirme dosyasında kullanıcı tarafından gerekli değişikliklerin yapılabilmesi için R2RML (RDB to RDF Mapping Language) ve RML (RDF Mapping Language) dilleri kullanılır. RML, R2RML dilini genişleten farklı veri yapılarından RDF veri modeline göre tanımlanmış eşleştirme kurallarını ifade etmek için kullanılan bir eşleştirme dilidir. R2RML ise ilişkisel veri tabanlarından RDF veri modeline dönüşüm için gerekli eşleştirme kurallarını ifade etmek için kullanılan bir dildir ve W3C standardıdır. RML, ilişkisel veri tabanları, XML, JSON, CSV gibi birçok formatta sunulan verileri girdi olarak alır ve RDF üçlülerinin oluşturulması için gerekli eşleştirme kurallarını tanımlar. RML, RDF kullanılarak yapılandırılmıştır ve böylece girdi verilerinin alınarak RDF veri modeline göre eşleştirme kurallarının ifade edilmesini sağlar. Şekil 4'te gösterildiği gibi gerekli eşleştirme kuralları tanımlanarak eşleştirme dosyası oluşturulmuştur. Sonrasında RDF verisi elde edilir.



Şekil 4. GeoTriples eşleştirme dosyası

Konumsal verilerin bağlantılı veri olarak yayınlanması için konumsal veri tabanlarındaki verilerin, ontolojilerdeki ilgili sınıf ve özniteliklerle ilişkilendirilmesi gerekir. Diğer bir deyişle ontoloji mimarisinde veri ontolojisi olarak adlandırılan düzeyde veri tabanı şemasını tanımlayan ontoloji ile veri kaynağındaki tablo ve öznitelikler arasındaki eşleştirmeler tanımlanmalıdır. Bu nedenle veri ontolojisine ihtiyaç vardır. Ontoloji geliştirme aşamasında öncelikle mevcut ontolojilerin incelenmesi gerekir. Mevcut ontoloji bulunamaması

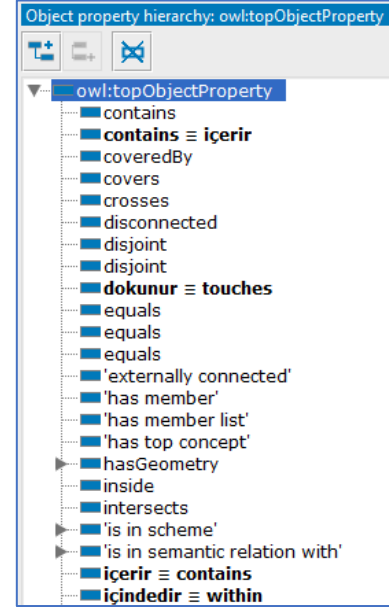
durumunda ontoloji yeni baştan geliştirilmelidir. Her veri seti için mevcut ontolojileri incelemeksizin yeni baştan ontoloji geliştirmek Semantik Web' in doğasına aykırıdır. Uygulama senaryosunda seçilen idari sınırlar veri seti için mevcut ontolojiler araştırılarak incelenmiştir ve ontoloji tanımlarının kullanılan standartlar ve idari alanların ve tanımlarının farklı olması sebebiyle Türkiye için İdari Birim Ontolojisi yeni baştan geliştirilmiştir (Şekil 5).



Şekil 5. TR İdari Birim Ontolojisi

Konumsal alanda konumsal varlıkları ve öznitelikleri ifade etmek için mevcut ontolojiler W3C GeoXG raporunda (Lieberman, Singh & Goad, 2007) verilmiştir. Bu rapor kapsamında ISO ve OGC standartlarının taslak ontolojileri, Ordnance Survey tarafından geliştirilen ontolojiler, Geonames ontolojisi, koordinat referans sistemleri Ontolojisi, Metaveri Ontolojisi gibi ontolojiler yer almaktadır. W3C Web Çalışma Grubu Sözleşmesi¹⁸ (Web Working Group Charter) gereği konumsal veriler ISO 19107 soyut modeli ile uyumlu ve GeoSPARQL, NeoGeo gibi yaygın olarak kullanılan bir konumsal ontoloji içermelidir ifadesi yer almaktadır. Bu nedenle çalışma kapsamında konumsal alanda detayların ve özniteliklerin tanımlanmasında yaygın olarak kullanılması sebebiyle GeoSPARQL ontolojisi seçilmiştir. İdari birim Ontolojisi' ndeki sınıf ve özelliklerin tanımları için GeoSPARQL Ontolojisi yeniden kullanılmıştır. TR İdari Birim Ontolojisi ve GeoSPARQL Ontolojisi Şekil 6' da verildiği gibi gerekli konumsal ilişkiler tanımlanarak Protégé

ontoloji editörü Nesne özelliği hiyerarşi sekmesi kullanılarak ilişkilendirilmiştir. İdari Birim Ontolojisi ve GeoSPARQL Ontolojisi sınıfları arasında ilişkilerin tanımlanması için owl:equivalentproperty ilişkisi kullanılmıştır. İdari Birim Ontolojisi http://purl.org/net/gkara/ontology/administrative_units.owl linkinde yayınlanmıştır.



Şekil 6. Nesne özelliği ilişkilendirme

İdari birimler veri setindeki sınıf ve öznitelik isimleri Türkiye İdari Birim Ontolojisi ile ilişkilendirilerek RDF formatında verinin oluşturulması için GeoTriples yazılımı kullanılmıştır. Bu aşamada kullanılan kelime hazineleri Tablo 2'de verilmiştir.

Tablo 2. Kullanılan kelime hazineleri

Vocabulary	IRI
Web Ontology Language – OWL*	http://www.w3.org/2002/07/owl#
Resource Description Framework – RDF*	http://www.w3.org/1999/02/22-rdf-syntax-ns#
Resource Description Framework Schema – RDFS*	http://www.w3.org/2000/01/rdf-schema#
eXtensible Markup Language – XML*	http://www.w3.org/XML/1998/namespace
Simple Knowledge Organization System – SKOS	http://www.w3.org/2004/02/skos/core#
Dublin Core – DC	http://purl.org/dc/elements/1.1/
Dublin Core Terms	http://purl.org/dc/terms/

*işareti ile gösterilen kelime hazineleri, ontoloji geliştirilmesi sırasında varsayılan olarak kullanılan kelime hazineleridir.

2.2. Link Kurulacak Veri Setlerinin Belirlenmesi

Bağlantılı verilerin farklı uygulamalar tarafından bulunabilir olması, web üzerinde standart veya yaygın kullanıma sahip veri setlerine link verilmesiyle sağlanır. Konumsal verilerin bağlantılı veri olarak yayınlanması için verilerin RDF dönüşümünden sonra link kurulacak veri setlerinin

belirlenmesi gerekir. Veri setlerini seçmek için mevcut araçlar ve kelime hazineleri aşağıdaki gibidir:

- LOD Cloud (Verilerin link kurulacağı veri setlerinin belirlenmesi için bir standart oluşturulması amacıyla Bağlantılı Açık Veri Bulutu oluşturulmuştur.)
- Linked Open Vocabularies¹⁹

¹⁸ <https://www.w3.org/2015/spatial/charter#bp>

¹⁹ <https://lov.linkeddata.es/dataset/lov/>

- Bağlantılı veri tarayıcıları: Tabulator²⁰, LodMilla²¹, LodLive²², CubicWeb²³, Disco²⁴, Tabulator²⁵, sameAs²⁶
- Ontoloji arama motorları (Swoogle, Watson, Schema.org, vb.)
- Bağlantılı veri mashupları Marbles²⁷, Revyu.com²⁸, DBpedia Mobile²⁹, SIGMA³⁰ olarak sıralanır.

Web üzerinde RDF olarak mevcut çok sayıda veri seti vardır. Bağlantılı veri olarak yayınlanacak veriler için uygun veri setlerinin bulunması gerekir. Bunun için öncelikle bilinen kelime hazineleri seçilmelidir. Veri setlerinin belirlenmesi aşamasında;

- Kelime hazineleri belirli bir alandaki kavramları ve kavramlar arasındaki ilişkileri modeller. İlgili alana yönelik kelime hazinelerinin belirlenmesi gerekir.
- Yaygın kullanılan kelime hazinelerindeki kavramlar mümkün olduğu kadar yeniden kullanılmalıdır.
- Mevcut kelime hazinelerinde uygulama senaryosu için gerekli kavramlar bulunamıyorsa yeni kavramlar tanımlanmalıdır.

Konumsal verileri tanımlamak amacıyla yaygın olarak kullanılan bağlantılı veri setleri aşağıda verilmiştir:

- Basic Geo (WGS84 lat/long) Vocabulary³¹
- OGC GML³²
- OGC GeoSPARQL³³
- OGC Simple Features³⁴
- NeoGeo Geometry Ontology³⁵
- NeoGeo Spatial Ontology³⁶
- OWL representation of ISO 19107 (Geographic Information - Spatial Schema)³⁷

- OWL representation of ISO 19109 (General Feature Model)³⁸
- GeoNames Ontology³⁹
- OS Spatial Relations Ontology⁴⁰
- OS Geometry Ontology⁴¹

Uygulama senaryosu kapsamında idari sınırlar veri setini tanımlamak amacıyla aşağıdaki veri setleri seçilmiştir.

- DBpedia (Endpoint URL-
http://dbpedia.org/sparql/)
- GeoNames (www.geonames.org/ontology#)
- Basic Geo
(https://www.w3.org/2003/01/geo/)

seçilmiştir.

2.3. RDF Linklerinin Kurulması

RDF veri setleri arasında link kurmak Bağlantılı Veri İlkeleri' nin gereğidir. Bağlantılı Veri İlkeleri' nden 4. sünde "kullanıcıların çok daha fazla bilgi ve veri bulmaları için diğer URI' lere link kurulmalıdır" ifadesi yer almaktadır. Bağlantılı veri tarayıcılarının veri setleri arasında gezinmelerini ve daha fazla veriler bulunmasını sağlar. Böylece küresel veri uzayı olarak adlandırılan çizge üzerinde semantik web teknolojileri kullanılarak veriler hakkında çok daha fazla bilgiye ulaşılır. Her bir düğüm ilgili olduğu tüm verilerle iç ve dış RDF linkleriyle ilişkilendirildiği için çizge üzerinde gezinerek gerekli tüm verilere ve bilgilere ulaşmak mümkündür. Çizge üzerinde gezinerek verilere nasıl ulaşıldığına örnek vermek gerekirse; Şekil 7' de belirtilen çizgede veri düğümleri ile ilişkiler RDF linkleri sayesinde linkler üzerinde gezindikçe çok daha fazla bilgi elde etmeye izin verir.

²⁰ <https://github.com/linkedata/tabulator>

²¹ <http://lodmilla.szaki.hu/lodmilla/>

²² <http://en.lodlive.it>

²³ https://fosdem.org/2019/schedule/event/collab_cwldbe/

²⁴ <http://wifo5-03.informatik.uni-mannheim.de/bizer/ng4j/disco/>

²⁵ <http://dig.csail.mit.edu/2005/ajar/release/tabulator/0.7/tab>

²⁶ <http://sameas.org>

²⁷ <http://mes.github.io/marbles/>

²⁸ <http://revyu.com>

²⁹ <https://wiki.dbpedia.org/projects/dbpedia-mobile>

³⁰ <http://sig.ma/>

³¹ <https://www.w3.org/2003/01/geo/>

³² http://schemas.opengis.net/gml/3.2.1/gml_32_geometries.rdf

³³ http://schemas.opengis.net/geosparql/1.0/geosparql_vocab_all.rdf

³⁴ http://schemas.opengis.net/sf/1.0/simple_features_geometries.rdf

³⁵ <http://geovocab.org/geometry>

³⁶ <http://geovocab.org/spatial>

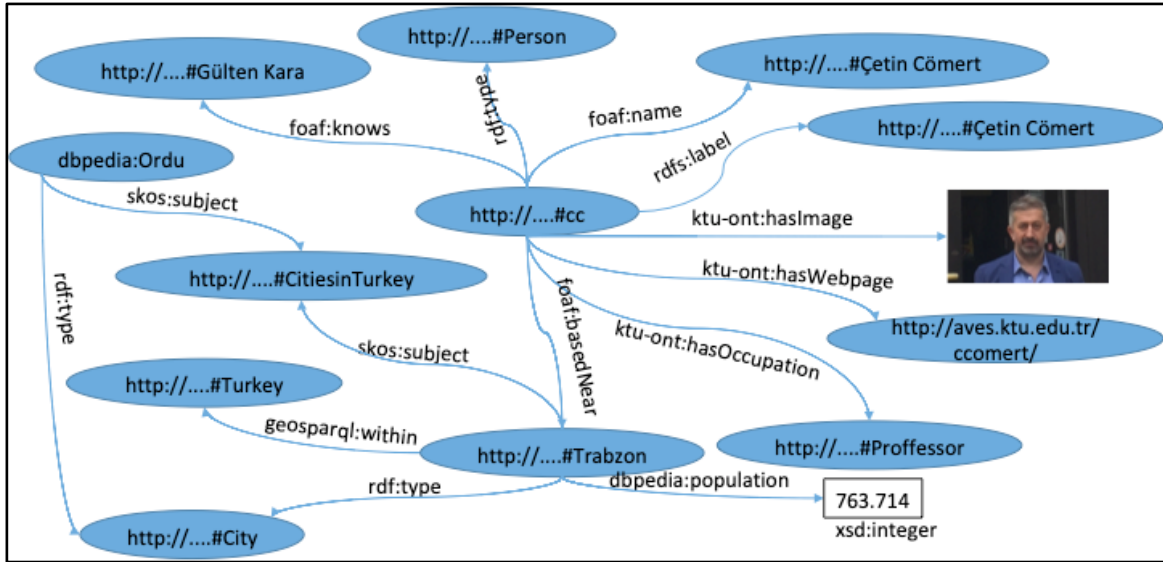
³⁷ <http://def.seegrid.csiro.au/isotc211/iso19107/2003/geometry>

³⁸ <http://def.seegrid.csiro.au/isotc211/iso19109/2005/feature>

³⁹ <http://www.geonames.org/ontology/documentation.html>

⁴⁰ <http://data.ordnancesurvey.co.uk/ontology/spatialrelations/>

⁴¹ <http://data.ordnancesurvey.co.uk/ontology/geometry/>



Şekil 7. RDF iç ve dış linkleri

Verilerin kapsamı diğer veri setlerinde referans verilecek özne, yüklem (predicate) ve nesne değerlerini belirler. Veriler ve diğer veri setleri arasında owl:sameAs yüklemi yaygın olarak kullanılmaktadır. owl:sameAs ilişkisi özne ve nesne tarafında bulunan iki URI referansının birbirinin aynısı olduğunu ifade ettiğini göstermek için kullanılır. Aynı zamanda aynı öznitelik değerlerini ilişkilendirmek için kullanılır. Bağlantılı Veri İlkeleri kullanılarak link kurulan veriler üzerinde gezindikçe çok daha fazla veriye ulaşılır ve hem konumsal detaylar hem de geometri tiplerini tanımlayan mevcut ontolojileri kullanarak konumsal veriler etkili bir şekilde modellenir.

Bağlantılı veri yaklaşımında kullanılan RDF linkleri iç RDF linkleri ve dış RDF linkleri olmak üzere ikiye ayrılır. İç RDF linkleri bir bağlantılı veri setinde bulunan veriler arasında kurulan linklerdir. Yani aynı ad alanına sahip özne ve nesne arasında kurulan linklerdir. Dış RDF linkleri ise bir veri seti ile farklı bağlantılı veri setleri arasında kurulan linklerdir ve farklı ad alanlarına sahiptir. Farklı bağlantılı veri setleri arasındaki linklerin kurulması için ilişki (Relationships) linkleri, kimlik (identity) linkleri ve kelime hazinesi (vocabulary) linkleri olmak üzere üç farklı link tipi kullanılır (Heath & Bizer, 2011).

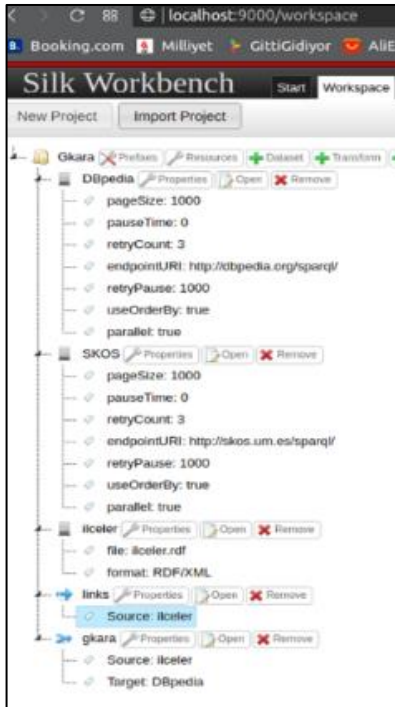
Veriler ve bağlantılı veri setleri arasındaki RDF linkleri elle ve otomatik olmak üzere iki yöntemle gerçekleştirilir. RDF linklerini elle oluşturmadan önce Bağlantılı Açık Veri Bulutu'nda yer alan veri setleri içerisinde verilerin link kurulacak veri setleri belirlenmelidir. Uygun veri setleri belirlendiğinde bağlantılı veri setlerinin URI referanslarını bulmak için bağlantılı veri tarayıcıları (Disco, Tabulator, Lodmilla, vb.) kullanılabilir. Bu yöntem, büyük veri setlerinin bağlantılı veri setleri ile ilişkilendirilmesinde kullanılamaz. Bu durumda RDF

linklerinin oluşturulması için otomatik eşleştirme algoritmaları kullanılır. Bağlantılı veri setleri, konumsal bağlantılı verilerin web üzerinde bulunabilir olmasında önemli rol oynar. Bu nedenle link kurulacak uygun veri setlerinin seçimi çok önemlidir. Link kurma kurallarına dayanarak link bulma işlemini gerçekleştirmek için kullanılan yazılım araçları SILK, Limes, xCurator, vb. gibi yazılım araçları geliştirilmiştir. Bağlantılı veri kapsamında geliştirilen yazılım araçları, gerçekleştirilen girişimler ve akademik çalışmalar⁴² Linking Open Data Projesi kapsamında incelenmiştir. İki farklı RDF veri seti arasında RDF linklerinin kurulması için aşağıdaki eşleştirme yöntemleri vardır:

Instance düzeyinde eşleştirme: doğrudan yeniden kullanım, owl:sameAs, rdfs:seeAlso
Şema düzeyinde eşleştirme: sınıf/öznitelik doğrudan yeniden kullanım, rdfs:subclass, rdfs:subproperty, owl:equivalent class, owl:equivalent property, SKOS broad match gibi link türleri vardır. Uygun linkler, link bulma (link discovery) yoluyla gerçekleştirilir. Konumsal veri sağlayıcıları kendi veri setleri ve web üzerindeki veri kaynakları arasında RDF linkleri kurarak bu linklerin takip edilmesiyle yeni bilgilerin bulunmasını sağlar.

Örneğin, owl:sameAs individual lar arasındaki linkleri oluşturur veya iki URI' nin aynı individual olduğunu ifade eder. rdfs:seeAlso bir kaynağın özne kaynak hakkında ek bilgi sağlayabileceği anlamına gelir. SKOS (Simple Knowledge Organization System) bilgi organizasyon sistemleri için geliştirilen veri modelidir. SKOS verisi RDF üçlülere olarak ifade edilir. SKOS eşleştirme özelliklerinin (skos:closeMatch, skos:exactMatch, skos:relatedMatch) kullanılmasıyla farklı veri setleri arasında RDF linklerinin kurulmasını sağlar.

Uygulama senaryosu kapsamında linklerin kurulması için SILK yazılımı seçilmiştir. 1. aşamada üretilen RDF verisi Şekil 8’ de belirtilen SILK Workbench yazılımına yüklenerek DBpedia, GeoNames ve Basic Geo veri kaynakları girdi kaynak ve hedef veri setleri tanımlanmış olur. Bağlantılı veri kaynaklarında bulunan verilerle link kurulması işlemi bu aşamada gerçekleştirilir. Böylece idari sınır RDF verisi konum verileri için GeoNames ile yer adları için DBpedia ile ilişkilendirilmiştir. Bu linklerin kullanılması ile idari sınır veri seti web üzerindeki bağlantılı veri kaynakları ile ilişkilendirilerek diğer veri kullanıcıları tarafından bulunabilirliği sağlanmıştır. Bu işlem adımı ile Bağlantılı Veri İlkeleri’ nden 4. Maddenin gereği yerine getirilmiş olur.



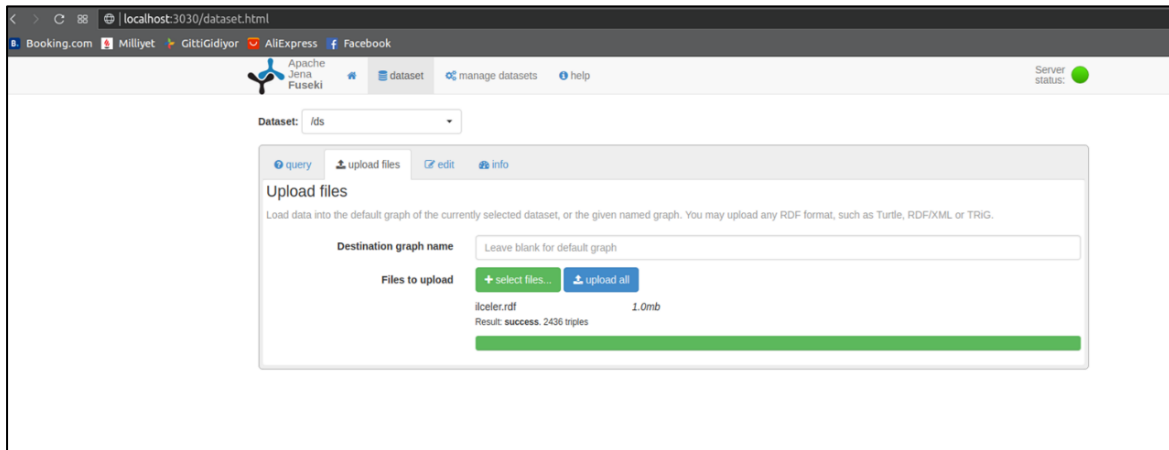
Şekil 8. SILK Workbench Eşleştirme İşleminin Tanımlanması

2.2. Bağlantılı Verinin Yayınlanması

Bağlantılı verilerin yayınlanması için bağlantılı veri hakkındaki meta verinin oluşturularak veriye erişimin sağlanması gerekir. Bağlantılı verilerin yayınlanması için mevcut sistemler Apache Jena, D2RQ, OpenLink Virtuoso Open-Source Edition , OpenLink Virtuoso Universal Server , vb. yazılım araçları geliştirilmiştir. Bu aşamada oluşturulan bağlantılı veri için meta verilerin oluşturulması ve bağlantılı verinin erişilebilir hale getirilmesi işlemleri gerçekleştirilir. Böylece Bağlantılı Veri İlkeleri’ nin 3. Maddesi de gerçekleştirilmiş olur. Web üzerinde RDF formatında sunulan veri setlerini sorgulamak için SPARQL Protokol ve RDF Sorgu Dili (SPARQL Protocol and RDF Query Language – SPARQL) kullanılır. OGC tarafından ek protokollerle genişletilen GeoSPARQL, SPARQL sorgulama diline konumsal sorgulama ifadeleri eklenmiştir. Bağlantılı verilerin yayınlanması ile herhangi bir veri kaynağındaki veriler farklı veri kaynaklarındaki kavramlarla aralarında semantik ilişkilerin kurularak Semantik Web Teknolojileri tarafından bilginin veya verinin kolayca anlaşılmasını bulunmasını ve paylaşılmasını sağlar. Böylece veri formatına ve belirli bir alana bağlı kalınmaksızın veriler ilgili oldukları veri kaynakları ile ilişkilendirilerek web üzerinde yayınlanmış olur.

Konumsal veri tabanlarındaki verilerin RDF dosyaları üzerinde sorgulamaların gerçekleştirilmesi için veri tabanlarında depolanması gerekir. Konumsal verilerden elde edilen RDF üçlülerini depolamak için Triple Store olarak adlandırılan veri tabanlarına ihtiyaç vardır. Konumsal RDF verilerini depolamak için etkin olarak kullanılan Triple Store lardan bazıları, AllegroGraph , Oracle Spatial and Graph , GraphDB , Parliament , Strabon , uSeekM , Virtuoso RDF Triple Store , Pubby vb. dir.

Uygulama senaryosu kapsamında verilerin yayınlanması ve sorgulanması için SPARQL Uç Noktalarından (SPARQL Endpoints) biri olan Fuseki kullanılmıştır. Şekil 9’da Fuseki ortamına dosya yükleme işlemi gösterilmiştir.



Şekil 9. Fuseki üzerinde dosya yükleme

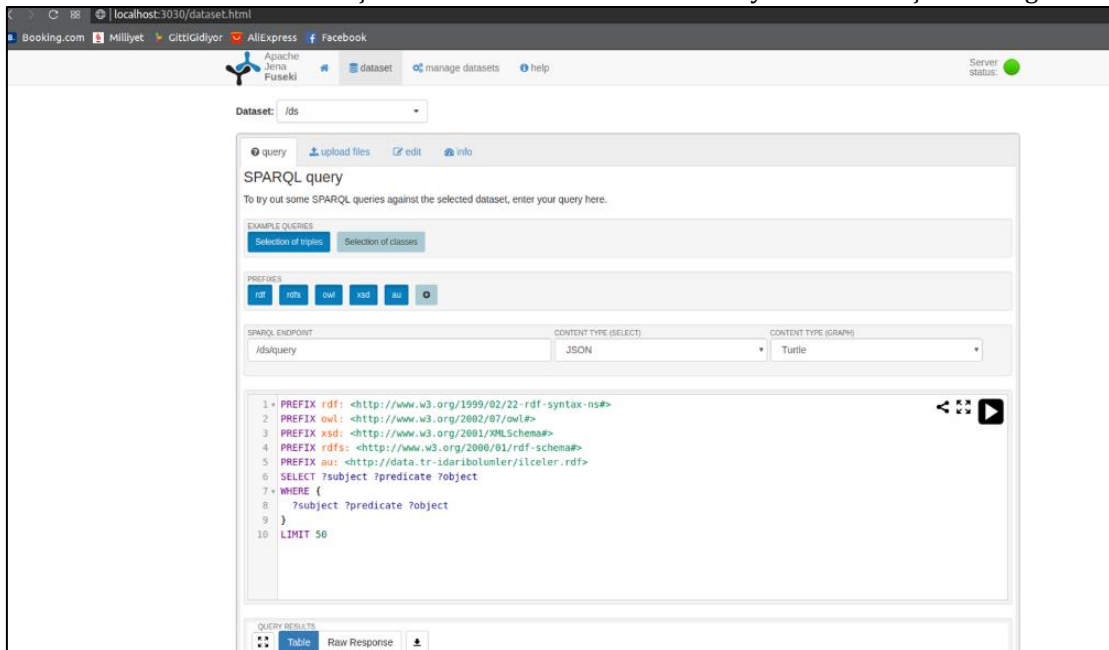
3. BULGULAR ve İRDELEMELER

Bağlantılı veri yaklaşımı herhangi bir veri kaynağında bulunan verilerin bağlantılı veri standartları temel alınarak web üzerinde yayınlanması anlamına gelir. Konumsal veri tabanlarında tutulan verilere ihtiyaç duyan uygulamalar, farklı veri kaynaklarında bulunan verileri arasında eşleştirme işlemlerinin yapılmasını gerektirir. Ayrıca web üzerinde sunulan konumsal verilerin her geçen gün daha da artması ile veri kaynakları arasında otomatik eşleştirme işlemlerinin anlamsal olarak gerçekleştirilmesi artık kaçınılmaz olmuştur. Konumsal alanda Semantik Web uygulamalarının hız kazanması ile birlikte Semantik Web Teknolojileri de paralel olarak gelişimini sürdürmektedir. Bu bağlamda konumsal alanda üretilen/sunulan verilerin Semantik Web uygulamalarında kullanılabilir hale getirilmesi gerekir. Bu nedenle konumsal verilerin anlamsal olarak tanımlanması ve web üzerinde yayınlanması anlamına gelen konumsal bağlantılı verilerin oluşturulması gerekir. Bu amaçla gerçekleştirilen bu çalışmada Trabzon idari sınır veri seti seçilerek bağlantılı verilerin oluşturulmasından yayınlanmasına kadar gerçekleştirilen tüm işlemler bu çalışmada ele alınmıştır. Öncelikle mevcut verilerin bağlantılı veri olarak yayınlanması için RDF formatına dönüşüm gerekir. Bu işlem adımıyla özellikle konumsal alanda mevcut yazılım araçları incelenmiştir. Mevcut yazılım araçlarından konumsal ve konumsal olmayan tüm özniteliklerin RDF dosyasına aktarılmasına imkân veren GeoTriples seçilmiştir. Ancak eşleştirme dosyasının oluşturulması aşaması, kullanıcının R2RML dilini kullanabilme yeteneğinin olmasını gerektirir. Ayrıca konumsal alanda kullanılabilecek veri kaynaklarını bilmesini gerektirir. Bu yazılım boyutu büyük olan verilerin RDF formatına dönüştürülmesinde

kullanıcı müdahalesi gerektirdiği için hata eğilimli olacaktır. Bu bağlamda bu aşamanın çok az kullanıcı müdahalesine ihtiyaç duyması ya da otomatik hale getirilmesi gerekmektedir. Ayrıca bu aşamada veri tanımlarının oluşturulması için ontoloji oluşturulmuştur. Bu ontoloji mevcut ontolojilerin yeniden kullanılmasıyla geliştirilmiştir ve ihtiyaç duyulması halinde geliştirilebilir. Veri ontolojisi olarak nitelendirilen bu ontoloji çalışma kapsamında geliştirilmiş olup RDF formatına dönüşüm aşamasında kullanılmıştır.

RDF formatına dönüştürülen verilerin bağlantılı veri olarak yayınlanması için mevcut veri setleri içinde link kurulacak veri setlerinin belirlenmesi gerekir. Bunun için çok sayıda bağlantılı veri tarayıcıları geliştirilmiş olsa da ilgili veri setlerinin bulunması son derece zordur. Bağlantılı veri tarayıcıları anahtar kelime tabanlı arama sonucunda veri setlerini listelemektedir. Bu nedenle gerçekte kullanılabilecek veri setleri anahtar kelime tabanlı arama ile bulunamayabilir. Bunun yanında veri setlerinin şemasının kullanıcı tarafından bilinmesi gerekir. Uygulama senaryosu kapsamında yaygın kullanılan veri setleri belirlenmiştir. Veri setlerinin belirlenmesinin ardından bu veri setleri ve RDF formatına dönüştürülen veriler arasında RDF linklerinin kurulması gerekir. Bu aşamada kurulan linkler owl:sameAs ilişkisidir ve SILK yazılımı ile gerçekleştirilmiştir. SILK üzerinde kaynak veri setinin ve hedef veri setlerinin belirlenmesinden sonra linkler otomatik olarak bulunmaktadır ve RDF olarak yayınlanmaktadır.

Bağlantılı veri setleri ile ilişkilendirilen idari birimler veri seti Fuseki ortamına aktararak SPARQL ile sorgulama gerçekleştirilmiştir. Şekil 10' da Fuseki ara yüzünde SPARQL sorgulama ekranı verilmiştir. Şekil 11' de ise Şekil 10' da verilen sorgu ile Trabzon idari birimler bağlantılı verisi üzerinde özne nesne yüklem RDF üçlülere sorgulanmıştır.



Şekil 10. Fuseki üzerinde SPARQL sorgulama

Sorgulama ara yüzünde PREFIX ön ekleriyle verilen veri setleri ve idari birimler veri seti üzerinde SPARQL kullanılarak sorgulama yapılmıştır. Sorgulama sonucunda idari birimlere ait bütün bilgiler listelenmiştir. İzleyen çalışmalarda

konumsal ilişkiler kullanılarak sorgulama yapılarak sorgu sonuçlarının harita üzerinde çevrimiçi olarak sunumu ve görselleştirilmesi ile ilgili çalışmalar yapılacaktır.

21	<http://data.linkedeodata.eu/iceSinirlari/21>	<http://data.linkedeodata.eu/ontology#has_SHAPE_Area>	"0.120136902867***xsd:double"
22	<http://data.linkedeodata.eu/iceSinirlari/21>	<http://data.linkedeodata.eu/ontology#has_NT2000>	"25865"
23	<http://data.linkedeodata.eu/iceSinirlari/21>	<http://data.linkedeodata.eu/ontology#has_KZ_SIRA>	"01"
24	<http://data.linkedeodata.eu/iceSinirlari/21>	<http://data.linkedeodata.eu/ontology#has_NS1997>	"12472"
25	<http://data.linkedeodata.eu/iceSinirlari/21>	<http://data.linkedeodata.eu/ontology#has_NT1990>	"21505"
26	<http://data.linkedeodata.eu/iceSinirlari/21>	<http://data.linkedeodata.eu/ontology#has_NS1990>	"11824"
27	<http://data.linkedeodata.eu/iceSinirlari/21>	<http://data.linkedeodata.eu/ontology#has_NK2000>	"11500"
28	<http://data.linkedeodata.eu/iceSinirlari/21>	<http://data.linkedeodata.eu/ontology#has_DSI_SUBE>	"226***xsd:integer"
29	<http://data.linkedeodata.eu/iceSinirlari/21>	<http://data.linkedeodata.eu/ontology#has_YOG2000>	"24***xsd:integer"
30	<http://data.linkedeodata.eu/iceSinirlari/21>	rdf:type	<http://data.linkedeodata.eu/ontology#iceSinirlari>
31	<http://data.linkedeodata.eu/iceSinirlari/21>	<http://data.linkedeodata.eu/ontology#has_URBAN_PO_P_>	"12436.0***xsd:double"
32	<http://data.linkedeodata.eu/iceSinirlari/21>	<http://data.linkedeodata.eu/ontology#has_NK1997>	"5449"
33	<http://data.linkedeodata.eu/iceSinirlari/21>	<http://data.linkedeodata.eu/ontology#has_YOG1990>	"20"
34	<http://data.linkedeodata.eu/iceSinirlari/21>	<http://data.linkedeodata.eu/ontology#has_TOTAL_PO_P_>	"18470.0***xsd:double"
35	<http://data.linkedeodata.eu/iceSinirlari/21>	<http://data.linkedeodata.eu/ontology#has_YOG1997>	"17***xsd:integer"
36	<http://data.linkedeodata.eu/iceSinirlari/21>	<http://data.linkedeodata.eu/ontology#has_KZ_KOD_1>	"2801"
37	<http://data.linkedeodata.eu/iceSinirlari/21>	<http://data.linkedeodata.eu/ontology#has_NT1997>	"17921"
38	<http://data.linkedeodata.eu/iceSinirlari/21>	<http://data.linkedeodata.eu/ontology#has_YUZOLCUM>	"1083"
39	<http://data.linkedeodata.eu/iceSinirlari/21>	<http://data.linkedeodata.eu/ontology#has_NS2000>	"14365"
40	<http://data.linkedeodata.eu/iceSinirlari/21>	<http://data.linkedeodata.eu/ontology#has_RURAL_PO_P_>	"6034.0***xsd:double"
41	<http://data.linkedeodata.eu/iceSinirlari/21>	<http://data.linkedeodata.eu/ontology#has_IL_KOD>	"28"
42	<http://data.linkedeodata.eu/iceSinirlari/21>	<http://www.opengis.net/ont/geosparql#hasGeometry>	<http://data.linkedeodata.eu/iceSinirlari/Geometry/21>
43	<http://data.linkedeodata.eu/iceSinirlari/21>	<http://data.linkedeodata.eu/ontology#has_NAME>	"ALUCRA"

Şekil 11. Fuseki üzerinde RDF üçlüsü sorgulama sonucu

4. TARTIŞMA

Semantik web teknolojileri, verilerin bilgisayar tarafından anlaşılabilir hale getirilerek yeni bilgiler çıkarılmasını sağlar. Bu teknolojinin birçok alanda kullanımı konumsal alanda bir paradigma kaymasına sebep olmuştur. Bununla birlikte konumsal veri altyapılarını oluşturmak için kullanılan WMS WFS gibi geleneksel teknolojilerle veri entegrasyonu ve veri arama işlemlerinin gerçekleştirilmesi çok zordur. Bu nedenle konumsal alanda semantik web teknolojilerinin avantajlarından yararlanarak farklı veri formatında sunulan verilerden yeni bilgiler çıkarmak için verilerin semantik tanımlarına ihtiyaç vardır. Bu çalışmada konumsal veri üreticilerinin verilerini web üzerinde yayınlamaları için bir yöntem önerilmiştir. Semantik web teknolojilerine dayanan bu yöntem, uygulama senaryosu kapsamında farklı veri formatındaki konumsal verilerin bağlantılı veri olarak yayınlanması için işlem adımları ile birlikte verilmiştir. Bu çalışmanın diğer çalışmalardan farkını ortaya koymak için literatürde ilgili kaynaklar belirlenerek karşılaştırmalar aşağıda özetlenmiştir.

Vilches-Blázquez, Villazón-Terrazas, Leon, Priyatna, & Corcho (2010), hidrografik verilerin bağlantılı veri olarak yayınlanması için yöntem önermiştir. Önerilen yöntemde veriler sadece

geliştirilen ontoloji ile ilişkilendirilmiştir ve harita üzerinde görselleştirilmiştir. Konumsal ilişkiler tanımlanmamıştır ve diğer bağlantılı veri kaynakları ile ilişkilendirilmemiştir. Bu çalışmada konumsal veriler geometri tanımları için GeoSPARQL ile ilişkilendirilmiştir. İdari sınır veri seti için gerekli bağlantılı veri kaynakları belirlenmiştir ve bu veri kaynakları ile linkler otomatik olarak bulunmuştur ve ilişkilendirilmiştir.

Barnaghi, Presser & Moessner (2010), sensör verilerinin bağlantılı veri olarak yayınlanması için Sense2Web adında bir bağlantılı veri platformu tanımlamıştır. Sense2Web farklı sensör verilerinin tanımlanması ve ilişkilendirilmesini sağlar. Konumsal ilişkiler tanımlanmamıştır.

Lehmann, Athanasiou, Both, García-Rojas, Giannopoulos, Hladky, Le Grange, Ngonga Ngomo, Sherif, Stadler, Wauer, Westphal & Zaslowski (2015), GeoKnow Projesi kapsamında web üzerinde konumsal bağlantılı verilerin yayınlanması için yazılım araçları geliştirmiştir. GeoKnow Generator açık konumsal bağlantılı verilerin yaşam döngüsünü içeren bütün yazılım araçlarını içerir. GeoKnow Generator konumsal verinin RDF verisine dönüşümü, diğer konumsal ve konumsal olmayan bağlantılı veri kaynaklarıyla ilişkilendirme, yeni bilgilerin çıkarsanması için veri birleştirme, web üzerinde görselleştirme gibi işlemleri destekler. Konumsal verilerin RDF dönüşümü için TripleGeo yazılımı kullanıldığı için bu çalışmada tercih

edilmemiştir. TripleGeo konumsal özniteliklerin dönüşümünü desteklerken konumsal olmayan özniteliklerin sadece bir tanesinin dönüşümünü destekler.

Iwaniak, Leszczuk, Strzelecki, Harvey, & Kaczmarek (2017), Polonya Mazovya eyaleti için geliştirilen prototip bir sistem önermiştir. Sistem, kadastral verilerin bağlantılı açık veri olarak yayınlanması için gerekli teknolojiler üzerine kurulmuştur. Kadastral veriler, GeoNames, DBpedia veri kaynakları ile ilişkilendirilmiştir. Konumsal ilişkiler tanımlanmamıştır. Kadastral verilerin bağlantılı veri olarak yayınlanması için semantik web teknolojileri tanıtılmıştır ancak işlem adımları verilmemiştir.

5. SONUÇLAR

Konumsal verilerin bağlantılı açık veri olarak yayınlanmasının önemi, Semantik Web uygulamalarının konumsal alanda da hızla yaygınlaşmasıyla daha da artmıştır. Buradan hareketle bu çalışma kapsamında konumsal verilerin bağlantılı açık veri olarak yayınlanması için bir metodoloji geliştirilmiştir. Uygulama senaryosu kapsamında ulusal düzeyde idari bölünmeler dikkate alınarak TR İdari Bölümler Ontolojisi geliştirilmiştir. İdari bölümler arasındaki topolojik ilişkilerin tanımlanması için OS Ordnance Survey Konumsal İlişkiler Ontolojisi yeniden kullanılmıştır.

Konumsal verilerin bağlantılı açık veri olarak yayınlanması için konumsal veri altyapıları kapsamında belirlenen veri setlerinin RDF formatına dönüştürülmesi gerekir. Bu dönüşüm sırasında verinin yeterli sayıda kelime hazinesi ile eşleştirilmesi gerekir. Bunun için mevcut kelime hazineleri içinde ilgili kelime hazineleri seçilmelidir. Seçilen kelime hazinelerinin konumsal verilerle ilişkilendirilmesi işlemi kullanıcı tarafından gerçekleştirilir. Bu aşama kullanıcıların RML kodlama tekniklerini bilmesini gerektirir. Ayrıca mevcut yazılım araçlarının konumsal verilerin RDF formatına dönüşümü için daha fazla veri formatını desteklemesi gerekir.

RDF formatındaki verilerin web üzerindeki veri setleri ile ilişkilendirilmesi için uygun veri setlerinin seçilmesi gerekir. Bu bağlamda Bağlantılı Açık Veri Bulutu içerisinde yer alan veri setleri incelendiğinde konumsal veri setleri sınırlı sayıdadır. Bu nedenle Bağlantılı Açık Veri Bulutu veri setlerinin konumsal alanda genişletilmesi gerekir. Ayrıca uygun veri setlerinin belirlenmesi ve eşleştirilmesi, kullanıcıların veri setlerinin içeriği hakkında bilgi sahibi olmalarını gerektirir. Bu durum bağlantılı veri yaklaşımının önündeki en büyük engellerden birisidir.

Konumsal verilerin bağlantılı açık veri olarak yayınlandığı bu çalışmada oluşturulan bağlantılı veri Fuseki ile localhost üzerinden yayınlanmıştır. Gelecek çalışmalarda topolojik ilişkiler kullanılarak SPARQL sorgu sonucu, SEXTANT yazılımı

kullanılarak harita üzerinde konumsal bağlantılı veriler görsel olarak sunulacaktır.

TEŞEKKÜR

Bu çalışma, TÜBİTAK 3001 Başlangıç Ar-Ge Projeleri Destekleme Programı kapsamında desteklenmiştir. Katkılarından dolayı TÜBİTAK Araştırma Destek Programları Başkanlığı (ARDEB), Çevre, Atmosfer, Yer ve Deniz Bilimleri Araştırma Destek Grubu (ÇAYDAG)' na teşekkür ederiz.

KAYNAKÇA

- Auer, S., Dietzold, S., Lehmann, J., Hellmann, S., & Aumuell, D. (2009). Triplify: light-weight linked data publication from relational databases. In *Proceedings of the 18th international conference on World wide web* (pp. 621-630). ACM.
- Barnaghi, P., Presser, M., & Moessner, K. (2010). Publishing linked sensor data. In *CEUR Workshop Proceedings: Proceedings of the 3rd International Workshop on Semantic Sensor Networks (SSN)*, Organised in conjunction with the International Semantic Web Conference (Vol. 668).
- Berners-Lee, T. (2006). *Design issues: Linked data*, Retrieved from <https://www.w3.org/DesignIssues/LinkedData.html>
- Bizer, C., Heath, T., & Berners-Lee, T. (2009). Linked Data - the story so far. *International Journal on Semantic Web and Information Systems*, 5 (3), 1-22. doi:10.4018/jswis.2009081901.
- Heath, T., & 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.
- ISO. (2011). *ISO 19152 Draft International Standard (DIS), Geographic information – Land Administration Domain Model (LADM)* Geneva, Switzerland.
- Iwaniak, A., Leszczuk, M., Strzelecki, M., Harvey, F. and Kaczmarek, I. (2017). A novel approach for publishing linked open geodata from national registries with the use of semantically annotated context dependent web pages. *ISPRS International Journal of Geo-Information*, 6(8), p.252.
- Lehmann, J., Athanasiou, S., Both, A., García-Rojas, A., Giannopoulos, G., Hladky, D., Le Grange, J.J., Ngonga Ngomo, A. C., Sherif, M. A., Stadler, C., Wauer, M., Westphal, P., Zaslowski, V. (2015).

Managing Geospatial Linked Data in the GeoKnow Project.

- Liebermen, J., Singh, R., & Goad, C. (2007). *W3C Geospatial Ontologies*, W3C Incubator Group Report, 23 October 2007.
- OGC. (2012). *OGC GeoSPARQL - A Geographic Query Language for RDF Data*, OGC® Implementation Standard, Editors: Matthew Perry and John Herring, Open Geospatial Consortium. Retrieved from <http://www.opengis.net/doc/IS/geosparql/1.0>
- Page, K. R., De Roure, D. C., Martinez, K., Sadler, J. D., & Kit, O. Y. (2009). *Linked sensor data: Restfully serving rdf and gml*. In Proceedings of the Semantic Sensor Networks 2009. Retrieved from <http://ceur-ws.org/Vol-522/p10.pdf>
- Saavedra, J., Vilches-Blázquez, L. M., ve Boada, A. (2014). *Cadastral data integration through Linked Data*, Huerta, Schade, Granell (Eds): Connecting a Digital Europe through Location and Place. In *Proceedings of the AGILE'2014 International Conference on Geographic Information Science*, Castellón. ISBN: 978-90-816960-4-3, Retrieved from https://agile-online.org/conference_paper/cds/agile_2014/agile2014_153.pdf
- Vilches-Blázquez, L.M., Villazón-Terrazas, B., Leon, A.D., Priyatna, F. & Corcho, O. (2010). An approach to publish spatial data on the web: The geolinked data case, Retrieved from <http://oa.upm.es/5657/1/Workshop41.LSTD2010.pdf>
- W3C. (2004). *Resource Description Framework (RDF): Concepts and Abstract Syntax*, W3C Recommendation, Retrieved from <https://www.w3.org/TR/rdf-concepts/>
- W3C. (2012). *R2RML: RDB to RDF Mapping Language*, Editors: Souripriya DasOracle, Seema Sundara-Oracle, Richard Cyganiak-DERI, National University of Ireland, Galway, W3C Recommendation. Retrieved from <https://www.w3.org/TR/r2rml/>
- URL-1. Ordnance Survey Ontologies, <https://data.ordnancesurvey.co.uk/ontology>