

# Baęlı Veri Entegrasyonu İin Bir Masaüstü Uygulama Mimarisinin Medikal Alanda Geliştirilmesi

Umutcan Şimşek, Rıza Cenk Erdur

Ege Üniversitesi Bilgisayar Mühendisliği Bölümü, Bornova, İzmir

s.umutcan@gmail.com, cenk.erdur@ege.edu.tr

**Özet.** Farmakolojiye yönelik araştırma süreçlerinde yaşanan en büyük problemlerden birisinin bir aktif madde ile ilgili olarak internet üzerinde geleneksel arama motorlarıyla yapılan aramalardan dönen sonuçların, işe yararlığa göre süzülüp entegre edilmesi olduğu görülmüştür. Bu sorunun önüne geçmek için baęlı verinin gücünden yararlanan ve yüksek erişilebilirlikli bir uygulama geliştirilmiştir. Ege Üniversitesi İla Araştırma ve Farmakokinetik Geliştirme-Uygulama Merkezi (ARGEFAR) danışmanlığında geliştirilen bu uygulama, baęlı açık ila verisi bulutu üzerindeki verilerin, bir masaüstü uygulamasında entegrasyonu için geliştirilen bir mimariye dayanmaktadır.

Veri erişiminde güncellięin sağlanması için SPARQL uçnoktalarında sorgu dağıtımı erişim deseni kullanılmıştır. Bu desen her zaman güncel veriyi getirmesine karşın, erişilebilirlik sorunlarına sahip olduğundan, mimaride, desen üzerinde önbellekleme ve indeksleme gibi mekanizmalar geliştirilmiştir.

Mimarinin, nesneye dayalı programlama paradigmasına dayalı esnek yapısı sayesinde yeni veri kaynaklarıyla genişlemeye uygun olduğu görülmüştür. Ayrıca paralelleştirmeye uygun yapısı sayesinde, dağıtık olarak da çalıştırılabileceęi düşünülmektedir. ARGEFAR araştırmacılarıyla yapılan testlerde, sadece isminin bir kısmı dahi bilinen bir aktif madde ile ilgili gerekli bilgilere ulaşımın oldukça kolaylaştığı gözlemlenmiştir.

**Anahtar Kelimeler.** RDF, SPARQL, veri entegrasyonu, baęlı veri, medikal bilişim

## 1 Giriş

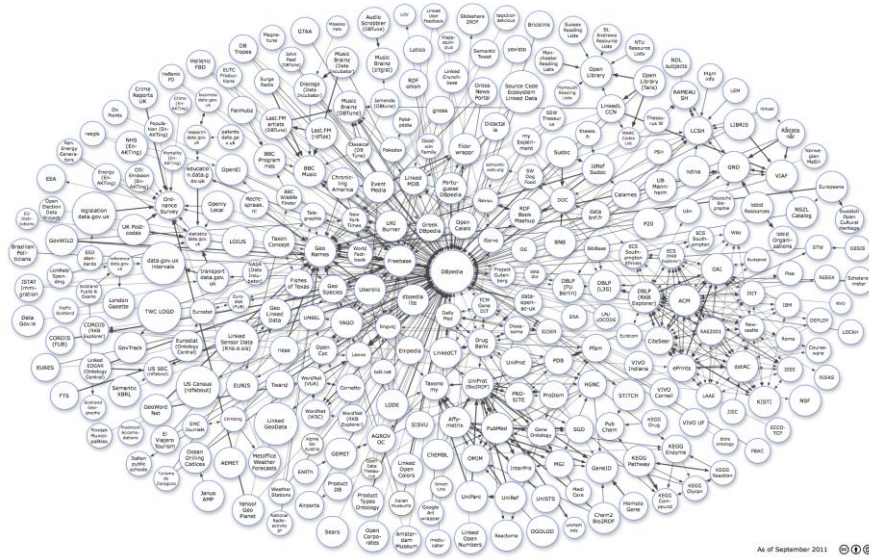
Birok alandaki araştırma-geliştirme çalışmalarının önemli bir bölümünü, verinin deęişik kaynaklardan toplanması ve bu toplanan verinin entegre edilmesi oluşturmaktadır. Bu süreçteki en büyük problemlerden birisi de, geleneksel arama motorlarıyla yapılan aramalar içinden, işe yarar verilerin süzülmesi ve entegre edilmesidir. Günümüz interneti, dökümanların birbirlerine bağlanmasıyla oluştuğundan, arama motorları sadece metin tabanlı arama yapmakta, direkt olarak aranılan kavramı deęil,

kavramın içinde bulunduğu dökümanı büyük ölçüde anlamsal olarak incelemeksizin bulmaktadır [15].

Bu çalışmada, bağlı verinin gücünü kullanarak, farmakoloji araştırmacılarının uygun veriyi bulma ve birleştirme sürelerini kısaltmaya, araştırma sürecinin etkinliğini artırmaya yönelik bir uygulama geliştirilmiştir.

Bağlı veri, temel olarak farklı kaynaklardaki verinin anlamsal çıkarımlar yapmaya uygun şekilde bağlanmasıyla oluşur. HTML dökümanlarına dayalı web dökümanların birbirine bağlanmasıyla oluşurken; bağlı veri, veriyi RDF (Resource Description Framework) [2] formatında dökümanlarda tutar ve dökümanlar yerine verinin kendisini yani kavramları bağlar [5].

Bağlı verinin en somut uygulaması Linking Open Data Project'tir (Açık Veriyi Bağlama Projesi) [5]. İnternet üzerinde, çeşitli formatlarda açık olarak yayınlanan verinin, Tim Berners Lee'nin öne sürdüğü "5 yıldız kuralı"na [4][9] uydurularak birbiriyle bağlanmasını hedefleyen bu projede oluşan veri bulutu her gün giderek büyümektedir. 2009 yılında bile bu bulut milyarlarca kayıt içermektedir. Açık bağlı veri bulutunda 2011 yılında yer alan veri kümeleri Şekil - 1'deki gibidir [11].

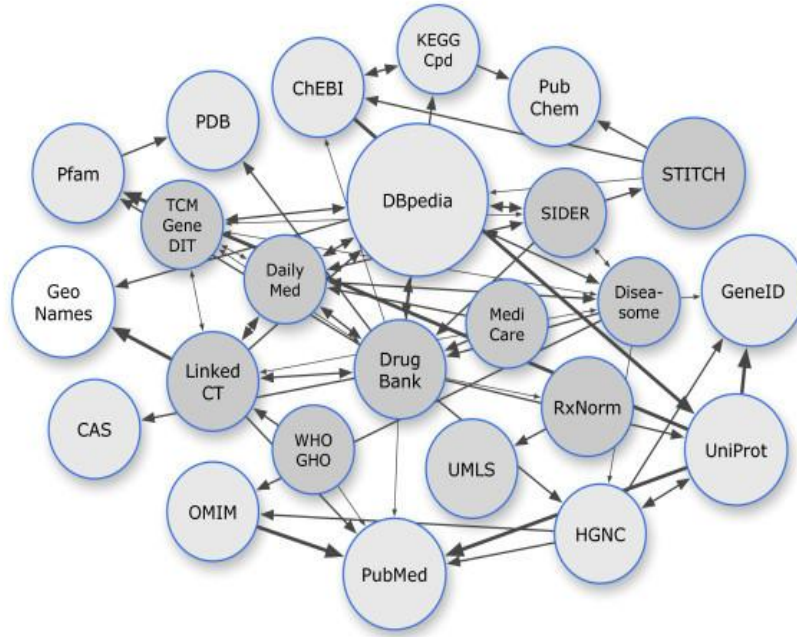


Şekil 1. 2011 yılında açık bağlı veri bulutu

Bu çalışmada, farmakoloji araştırmaları için geliştirilen uygulamada, bağlı açık veri bulutunun bir parçası olan Linked Open Drug Data (LODD)<sup>1</sup> Verisi bulutunda yer alan veriden yararlanılmıştır (Şekil - 2) [14]. Bu bulut, internette dağınık olarak bulunan onlarca biyomedikal veri kümesini birbirine bağlamaktadır. Örneğin, Aspirin

1 Bağlı Açık İlaç Verisi

ilacı ile ilgili kimyasal yapı verisi ve klinik deney verisi bu bulut kullanılarak farklı kaynaklar üzerinden getirilebilir [14].



Şekil 2. LODD (Linked Open Drug Data) bulutu

Bu çalışma kapsamında geliştirilen uygulama, LODD üzerinde ARGEFAR danışmanlarıyla belirlenen veri kümeleri üzerindeki verileri bir masaüstü uygulamasında entegre etmektedir. Uygulama, başka alanlara özgü masaüstü bağlı veri uygulamaları için de kullanılabilir katmanlı bir mimari önermektedir.

Bildiri şu bölümlerden oluşmaktadır: Bölüm 2'de bu çalışma ile ilgili daha önceden gerçekleştirilen çalışmalar kısaca incelenmiştir. Bölüm 3'te uygulamanın gerçekleştirimi ve mimarisi açıklanmıştır. Bölüm 4'te ise uygulamanın sonuçları tartışılmıştır.

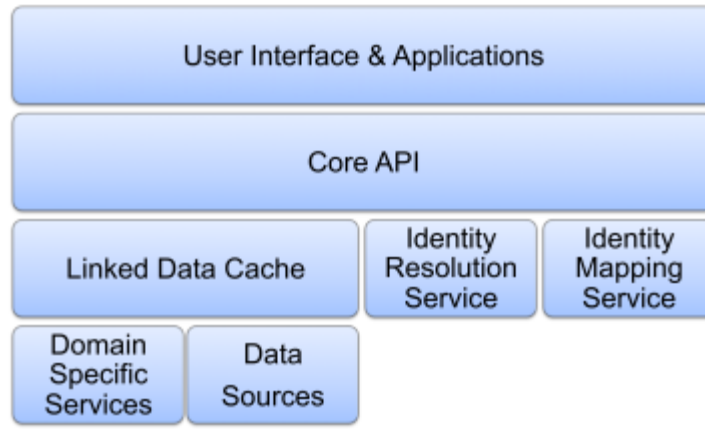
## 2 İlgili Çalışmalar

### 2.1 OpenPHACTS

OpenPHACTS projesi, üniversitelerin, araştırma kurumlarının ve özel sektörden birçok firmanın işbirliği ile hazırlanmış ve hayata geçirilmeye başlanmıştır. Proje, farmakoloji araştırmalarında ortaya çıkan veriye ulaşım ve ulaşılan verinin kullanıla-

bilecek hale getirilmesi problemini bağılı veriden yararlanarak ortadan kaldırmayı amaçlar. [6]

OpenPHACTS kapsam itibariyle oldukça geniş bir alana hitap eder, ilaçların farmasötik özelliklerinin yanında kimyasal yapılarıyla ilgili de bilgi sunar. OpenPHACTS'in kullanıma sunduğu internet üzerinde çalışan OPS [6] platformu, kendi başına bir uygulama olmaktan ziyade, geliştirilecek bağılı veri uygulamalarına bir altyapı hazırlamaktadır. Utopia Documents isimli makale zenginleştirme uygulaması buna örnek olarak gösterilebilir [1].



Şekil 3. OPS mimarisi[6]

OPS platformu, yedi bileşenli bir mimari kullanmayı seçmiştir. Şekil - 3'te görülebilecek, dört katmana yayılan bu yedi bileşen şu şekildedir [6]:

1. Data Sources (Veri Kaynakları): Bu bileşen, mevcut RDF verikümelerinin bulunduğu katmandır. Enzimlerden, diğer kimyasal bileşeklere, birçok veriyi sunan UniProt[12], LODD[10] ve Bio2RDF[3] kaynakları bu katman aracılığı ile kullanılmaktadır.
2. Linked Data Cache (Bağılı Veri Önbelleği): Bu bileşen, güvenilirlik ve performans açısından iyileştirme yapılması amacıyla kullanılan veri kümelerinin tamamıyla yerel olarak saklanması amacıyla kullanılmaktadır. SPARQL sorgularının sunucuları yoracağı varsayılarak tüm sorgulama işlemlerinin burada yapılması öngörülmüştür.
3. Identifier Resolution Service (Kimlik Çözümleme Servisi): Bu katman, serbest metin arama ile yapısal sorguların birbirinden ayrıldığı katmandır.
4. Identifier Mapping Service (IMS): Eşitliğin bağlam bağımlı olmasından yola çıkarak, iki farklı veri kaynağından gelen aynı kavrama ait verilerin eş olup olmadığını kontrol eden bileşendir.

5. Domain Specific Services (Alana Özgü Servisler): OPS, alana özgü servislere örnek olarak, farklı veri kaynaklarındaki küçük kimyasal bileşenleri, yapılarına göre birbirlerine entegre edip eşsiz bir kimlik numarası veren ChemSpider isimli servisten yararlanmaktadır.
6. Core API (Çekirdek API): OPS platformunun ilk sürümlerinde yer almayan bu katman, SPARQL sorgularını, OPS üzerinde geliştirilen uygulamalardan soyutlar. Uygulama geliştiricilere önceden tanımlanmış yöntemler sunarak, geliştiricilerin düşük performanslı sorguları çalıştırmalarının önüne geçer. Bu katman, uygulama geliştiricilerin sisteme dahil oldukları katmandır.
7. User Interfaces/Applications(Kullanıcı Arayüzü/Uygulamalar)

Mimarinin diğer bileşenleri ile ilgili ayrıntılı bilgiye [6] kaynağından erişilebilir.

## 2.2 TripleMap

TripleMap2, ağırlıklı olarak sağlık alanında olmakla birlikte, değişik veri kaynaklarındaki verinin grafik olarak gösterilmesini sağlayan bir internet uygulamasıdır. Bu uygulama, kavramları simgelerle temsil edip kavramlar arasındaki bağları etiketleyerek bir çizge düzeninde gösterir [14]. Bir veri kümesinde neyin neyle bağlı olduğunu görmek için kullanışlı bir uygulamadır.

## 2.3 Tartışma

OPS platformu, verinin anlık olarak güncel bir şekilde getirilmesinde bazı sorunlar yaşayabilmektedir. Çünkü Şekil - 3'te görülen mimarideki Linked Data Cache bileşeni, kullanılan veri kümelerindeki tüm verinin OPS sunucularına alınmasıyla oluşur ve sorgular bu sunucular üzerinde çalışır. Bu da programın zaman zaman güncel olmayan veri ile çalışmasına neden olabilir. Bunların dışında, OPS platformunun kimlik çözümü servisi ve bazı eşleştirme mekanizmaları oldukça gelişmiştir. Yine de OPS platformu, bir uygulamadan ziyade bir altyapı sunduğundan, bu çalışma kapsamında geliştirilen uygulama ile işlevsel açıdan birebir karşılaştırmak pek mümkün değildir. Zira bu çalışma kapsamında geliştirilen uygulamada, ilerleyen sürümler OPS platformunun Core API bileşeninden yararlanabilir.

TripleMap uygulaması oldukça başarılı bir görselleştirme sunmanın yanında, arama sonuçlarına cevap vermede eksiklikleri görülmüştür. Bu çalışma kapsamında geliştirilen uygulamada yapılan Propranolol araması ilacın farmakolojik özelliklerine dair verilerini getirirken, TripleMap'te yapılan aramada zaman zaman sonuç gelmemekte, zaman zaman da sadece belirli veriler dönmektedir. Bunun dışında, TripleMap'in mimarisine dair bir içerik bulunmadığından iki çalışma bu açıdan karşılaştırılamamaktadır.

---

2 <http://www.triplemap.com>

### 3 Gerçekleştirme

Bir bağılı veri uygulamasında, en önemli karar noktası, bağılı veriye nasıl erişileceğidir. Bağılı veri araştırmacılarının genel olarak kabul ettiği üç erişim deseni bulunmaktadır: Emekleme Deseni, Yerinde Çözümleme Deseni, Sorgu Dağıtım Deseni [9].

Emekleme Deseni, temel olarak günümüz arama motorları ile benzer şekilde çalışır. Bu desende tüm veri uzayı gezilerek sonuçlar bir yerde depolanır ve arama sonuçları bu depolanan verilerden getirilir. Her ne kadar büyük veri kümeleri için uygun olsa da güncel olmayan veriyle çalışma ihtimali oldukça yüksektir. Bir diğer desen olan Yerinde Çözümleme deseni de emeklemeye benzerdir fakat emekleme işlemi sadece ihtiyaç duyulduğunda yapılır. Bu desenin de performansı işlemler karmaşıklıkça önemli ölçüde düşer.

Üçüncü ve son desen ise Sorgu Dağıtım Deseni'dir. Bu desen, SPARQL sorgularının arama ile eş zamanlı olarak çevrimiçi SPARQL uçnoktalarına dağıtılması ile çalışır. Ulaşılan veri sürekli güncel olmasına rağmen, uçnoktaların zaman zaman performans sorunları yaşamaları ve karmaşık sorgularda, özellikle serbest metin aramada çabuk sonuç döndürememeleri sebebiyle sorgu etkinliği önemli ölçüde düşebilmektedir. Gerçekleştirilen uygulama, indeksleme ve kalıcı katman oluşturma(önbellekleme) gibi yöntemler kullanarak sorgu dağıtım deseni iyileştirilmesini amaçlar.

#### 3.1 Veri Kümeleri

Uygulama gerçekleştirilirken, ARGEFAR danışmanlarıyla birlikte ilk prototiplerde<sup>3</sup> kullanılmak üzere dört adet veri kümesi seçilmiştir.

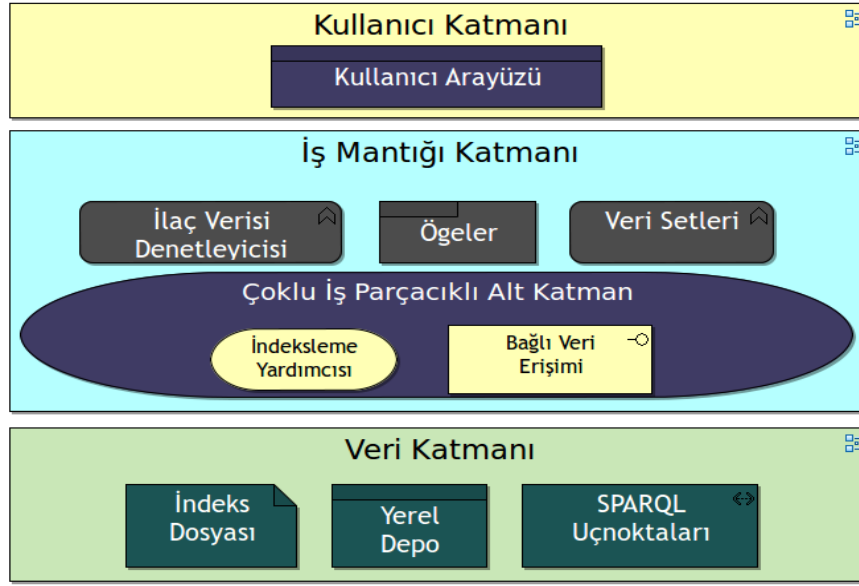
- DBPedia<sup>4</sup> <http://dbpedia.org>: Wikipedia'nın bağılı veri versiyonudur. Neredeyse tüm wikipedia verisine bağılı veri formatında erişilebilir, verilerin diğer veri kümeleriyle bağlar da kurulmuştur.
- Drugbank: Yaklaşık 5000 FDA onaylı ilaç ile ilgili gerek farmakolojik gerek biyolojik gerekse kimyasal bilgiyi sunan güncel bir veri kaynağıdır [10].
- DailyMed: Amerikan Ulusal Tıp Kütüphanesi tarafından yayınlanan FDA onaylı ilaçların prospektüs bilgilerini yayınlayan bir veri kaynağıdır. Bu veri kaynağı RSS beslemeleri kullanılarak bir üniversite tarafından bağılı veri olarak yayınlanmaktadır [10].
- LinkedCT: Dünya üzerinde tamamlanmış veya hala süren klinik ilaç deneylerinin yayınlandığı [ClinicalTrials.gov](http://ClinicalTrials.gov) sitesinin bağılı veri olarak yayınlanmış halidir. 158 ülkeden 60000 üzerinde klinik deney verisi yer almaktadır [8].

<sup>3</sup> Prototipin kaynak koduna <https://github.com/sumutcan/Pharmeshup> adresinden ulaşılabilir.

<sup>4</sup> <http://www.dbpedia.org>

### 3.2 Yazılım Mimarisi

Geliştirilen üç katmanlı mimari Şekil - 4'te görülmektedir. Mimari, en üstte kullanıcıyla etkileşimin sağlandığı arayüz katmanı, ortada çoklu iş parçacıklı alt katman ile birlikte uygulamanın asıl yükünü taşıyan iş katmanı ve en altta da veri kaynaklarının bulunduğu veri katmanından oluşur.



Şekil 4. Geliştirilen katmanlı mimari

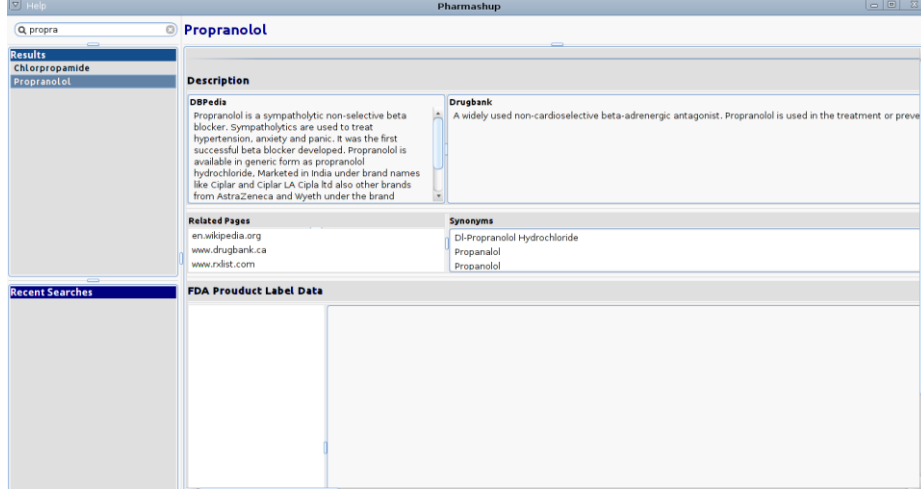
**Kullanıcı Katmanı.** Bu katman kullanıcı ile uygulamanın etkileşime girdiği katmandır (Şekil - 5). Kullanıcıya sunulan veriler, bu katmanda Genel Bilgiler, Farmakokinetik, Farmakodinamik ve Klinik Deneyler olarak dört sekme halinde gösterilmektedir.

Genel Bilgiler sekmesinde, aktif madde ile ilgili kısa açıklamalar ve prospektüs bilgileri gibi bilgiler yer almaktadır.

Farmakokinetik sekmesinde, canlı metabolizmasının ilacın aktif maddesine ne yaptığı ile ilgili bilgiler yer almaktadır. Emilim, vücuttan uzaklaştırılma süresi gibi bilgiler bu kategoriye girer.

Farmakodinamik sekmesinde ise, ilacın canlı metabolizmasına ne yaptığı ile ilgili bilgiler verilmektedir.

Klinik Deneyler sekmesi de, o aktif madde ile ilgili klinik deney verilerinin güncel bir şekilde ClinicalTrials.gov kaynağından getirilip kullanıcıya sunulduğu sekmedir.



Şekil 5. Kullanıcı arayüzünden bir ekran görüntüsü

**İş Katmanı.** Bu katman, iki tanesi çoklu iş parçacıklı alt katmanda yer almak üzere, toplamda beş bileşenden oluşur. Bu bileşenlerden ilki olan *İlaç Verisi Denetleyicisi*, GRASP [13] desenlerinde Controller'a karşılık gelen bileşendir ve uygulamaları kullanıcı arayüzünden diğer katmanlara çıkış noktasını oluşturur.

*Ögeler* bileşeni, iş katmanında gerçekleşen işlemlerde bazı gerekli verilerin tutulması için yaratılan sınıfları içerir. Bu bileşendeki en önemli sınıf, arama yapılan karakter katarını bellekteki indekslenen elemanlarla eşleştirip bir RDF kaynağına dönüştüren sınıftır.

*Veri Setleri* bileşeni de, farklı veri kümelerini temsil eden sınıfları içerir. Nesneye dayalı paradigma ile sağlanan esnek yapı sayesinde yeni bir veri kümesi eklemek, neredeyse sadece yeni bir sınıf oluşturmak kadar basitleşmiştir.

Bu katmanda yer alan Çoklu İş Parçacıklı Alt Katman, uygulamanın veri erişimi, indeksleme, önbellekleme gibi işlemlerini yaparken hem performansı artırması hem de yapılan yoğun işlemlerin kullanıcının aramalarını bölmemesi için oluşturulmuştur.

Bu alt katmanda iki bileşen yer almaktadır:

- **İndeksleme Yardımcısı:** Bu bileşen, belirlenen aralıklarda, veriye ulaşım için anahtar olan verileri, SPARQL uçnoktalarına yaptığı sorgularla bir JSON<sup>5</sup> dosyasında indeksler. Uygulamanın şu anki halinde yaklaşık 15000 ilaç verisi ve bunların eşdeğerleri indekslenmektedir. Kullanıcı tarafından girilen arama ifadesi, bu JSON dosyasının belleğe eşlenmiş hali üzerinde aranarak bulunan RDF kaynakları döndürülür. Bu RDF kaynakları aracılığı ile istenen ilacın bağlı veri bulutu üzerindeki

<sup>5</sup> Nesneye dayalı paradigmaya göre tasarlanmış bir veri gösterim formatıdır.



bilgilerine daha hızlı bir şekilde ulaşılır. Tablo - 1'de, indeksleme işleminin doğrudan SPARQL uçnoktaları üzerinde yapılan serbest metin aramasına göre getirdiği iyileştirme açığı görülmektedir.

**Tablo 1.** DBPedia verisi üzerinde, isminde “pro” ifadesi geçen aktif maddelerin ve eşdeğerlerinin aranması

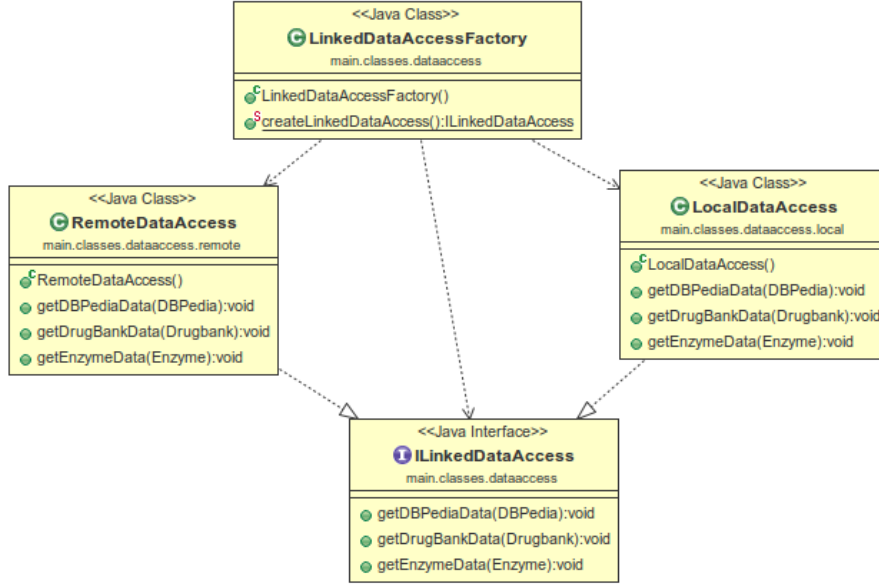
SPARQL Uçnoktası Tepki Süresi	İndeksleme Mekanizması Tepki Süresi
32 saniye	0.3 saniye

- **Bağlı Veri Erişimi:** Bu katmanda bağlı veriye erişim gerçekleşir. Sorgu performanslarının ve erişilebilirliğin artırılması, internet bağlantısına bağımlılığın azaltılması için oluşturulan kalıcı katmana (önbellek) erişim de yine bu katmanda gerçekleştirilir. Strateji tasarım deseni ile oluşturulan bu bileşende (Şekil - 6), çalışma zamanında SPARQL uçnoktalarından ulaşılan bir aktif maddeye ait alt çizge arkaplanda yerel üçlü deposuna kaydedilir. Uçnoktaların ulaşılamaz olduğu zamanlarda strateji tasarım deseni ile oluşturulan yapı yerel depoya ulaşarak daha önceden önbelleklenmiş veriyi getirir. Böylece daha önce bir kere ulaşılmış veri, uygulamanın kullandığı her an erişilebilir olur.

Kalıcı katman aynı zamanda basit bir kimlik çözümleme servisi görevi de görür. Bağlı veride, iki RDF kaynağı owl:sameAs özelliği ile birbirine bağlanabilir. Bu özellik iki kaynağın tüm yönleriyle eş olduğunu gösterir [7]. Fakat açık bağlı veri bulutunda owl:sameAs linklerine her zaman güvenmek doğru olmaz. Örneğin Metamizole etken maddesi için, DBPedia kaynağı, Drugbank kaynağı ile owl:sameAs aracılığı ile bağlanmıştır. Halbuki, isim olarak aynı da olsalar, DBPedia'daki Metamizole, metamizol maddesinin sodyum tuzudur. Bu durum maddenin kimyasal formülünü değiştirdiğinden, iki kaynaktaki RDF ifadelerinin owl:sameAs ile bağlanabilmesi mümkün değildir. Bu nedenle, bir etken madde ile ilgili farklı kaynaklardan gelen alt çizgeler, isimlendirilmiş çizgeler olarak tutularak verinin kaynağının da belli olması sağlanır [16].

Çoklu İş Parçacıklı Alt Katman'da, geleneksel Thread yapısının yanı sıra, Java 5.0 ile birlikte gelen Callable API'si<sup>6</sup> de kullanılmıştır. Bu API, Thread'den farklı olarak, iş parçacıklarından sonuç dönmesine ve zaman aşımı gibi çeşitli istisnai durumların yakalanmasına olanak sağlar.

<sup>6</sup> <http://docs.oracle.com/javase/1.5.0/docs/api/java/util/concurrent/Callable.html>



Şekil 6. Veri erişimi için kullanılan strateji tasarım deseni UML şeması

**Veri Katmanı.** Bu katmanın bir yerel bir de uzak tarafı vardır. Yerel tarafta, yani uygulama tarafında, indekslenen anahtar verilerin tutulduğu JSON dosyası ve önbelleklenmiş ilaç verisinin tutulduğu yerel üçlü deposu; uzak tarafta ise verilerin getirildiği SPARQL uç noktası bulunmaktadır.

#### 4 Sonuç

Geliştirilen mimari, özellikle düşük yoğunluklu sorgularla veya büyük yoğunluklu sorguların daha küçük parçalara dağıtılabildiği sorgularla çalışacak, masaüstü bağlı veri uygulamaları için oldukça kolay uygulanabilir, genişleyebilir ve etkin bir jenerik çözüm ortaya koymuştur. Zira Veri Setleri bileşeninde her veri kümesini bir Java sınıfı temsil etmektedir ve tüm veri kümeleri ortak bir soyut sınıftan türediğinden, yeni veri kümeleri sadece bu sınıf implemente edildikten sonra soyut yöntemlerin implemente edilmesiyle eklenebilir. Aynı şekilde, sorgular da ayrı depo sınıflarında tutulduğundan; mimari, farklı veri kümeleri ve sorgular tanımlanarak farklı alanlara da uygulanabilir.

Mimarinin paralel yapısının, işlemleri tamamen veya kısmen dağıtık sistemler üzerinde gerçekleştirilebilmeye uygun olacağı düşünülmektedir. Şekil - 4'te görülen mimaride, çoklu iş parçacıklı alt katmandaki bileşenler birbirlerinden büyük ölçüde

bağımsız olduklarından, farklı sistemlere dağıtılıp farklı stratejiler (mesaj gönderme, paylaşılan bellek gibi) ile beraber çalıştırılabilecekleri öngörülmektedir.

Gerçekleştirilen uygulama, ARGEFAR araştırmacılarıyla yapılan ilk testlerden olumlu sonuçlar elde etmiş, araştırma sürelerini kısalttığı görülmüştür.

## Kaynaklar

1. TK Attwood, DB Kell, P McDermott, J Marsh, SR Pettifer and D Thorne. Utopia Documents: linking scholarly literature with research data. *Pub Med, Bioinformatics*, 41(5), 2010.
2. Dave Beckett, *RDF/XML Syntax Specification*, 2004.
3. F Bellau, MA Nolin, N Tourigny, P Rigault and J Morissette. Bio2RDF: towards a mashup to bioinformatics knowledge systems. *Journal of Biomedical Informatics*, 41(5):706-716, 2008.
4. Tim Berners Lee, *Linked Data – Design Issues*, 2006.
5. Cristian Bizer, Tom Heath, Tim Berners-Lee and M. Hepp. *Linked Data – The Story So Far*. *International Journal on Semantic Web and Information Systems* 5(3):1-22, 2009.
6. Alasdair J G Gray, Paul Groth, Antonis Loizou, Sune Askjaer, Christian Breninkmeijer, Kees Burger, Christine Chichster, Chris T Evelo, Carole Goble, Lee Harland, Steve Pettifer, Mark Thompson, Andra Waagmeester and Antohny J Williams. *Applying Linked Data Approach to Pharmacology: Architectural Decisions and Implementation*. *Semantic Web Journal*, 1(13), 2012.
7. Harry Halphin and Patrick J Hayes. When owl:sameAs isn't the Same: An Analysis of Identity Links on the Semantic Web, *ISWC*, 6496, 2010.
8. Oktie Hassanzadeh, Anastasios Kementsietsidis, Lipyew Lim and Min Wang. *LinkedCT: A Linked Data Space for Clinical Trials*. *CoRR*, 2009.
9. Tom Heath and Cristian Bizer. *Linked Data – Evolving the Web into a Global Data Space*. Morgan and Claypool Publishers, 1<sup>st</sup> edition, 2011.
10. Anja Jentzsch, Matthias Samwald and Bo Anderson. *Linking Open Drug Data. Triplification Challenge*, 2009.
11. Anja Jentzsch and Richard Cyganiak. *The Linking Open Data Cloud Diagram* – <http://lod-cloud.net>
12. Michele Magrane and UniProt Consortium. UniProt knowledgebase: a hub of integrated protein data. *BMC Bioinformatics*, 10(1):136+, 2009.
13. RM Noorullah. GRASP and GoF Patterns in Solving Design Problems. *International Journal of Engineering and Technology*, 1(3):196-205, 2011.
14. M Samwald, A Jentzsch, Christopher Boutton and Claus Stie Kalleso e. *Linked Open Drug Data for Pharmaceutical Research and Development*, *J Cheminform*, 3(19), 2011.
15. John F Sequeda, *Consuming Linked Data*, 2010.
16. Chen Zhangjian and Qian Nen. *Research of Semantic Web Publishing using Named Graphs*. In *Semantics, Knowledge and Grid*, page 132, 2005.