

Türkçe Kelime Ağı KeNet için Arayüz

User Interface for Turkish Word Network KeNet

Rıza Özçelik¹, Gökçe Uludoğan¹, Selen Parlar^{1,2}, Özge Bakay³, Özlem Ergelen⁴, Olcay Taner Yıldız⁵

¹Boğaziçi Üniversitesi Bilgisayar Mühendisliği Bölümü, İstanbul, Türkiye

²Starlang Yazılım Danışmanlık, İstanbul, Türkiye

³Boğaziçi Üniversitesi İngilizce Öğretmenliği Bölümü, İstanbul, Türkiye

⁴Boğaziçi Üniversitesi Dilbilim Bölümü, İstanbul, Türkiye

⁵Işık Üniversitesi Bilgisayar Mühendisliği Bölümü, İstanbul, Türkiye

{riza.ozcelik, gokce.uludogan, selen.parlar, ozge.bakay, ozlem.ergelen}@boun.edu.tr, olcaytaner@isikun.edu.tr

Özetçe—Kelime ağları, bir dildeki kelimeler arasındaki bağlantıları, eş anlam kümeleri oluşturarak ve bu kümeleri birbirine çeşitli anlamsal bağlantılar ile bağlayarak temsil eden bir çizge veri yapısıdır. Doğal dil işleme alanındaki en yaygın bilinen kelime ağı WordNet 1990 yılında İngilizce için oluşturulmuşken, Türkçe için en kapsamlı ağ, 2018 yılında oluşturulan KeNet'tir. Bildiğimiz kadarıyla, içinde 80000 eş anlam kümesi ve 25 farklı anlamsal bağlantı bulunan KeNet için şu ana kadar geliştirilen bir kullanıcı arayüzü yoktur. Bu çalışmada, KeNet çizgesinde, anlamsal bağlantıları kullanarak eş anlam kümeleri arasında çevrimiçi olarak gezinmeyi sağlayan bir arayüz sunuyoruz. Bu arayüz sayesinde, bir söz öbeği KeNet'te aranabilir ve eş anlam kümeleri arasındaki üst/alt anlam, parça-bütün ilişkileri gibi ilişkiler kullanılarak KeNet üzerinde gezilebilir. Ayrıca, herhangi bir eş anlam kümesinin, varsa, İngilizce karşılığının kimliği de görüntülenebilir ve bu kümeye WordNet'e ait internet sayfasından erişilebilir.

Anahtar Kelimeler—Kelime Ağı, KeNet, Kullanıcı Arayüzü

Abstract—Word networks are graphs that represent relations between literals by constructing synonym sets and connecting these sets via various semantic relations. WordNet was constructed in 1990 for English, which is the most popular word network in natural language processing, whereas, KeNet was proposed in 2018, which is the most comprehensive network for Turkish. To the best of our knowledge, there is no user interface for KeNet, which contains 80000 synonym sets and 25 different semantic relations. In this work, we present a user interface, which enables traversing synonym sets in KeNet via semantic relations. Thanks to this interface, KeNet can be searched for a literal and can be traversed using the relations such as hypernym/hyponym and holonym/meronym. Moreover, for any synonym set, ID of the corresponding WordNet synonym set can be seen, if exists and this synonym set can be viewed in the WordNet's web page.

Keywords—Word networks, KeNet, User Interface

I. GİRİŞ

Kelime ağı, dillerdeki eş anlam kümelerini ve bu kümeler arasındaki ilişkileri göstermek için oluşturulan çizge yapılarına verilen isimdir. Bu çizgede, her bir düğüm bir anlam kümesine karşılık gelirken, düğümler arasındaki bağlantılar da anlam kümeleri arasındaki bağlantıları temsil eder. Her bir bağlantı; zıt anlam, eş anlam, alt anlam ve üst anlam gibi farklı ilişkilere tekabül eden bir tipe sahiptir.

Dildeki anlamsal ilişkilerin bilgisayarlar tarafından işlenebilir bir yapıda sunulduğu bu anlamsal ağ, kelime ağlarını doğal dil işleme (DDİ) uygulamaları için oldukça faydalı bir araç haline getirmektedir. Çünkü; kelime ağları bir DDİ programının karşılaştığı her bir kelime için sistematik bir biçimde sorgulanabilir anlamsal bir bağlam sunmakta; ve bu bağlam, duygu analizi, soru cevaplama, duygu analizi gibi birçok DDİ uygulamasında kullanılabilir. KeNet de tıpkı WordNet gibi eş anlam kümelerinin birbirine çeşitli bağlantılarıyla bağlanmasıyla oluşturulmuş bir ağ yapısıdır.

Kelime ağlarının sunduğu bu potansiyel fayda, örnek ağlar oluşturmak için de bir teşvik olmuştur. Oluşturulan ağların en popüler olan WordNet [1], ilk olarak 1995 yılında İngilizce için ortaya konulmuştur ve bu ağda isimler, fiiller ve sıfatlar ayrı ayrı eş anlam kümelerinde gruplanmış ve bu kümeler çeşitli dilsel bağlantılarla birbirleriyle ilişkilendirilmiştir. Türkçe için de benzer çalışmalar yapılmış ve 2018 yılında KeNet ortaya konmuştur. KeNet de tıpkı WordNet gibi eş anlam kümelerinin birbirine çeşitli bağlantılarıyla bağlanmasıyla oluşturulmuş bir ağ yapısıdır.

Yaptığımız bu çalışmada, KeNet kelime ağındaki eş anlam kümelerini ve bağlantılarını çevrimiçi olarak gezinmeyi sağlayan bir arayüz sunuyoruz [2]. Sunduğumuz arayüz, bir söz öbeğini KeNet'te arama özelliğini desteklemekle kalmayıp bu söz öbeğinin dahil olduğu eş anlam kümesinin çeşitli bağlantılarıyla ilişkili olduğu eş anlam kümelerini de göstermektedir. Her ne kadar KeNet'te veri tipi olarak 25 tane anlamsal ilişki tipi olsa da, arayüz bu çalışmanın yazıldığı tarih itibarıyla 13 tanesini kullanarak gezinmeyi desteklemektedir. Desteklenen ilişkileri¹ şöyle sıralayabiliriz: Üst anlam (hypernym), alt anlam (hyponym), zıt anlam (antonym), bakınız (also see), üyelik bildiren bütün parça (member holonym), madde bütün parça (substance holonym), kısımlık bildiren bütün parça (part holonym), üyelik bildiren parça bütün (member meronym), madde parça bütün (substance meronym), kısımlık bildiren parça bütün (part meronym), üst anlam örneği (instance hypernym), alt anlam örneği, ilişkili konu (domain topic). WordNet'te de bulunan bu anlamsal ilişkilere ek olarak, KeNet'te diller arası ilişkiyi gösteren bir bağlantı tipi de vardır. Bu bağlantı kullanılarak, KeNet'te bulunan eş anlam kümesinin, WordNet'teki karşılığının kimliği bulunabilir ve bu

küme WordNet'e ait internet sayfasında detaylı olarak görümlenebilir. Bu özellik de arayüz tarafından desteklenmektedir.

İşbu makale şöyle düzenlenmiştir: II. kısımda kelime ağlarına dair geçmişteki çalışmalar özetlenmiş; III. kısımda KeNet hakkında detaylı bilgi verilmiş; IV. bölümde arayüz kullanımı anlatılmış ve V. bölümde çalışma özetlenmiş ve gelecekte yapılabilecek iyileştirilmeler aktarılmıştır.

II. LİTERATÜR ÖZETİ

Kelime ağı, düğümlerin (node) literalleriyle birlikte sözcük anlamlarını, bu düğümler arası bağlantılarınsa anlamsal ilişkileri temsil ettiği bir çizge veri yapısıdır. Farklı alanlardaki sözcük kullanım tarzları ve yeterli sözcük bilgisi gerektirmesi sebebiyle kelime ağı oluşturmak, yoğun bir çaba gerektirir.

Kelime ağlarının ilki, George Miller öncülüğünde Princeton Üniversitesi'nde oluşturulan WordNet'tir [1]. WordNet'te ana ilişki eş anlamlılıktır. Son sürüm olan WordNet 3.1, 117000 anlam kümesi (synset) ve 206941 sözcük-anlam çiftinden oluşmaktadır. Bu eş anlam kümeleri, birbirine kavramsal ilişkiler ile bağlıdır. Zıt anlam, alt anlam, üst anlam bu ilişkilerin örnekleridir.

WordNet'i takiben diğer diller için de kelime ağları oluşturulmuştur. EuroWordNet [3] ve Arapça WordNet [4] bunlardan birkaçıdır.

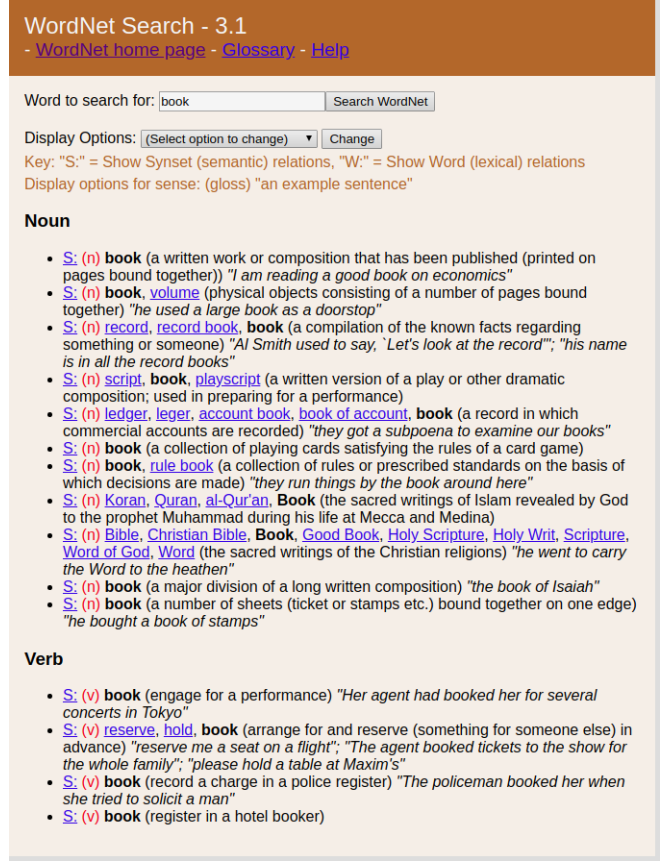
EuroWordNet, birçok Avrupa dili (Felemenkçe, İtalyanca, İspanyolca, Almanca, Fransızca, Çekçe ve Estonca) için kelime ağlarının bulunduğu çok dilli bir veritabanıdır. Kelime ağları, aralarındaki temel anlamsal ilişkilere sahip olan eş anlam kümeleri açısından WordNet ile aynı şekilde yapılandırılmıştır. Bu kelime ağları WordNet'e dayanarak, bir dillerarası dizine bağlanır. Bu indeks vasıtasıyla, diller birbirine bağlanır, böylece bir dildeki kelimelerden başka bir dildeki benzer kelimelere geçilebilir.

Arapça WordNet, EuroWordNet gibi Princeton'ın WordNet'i baz alınarak oluşturulmuştur. Hem EuroWordNet'te hem de Princeton WordNet'te doğrudan karşılığı bulunabilir. Bu kelime ağında standart gösterimin yanı sıra, kelime anlamları birinci dereceden mantıkla makine tarafından anlaşılabilir bir semantik ile de tanımlanır.

Bahsedilen bu çalışmalar, WordNet kadar kapsamlı değildir. Bugüne kadarki en kapsamlı çalışma, Balkan dilleri için ayrı kelime ağlarından oluşan çok dilli bir sözlük veritabanının geliştirilmesini amaçlayan BalkaNet'tir [5]. BalkaNet projesi Bulgarca, Çekçe, Yunanca, Rumence, Türkçe ve Sırpça olmak üzere altı dili kapsamaktadır. Bu projenin bir parçası olarak, Türkçe sözlükten eş anlam, zıt anlam ve alt anlam ilişkileri otomatik olarak çıkarılmıştır. [6]

110000 sözcük anlam çifti ve 80000 anlam kümesi içeren KeNet [7], Türkçe kelime ağlarının en kapsamlısıdır.

Var olan kelime ağlarından bir çevrimiçi kullanıcı arayüzüne sahip olan WordNet'tir. Bu arayüz, sözcüklerin anlamlarının ve anlamsal ilişkilerinin görüntülenmesi ve bunlar arasında gezilmesini mümkün hale getirir. WordNet arayüzü, Şekil 1'de görülmektedir. Bu çalışmada WordNet arayüzü temel alınarak KeNet kelime ağı için bir arayüz geliştirilmiştir.



Şekil 1: WordNet arayüzü

III. KENET

KeNet, Türkçe için yapılmış kapsamlı bir kelime ağıdır. Sıfırdan Türk Dil Kurumunun Güncel Türkçe Sözlüğü kullanılarak hem manuel hem de otomatik olarak bir kelime ağı oluşturulmuştur.

Kelime ağının manuel oluşturulması için öncelikle otomatik eş anlam kümeleri kullanılmıştır. Güncel Türkçe Sözlüğü'nden [8] her kelime için olası bütün anlamlar çıkarılmıştır ve her biri eş anlam kümesi olarak değerlendirilmiştir. İnsanlar tarafından anlamlar listesinden eşleşen anlamlar seçilerek bu eş anlam kümeleri indirgenmiştir. Zıt anlam, üst anlam ve konu gibi anlamsal ilişkiler içinse Güncel Türkçe Sözlük ve Vikipedi'den [9] kurallar çıkarılmış ve bu kurallarla elde edilen adaylar insanlar tarafından gözden geçirilmiştir.

Otomatik olarak bir kelime ağı oluşturmak için yine Güncel Türkçe Sözlük ve belirlenimci kurallar kullanılmıştır. Bu kurallar, sözlükteki belirli bir anlam tanımından olası ilişkili kelimeleri oluşturur. Eş anlamlı kelimeleri bulabilmek için bu kelimelerden bir çizge oluşturulmuştur ve bağlı olduğu diğer kelimeler de göz önüne alınarak eş anlam kümeleri yaratılmıştır.

IV. ARAYÜZ

KeNet arayüzü, kelime ağında temel gezinme yöntemlerini desteklemek için tasarlanmış bir arayüzdür. Bu arayüz kullanılarak, bir sözcüğün dahil olduğu eş anlam kümelerindeki

diğer kelimeler ve kümenin çizge yapısında giriş bölümünde sıralanan herhangi bir ilişki ile bağlı olduğu eş anlam kümeleri görüntülenebilir. Bu amaç doğrultusunda, ilk atılacak adım Şekil 2’de görülebilen arama çubuğu yardımıyla istenen sözcüğün dahil olduğu eş anlam kümesini ağaçta aramaktır. Eğer aranan sözcük, herhangi bir eş anlam kümesine dahilse, ait olduğu küme veya kümelerin tanımı, çizgede bağlı olduğu diğer kümeler, kümenin hangi sözcük türüne ait olduğu bilgileri bir ağaç yapısında görünecektir. Bu yapıda gezinmek için, her bir eş anlam kümesi en alt anlamına kadar açılıp en alt anlamdaki sözcüklere tıklanabilir. Bu tıklama sonucunda, eğer tıklanan kelime Türkçe ise istenen kelime tıpkı ilk adımdaki gibi KeNet’te aranacak ve sonuçlar aynı yapıda gösterilecektir. Aksi halde, kelime WordNet’te aranacak ve karşılık gelen internet sayfası açılacaktır.

Arayüzün kullanım akışının daha iyi aktarılması adına, iki adet örnek akış gösterebiliriz. Bu akışlardan ilkinde aranan kelimenin alt ve üst eş anlam kümelerini gezerken WordNet ile olan bağlantısını da ortaya koyacağız. Diğer akışta ise, parça-bütün ilişkilerini inceleyip, çizgede yatay bir eksenle gezeceğiz. Her ne kadar bu iki akış arayüzün bütün özelliklerini kapsamasa da, arayüzün genel işleyişi hakkında fikir verecektir.

A. Örnek Akış 1

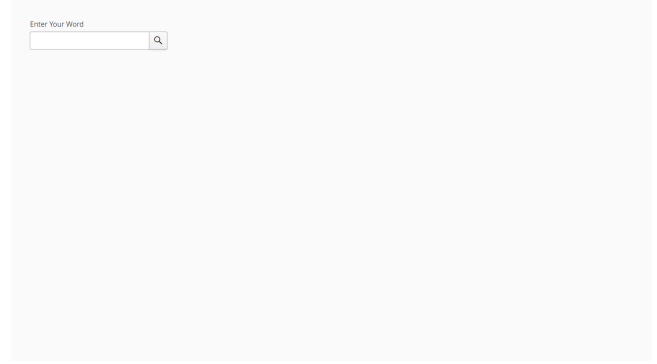
Arayüzün bu kullanım akışını Şekil 2, 3, 4 ve 5’te görebiliriz. Bu akışta, KeNet’te ilk olarak *nazım* kelimesi aranmış ve bu aramanın sonucunda ortaya çıkan ağaç yapısı Şekil 3’de gösterilmiştir. Şekilde de görüldüğü üzere *nazım* kelimesinin sözlükteki üç farklı karşılığı ağaçta kardeş olacak şekilde gösterilirken, bu karşılıkların ilişkili olduğu eş anlam kümeleri çocuklar olarak gösterilmiştir. Örneğin, ilk eş anlam kümesi incelediğinde, *nazım*’ın birer özel çeşidi olan *mersiye*, *naat*, *semai*, *nazire* ve *lirik şiir* alt anlamlar arasında sıralanmış, ilişkili olduğu *yazın* konusu ise benzer şekilde gösterilmiştir. Yine bu eş anlam kümesine karşılık gelen İngilizce eş anlam kümeleri ise, WordNet’teki kimlikleri ile birlikte gösterilmiştir. Bu aşamada, İngilizce maddelerden herhangi birine tıkladığı takdirde, WordNet’te bu maddeye karşılık gelen sayfa açılacaktır.² Örnek olarak, *poem* maddesine tıkladığında açılan sayfa, Şekil 5’te gösterilmiştir.

Bu aramanın ardından, alt eş anlam kümelerinden birisi olan *mersiye* tıklanmış ve bu aramanın sonucu da Şekil 4’te gösterilmiştir. Beklendiği gibi, *mersiye* anlam kümesinin üst kümeleri listesinde *nazım*’ı da içinde bulunduran eş anlam kümesi bulunmaktadır. Bir önceki adımla örtüşen bir biçimde, bu şekilde de eş anlam kümesinin bağlantılı olduğu konu ve WordNet’teki karşılığı görülebilmektedir.

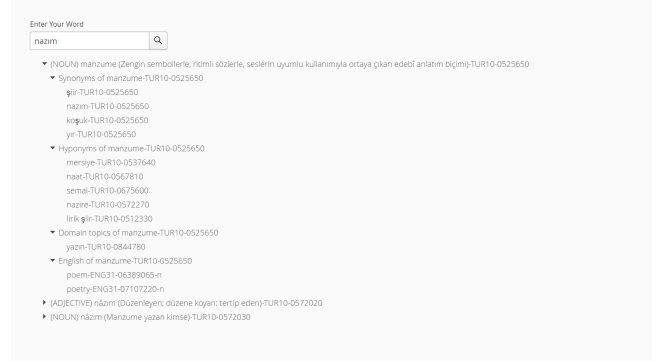
B. Örnek Akış 2

Arayüzün bu kullanım akışını Şekil 6, 7 ve 8’de görebiliriz. Bu akışta, KeNet’te ilk olarak *bitki* kelimesi aranmış ve bu aramanın sonucunda ortaya çıkan ağaç yapısı Şekil 6’da gösterilmiştir. Şekilde de görüldüğü üzere, *nazım* kelimesinin

²Eğer tarayıcı ayarlarında beliren pencereler engelleniyorsa, bu özelliğin çalışması için bu ayarın değiştirilmesi gerekebilir



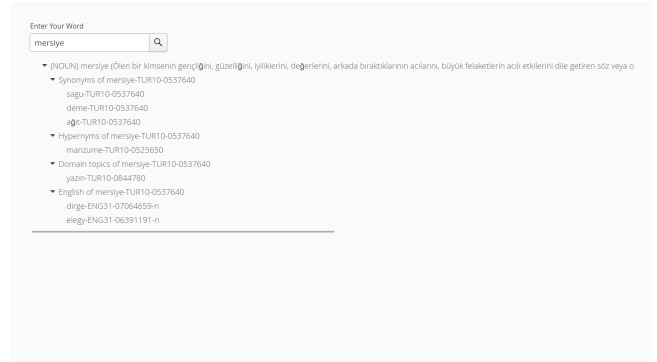
Şekil 2: Arayüzün Başlangıç Durumu



Şekil 3: Nazım Aramasının Sonucu

aksine *bitki* kelimesi sadece bir tane eş anlam kümesiyle eşleştirilmiştir. Ek olarak, akış 1’de görülemeyen kısımlık bildiren parça bütün ve kısımlık bildiren bütün parça ilişkileri Şekil 6’da görülebilmektedir. Akış 1’e benzer olarak ise, alt/üst eş anlam kümeleri, ilişkili konu ve eş anlam kümesinin WordNet karşılığı bu şekilde de görülebilmektedir.

Şekil 6 incelendiğinde *bitkinin yeşillik* ile kısımlık bildiren bütün parça ilişkisine sahip olduğu gözlemlenebilir. Bu ilişkiye tıkladığında *yeşillik* KeNet’te aranacak ve ortaya Şekil 7’deki ağaç yapısı çıkacaktır. Bu ağaçta, beklendiği üzere *yeşillik* kelimesinin *bitki* ile kısımlık bildiren parça bütün ilişkisine sahip olduğu görülebilir. Yine önceki şekillere benzer şekilde, *yeşillik* kelimesinin ait olduğu farklı eş anlam kümeleri, alt/üst eş anlam kümeleri ve WordNet karşılığı da bu ağaçta yer almaktadır.



Şekil 4: Mersiye Maddesine Tıklanmasının Sonucu

WordNet Search - 3.1

[WordNet home page](#) - [Glossary](#) - [Help](#)

Word to search for:

Display Options:

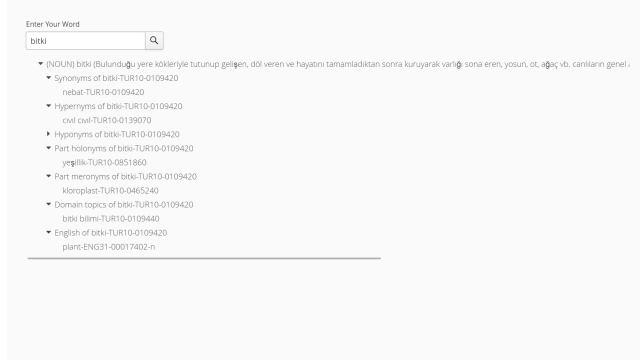
Key: "S:" = Show Synset (semantic) relations, "W:" = Show Word (lexical) relations
Display options for sense: (gloss) "an example sentence"

Noun

- **S: (n)** **poem**, [verse form](#) (a composition written in metrical feet forming rhythmical lines)

Şekil 5: Poem Maddesine Tıklanmasının Sonucu

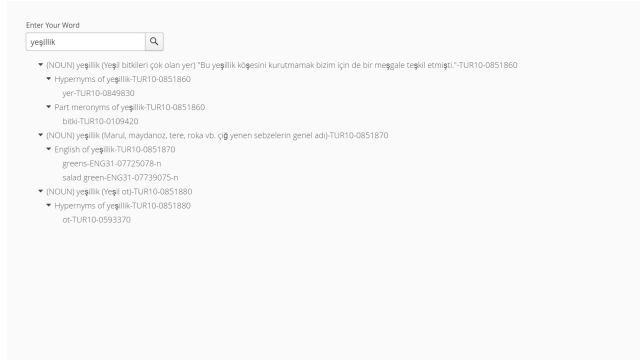
Yine Şekil 6 incelendiğinde *bitki* ile *kloroplast* arasında kısmılık bildiren parça bütün ilişkisi olduğu fark edilebilir. Bu durumda, *kloroplasta* tıkladığında ekrana Şekil 8'deki ağaç gelecektir. Bu ağaçta dikkat edilmesi gereken en önemli nokta ise *kloroplast* ile *bitki* arasında kısmılık bildiren bütün parça ilişkisi olmasıdır. Tıpkı bir önceki adımdaki gibi parça-bütün ilişkisi *bitki* ve *kloroplast* açısından farklı isimlere sahiptir.



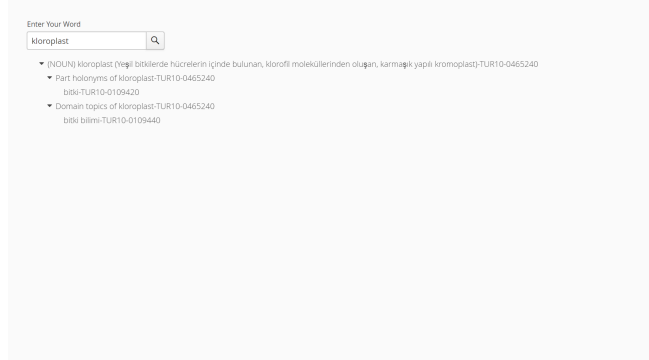
Şekil 6: Bitki Aramasının Sonucu

V. SONUÇ

Bu çalışmada WordNet için tasarlanan arayüzden yola çıkarak Türkçe kelime ağı KeNet için bir arayüz geliştirdik. Geliştirdiğimiz bu arayüz sayesinde, kelime ağı üzerindeki eş anlam kümelerinde, kümeler arasındaki çeşitli anlamsal bağlantıları kullanarak çevrimiçi olarak gezinmek mümkündür. Hatta, KeNet ve WordNet arasındaki bağlantılar kullanılarak,



Şekil 7: Yeşillik Maddesine Tıklanmasının Sonucu



Şekil 8: Kloroplast Maddesine Tıklanmasının Sonucu

bir eş anlam kümesinin WordNet karşılığına da erişilebilir. Her ne kadar arayüz şu an için KeNet içinde var olan her anlamsal ilişkiyi kullanmayı desteklemese de, hali hazırda desteklenen ilişkiler KeNet hakkında genel bir bilgi edinmek için oldukça uygundur.

Geliştirdiğimiz arayüzü daha iyi hale getirmek için atılacak adımlar arasında ilk akla gelen daha önce de bahsettiğimiz gibi KeNet'teki diğer anlamsal ilişkileri de desteklemek olacaktır. Ancak bundan daha önemli bir katkı, İngilizce olan WordNet'in de KeNet ile beraber gezilebilmesine olanak sağlamaktır. Bu şekilde, kullanıcı iki kelime ağı arasında kolaylıkla geçiş yapabilecek ve iki çizgeyi beraber dolaşabileceklerdir.

Arayüze yapılabilecek diğer bir geliştirme ise, hali hazırda var olan kelime ağı gezinme programını bir uygulama programlama arayüzü ile sarmaktır. Bu işlem, var olan programı, diğer bilgisayar programları tarafından daha kolaylıkla erişilebilir hale getirecektir.

TEŞEKKÜRLER

Bu çalışma Tübitak 116E104 nolu proje tarafından desteklenmiştir.

KAYNAKLAR

- [1] G. A. Miller, R. Beckwith, C. Fellbaum, D. Gross, and K. J. Miller, "Introduction to wordnet: An on-line lexical database," *International journal of lexicography*, vol. 3, no. 4, pp. 235–244, 1990.
- [2] KeNetUI, "KeNet için arayüz," <http://haydut.isikun.edu.tr/wordnet.ui-1.0/>, accessed: 2019-02-08.
- [3] P. Vossen, "Eurowordnet: a multilingual database for information retrieval," 1997.
- [4] W. Black, S. Elkateb, H. Rodriguez, M. Alkhalifa, P. Vossen, A. Pease, and C. Fellbaum, "Introducing the arabic wordnet project," in *Proceedings of the third international WordNet conference*. Citeseer, 2006, pp. 295–300.
- [5] D. Tufis, D. Cristea, and S. Stamou, "Balkanet: Aims, methods, results and perspectives. a general overview," *Romanian Journal of Information science and technology*, vol. 7, no. 1-2, pp. 9–43, 2004.
- [6] O. Bilgin, Ö. Çetinoğlu, and K. Oflazer, "Building a wordnet for turkish," *Romanian Journal of Information Science and Technology*, vol. 7, no. 1-2, pp. 163–172, 2004.
- [7] R. Ehsani, E. Solak, and O. T. Yildiz, "Constructing a wordnet for turkish using manual and automatic annotation," *ACM Transactions on Asian and Low-Resource Language Information Processing (TALLIP)*, vol. 17, no. 3, p. 24, 2018.
- [8] TDK, "Güncel türkçe sözlük," http://www.tdk.gov.tr/index.php?option=com_gts, accessed: 2019-02-06.
- [9] "Vikipedi," <https://tr.wikipedia.org/>, accessed: 2019-02-06.