Hacettepe Üniversitesi Sosyal Bilimler Enstitüsü

Bilgi ve Belge Yönetimi Anabilim Dalı

ATIF DİZİNLERİNDE ÜNİVERSİTE ADRESLERİNİN STANDARDİZASYON SORUNU

Zehra Taşkın

Yüksek Lisans Tezi

Ankara, 2012

ATIF DİZİNLERİNDE ÜNİVERSİTE ADRESLERİNİN STANDARDİZASYON SORUNU

Zehra Taşkın

Hacettepe Üniversitesi Sosyal Bilimler Enstitüsü

Bilgi ve Belge Yönetimi Anabilim Dalı

Yüksek Lisans Tezi

Ankara, 2012

KABUL VE ONAY

Zehra Taşkın tarafından hazırlanan "Atıf Dizinlerinde Üniversite Adreslerinin Standardizasyon Sorunu" başlıklı bu çalışma, 8 Haziran 2012 tarihinde yapılan savunma sınavı sonucunda başarılı bulunarak jürimiz tarafından yüksek lisans tezi olarak kabul edilmiştir.

Prof. Dr. İrfan Çakın (Başkan)

Prof. Dr. İlyas Çiçekli

Prof. Dr. Yaşar Tonta

Prof. Dr. Nazan Uçak

Doç. Dr. Umut Al (Danışman)

Yukarıdaki imzaların adı geçen öğretim üyelerine ait olduğunu onaylarım.

Prof. Dr. Ş. Armağan Tarım

Enstitü Müdürü

BİLDİRİM

Hazırladığım tezin/raporun tamamen kendi çalışmam olduğunu ve her alıntıya kaynak gösterdiğimi taahhüt eder, tezimin/raporumun kağıt ve elektronik kopyalarının Hacettepe Üniversitesi Sosyal Bilimler Enstitüsü arşivlerinde aşağıda belirttiğim koşullarda saklanmasına izin verdiğimi onaylarım:

- ■ Tezimin/Raporumun tamamı her yerden erişime açılabilir.
- ☐ Tezim/Raporum sadece Hacettepe Üniversitesi yerleşkelerinden erişime açılabilir.
- ☐ Tezimin/Raporumun …… yıl süreyle erişime açılmasını istemiyorum. Bu sürenin sonunda uzatma için başvuruda bulunmadığım takdirde, tezimin/raporumun tamamı her yerden erişime açılabilir.

08.06.2012

Zehra Taşkın

Biricik eşim ve minicik oğluma…

TEŞEKKÜR

Bu çalışmada yardım ve desteğini gördüğüm pek çok kişiye teşekkür borçluyum. İlk olarak çalışmanın her aşamasında değerli fikirlerini paylaşarak katkıda bulunan, defalarca tezi okumasına rağmen hiç şikayet etmeyen ve tezimin zamanında bitmesinde önemli destekleri olan değerli hocam ve danışmanım Doç. Dr. Umut Al'a sonsuz teşekkür ederim.

Çalışmamı okuyarak yorumlarını benimle paylaşan değerli hocalarım Prof. Dr. Yaşar Tonta ve Dr. İrem Soydal'a, ayrıca jürimde yer alarak değerli görüşlerini ileten ve çalışmanın daha iyi yerlere gelmesini sağlayan Prof. Dr. İrfan Çakın, Prof. Dr. İlyas Çiçekli ve Prof. Dr. Nazan Uçak'a teşekkürü borç bilirim.

Tezin en can alıcı noktasında 200.000 verinin temizlenmesi aşamasında bana yardımcı olan ahretliğim Sema Özgür'e, çalışma arkadaşlarım Tolga Çakmak ve İpek Şencan'a, güzel insan Şule Karacan'a; ve yaz tatili demeden yardımıma koşan başta Müge Akbulut olmak üzere tüm iyi kalpli ve yardımsever gönüllülerime (Meriç Dirik, Sibel Akın, Canan Cevahir, Yağmur Ege Zeybekoğlu, Sefa Mustafa Anwar, Tuğçe Erdoğan, Melisa Gelbal, Mehmet Laçinkaya, Ömer Çetin, Merve Ceren Tüzün) sonsuz teşekkür ederim.

Hiç bilmediğim bir tekniği anlama konusunda bana her zaman yol gösteren ve düşünmemi sağlayan sevgili arkadaşım Esra Özkan Çelik'e ve tez tamamlanana kadar annesine hiç sıkıntı vermeyerek bana yardımcı olmasını kolaylaştıran Doğu Çelik'e minnettarım.

Her sıkıntılı anımda ve aklıma bir soru takıldığında bana yardımcı olan, tüm sorularıma özenle cevap veren ve her konuda desteğini hissettiğim sevgili oda arkadaşım ve dostum Güleda Doğan'a teşekkürü borç bilirim.

Bu çalışma ile ilgili her türlü endişemi ve stresimi aşmamda yardımcı olan, her zaman desteğini hissettiğim, hem iyi bir eş, hem iyi bir baba, hem de çok iyi bir destekçi olan sevgili eşim Abdülkadir Taşkın'a gösterdiği tüm özveri ve anlayış için sonsuz teşekkür ederim. Sevgili annem, babam ve kardeşlerime de tüm destekleri için minnettarım.
Son olarak biricik oğlum Emir, sen bu kadar akıllı ve uslu bir çocuk olmasan bu tez bitmezdi. Benim oğlum olduğun için ve çalışma boyunca bana hiç sıkıntı vermediğin için teşekkürlerin en büyüğü sana…

ÖZET

TAŞKIN, Zehra. Atıf Dizinlerinde Üniversite Adreslerinin Standardizasyon Sorunu, Yüksek Lisans Tezi, Ankara, 2012.

Üniversitelerin etkinliğinin değerlendirilmesinde atıf dizinlerinde yer alan yayınlarının sayısı ölçüt olarak kullanılmaktadır. Ancak atıf dizinlerinde doğal dil ile yapılan indeksleme ya da yazarın veya editörün gözünden kaçan hatalar gibi bazı nedenlerle adres ve yazar alanlarında standardizasyon sorunları doğabilmektedir. Bu sorunlar, üniversitelerin uluslararası görünürlüğünü azaltmakta, değişik üniversite sıralamaları ortaya çıkarmakta ve bibliyometrik çalışmalarda farklı sonuçlara neden olabilmektedir. Bu durum atıf dizinleri için düzgün yapılmış bir adres standardizasyonunu önemli hale getirmektedir.

Bu çalışmada Türkiye'de eğitim veren üniversitelerin atıf dizinlerinde yer alan yayınlarının adreslerinde yapılan hataların tespiti, etkilerinin ölçülmesi ve hataları saptayabilen bir standardizasyon algoritması geliştirilmesi hedeflenmiştir. Bu doğrultuda temel hipotez "sonlu durum tekniği aracılığı ile atıf dizinlerinde yapılmış hatalar saptanabilir" şeklinde belirlenmiştir. Bu hipotezi test etmek amacı ile atıf dizinlerinde yer alan Türkiye adresli toplam 198.687 yayın incelenmiş ve tüm adresler üzerinden hata türleri ve bu hataların ortaya çıkardığı etkiler saptanmıştır. Ardından mevcut hataları saptayacak bir algoritma geliştirilerek üniversitelere ait yayınların kaybının önüne geçilmesi hedeflenmiştir.

Çalışmada üniversite adreslerinde yapılan hatalar ile üniversite isminde yer alan karakter sayısı, sözcük sayısı ya da Türkçe karakter sayısı arasında anlamlı bir ilişki olmadığı sonucuna ulaşılmıştır. En sık rastlanan hata türlerinin dizinleme, harf ve heceleme ile Türkçe isimlerin İngilizceye çevrilmesi olduğu sonucuna ulaşılmıştır.

Çalışma kapsamında Hacettepe Üniversitesi için sonlu durum tekniği ve Nooj sonlu durum dönüştürücüsü kullanılarak bir algoritma geliştirilmiş ve bu sayede Hacettepe Üniversitesi için atıf dizinlerinde 330 kez kullanılmış 49 farklı adres girişi saptanabilmiştir. Gözle yapılan kontroller sonucunda ise algoritmanın atıf dizinlerinde

10 kez kullanılan 10 farklı adres girişini belirleyemediği görülmüştür. Çalışmada elde edilen sonuçlar ana hipotezi kanıtlar niteliktedir. Sonlu durum tekniği aracılığı ile birçok üniversitenin yayınlarında yapılmış hatalar saptanabilmektedir. Türkiye kütüphanecilik ve bilgibilim literatüründe daha önce sonlu durum tekniği kullanılmadığından bu çalışma alanımızda ilk olma özelliği taşımaktadır.

Bu çalışma ile ortaya koyulan bir diğer önemli konu da Yükseköğretim Kurulu tarafından her yıl yapılan ve pek çok üniversite sıralamasının temelini oluşturan yayın sayısı sıralamalarında farklılıklar görülmesi ve bu farklılıkların büyük ölçüde adres hatalarına dayanıyor olmasıdır. Bu nedenle de yöneticiler ve karar vericiler yayın sayısına göre yapılan niceliksel değerlendirmelerin hatalı olabileceğini göz önüne almalı ve buna göre düzenlemeler gerçekleştirmelidirler.

Atıf dizinlerinde yer alan hatalar ile ilgili olarak yazarlar, editörler, kütüphaneciler, yöneticiler ve karar vericiler ile atıf dizinlerini sağlayan kurumlara büyük görevler düşmektedir. Bu görevler layıkıyla yerine getirildiğinde atıf dizinlerinde var olan standardizasyon sorunu en aza indirilebilecektir.

Anahtar Sözcükler

Atıf Dizinleri, Adres Standardizasyon Sorunu, Adres Birleştirimi, Sonlu Durum Tekniği, Nooj.

ABSTRACT

TAŞKIN, Zehra. *Standardization Problem of University Addresses in Citation Indexes*. Master's thesis, Ankara, 2012.

Effectiveness of academia has usually been evaluated by using the number of publications that took part in the citation indexes. However, indexing with natural language and mistakes made by authors or editors in the citation indexes make evaluation process problematic for author names and affiliations. Consequently, international visibility of the organizations in the citation indexes can reduce, different rankings for organizations can emerge and bibliometric studies can generate inaccurate results. Under these circumstances, the importance of well-structured address standardization becomes significant.

The main aim of this study is to identify mistakes on Turkish Universities' publication addresses, to measure the influences of these mistakes and to develop a standardization algorithm which can determine them. Accordingly, the main hypothesis is "the mistakes on citation indexes can be determined by using finite state technique". To test this hypothesis, we gathered 198,687 Turkey addressed publications from citation indexes. Types and effects of the address mistakes were clarified. Subsequently, an algorithm, which was thought to prevent the missing Turkey addressed publications, was developed to determine the address mistakes.

Findings indicate that there are no statistically significant correlations between character count, word count or Turkish character count and address mistakes. Mistakes generally depend on inexact indexing and wrong spelling, abbreviation or translation.

The algorithm for Hacettepe University is created by using finite state technique and Nooj finite state transducer. By this means, 49 different address variations was determined for Hacettepe University which was used 330 times. It was observed that only 10 different addresses (which were used 10 times) could not be specified by the algorithm. Therefore, it is quite clear that the results proved our main hypothesis. Finite state technique can be applied to universities that have distinctive names to find

mistakes on the address fields of citation indexes. This study is the first one that uses the finite state technique in the Turkish library and information science literature.

This study also reveals the dissimilarities in the number of publications of different universities, which creates the base of university rankings that is announced every year by the Council of Higher Education. According to our findings, these inconsistencies were mostly depend on address mistakes in the citation indexes. Therefore, managers and policy makers should consider the possibility of errors in the number of publications and they should perform regulations accordingly.

Address mistakes in the citation indexes concern authors, editors, librarians, managers and decision makers and they should be aware of the consequences of this phenomenon and should make provisions in order to minimize this standardization problem.

Keywords

Citation Indexes, Address Standardization Problem, Unification of Addresses, Finite State Technique, Nooj.

İÇİNDEKİLER

KABUL VE ONAY ii
BİLDİRİM iii
TEŞEKKÜR IV
ÖZET V
ABSTRACT VII
İÇİNDEKİLER IX
KISALTMALAR DİZİNİ XI
ŞEKİLLER DİZİNİ XII
TABLOLAR DİZİNİ XIII
1. BÖLÜM: GİRİŞ 1
1.1. KONUNUN ÖNEMİ 1
1.2. ARAŞTIRMANIN AMACI VE PROBLEMİ 3
1.3. ARAŞTIRMA SORULARI VE HİPOTEZLER 4
1.4. ARAŞTIRMANIN KAPSAMI 4
1.5. ARAŞTIRMANIN DÜZENİ 5
1.6. KAYNAKLAR 6
2. BÖLÜM: LİTERATÜR DEĞERLENDİRMESİ 7
2.1. GİRİŞ 7
2.2. SONLU DURUM TEKNİĞİ (FINITE STATE TECHNIQUE) 7
2.2.1. Sonlu Durum Tekniği Kullanılan Çalışmalar 9
2.2.1.1. Kütüphanecilik ve Bilgibilim Alanında Sonlu Durum Tekniği Kullanılan Çalışmalar 13
2.2.1.2. Kütüphanecilik ve Bilgibilim Alanında Adres ve Yazar Adı Standardizasyonu Üzerine Yapılan Diğer Çalışmalar 17
3. BÖLÜM: YÖNTEM VE TEKNİK 20
3.1. GİRİŞ 20
3.2. VERİLERİN TOPLANMASI VE ANALİZE UYGUN HALE GETİRİLMESİ 20
3.3. STANDARDİZASYON SORUNUNUN SAPTANMASI VE BÜYÜKLÜĞÜNÜN ÖLÇÜMÜ 22
3.4. SONLU DURUM DÖNÜŞTÜRÜCÜSÜNÜN UYGULANMASI 23

4. BÖLÜM: BULGULAR VE DEĞERLENDİRME 24
4.1. GİRİŞ 24
4.2. ATIF DİZİNLERİNDE ÜNİVERSİTE ADRESLERİNDE YAPILAN HATALAR 28
4.2.1. Harf ve Heceleme Hataları 28
4.2.2. Dizinleme Hataları 29
4.2.3. Üniversite İsimlerinin İngilizce Çevirilerinde Yapılan Hatalar 34
4.2.4. Diğer Hatalar 35
4.3. ÜNİVERSİTE ADRESLERİNDE STANDART YAPI SORUNU 37
4.4. ADRESLERDE YAPILMIŞ HATALAR VE STANDARDİZASYON SORUNLARININ ETKİLERİ 40
4.5. ÜNİVERSİTE ADRESLERİ STANDARDİZASYON SORUNU İÇİN ÇÖZÜM ÖNERİLERİ 43
4.6. TARTIŞMA VE YORUM 55
5. BÖLÜM: SONUÇ VE ÖNERİLER 57
5.1. SONUÇ 57
5.2. ÖNERİLER 59
KAYNAKÇA 64
EK 72

KISALTMALAR DİZİNİ

A&HCI	Arts and Humanities Citation Index
GATA	Gülhane Askeri Tıp Akademisi
LISA	Library and Information Science Abstracts
OCR	Optical Character Recognition (Optik Karakter Tanıma)
ODTÜ	Orta Doğu Teknik Üniversitesi
SCI	Science Citation Index
SSCI	Social Sciences Citation Index
TÜBİTAK	Türkiye Bilimsel ve Teknolojik Araştırma Kurumu
UBYTP	Uluslararası Bilimsel Yayınları Teşvik Programı
ULAKBİM	Ulusal Akademik Ağ ve Bilgi Merkezi
URAP	University Ranking by Academic Performance
WoS	Web of Science
YÖK	Yükseköğretim Kurulu

ŞEKİLLER DİZİNİ

Şekil 1: Metro turnikesi örneği (Scholl, 2008) 8
Şekil 2: Melvon Ankeny isminin farklı kullanımlarını saptamak üzere hazırlanmış bir sonlu durum grafiği (Galvez ve Moya-Anegón, 2007b) 16
Şekil 3: Fırat Üniversitesi adresli yayının orijinal metninde geçen adresler 30
Şekil 4: Şekil 3'te yer alan yayının *WoS* kaydı 31
Şekil 5: "Uslu Univ" adresli yayının *WoS* kaydı 33
Şekil 6: Şekil 6'da verilen kaydın orijinal metni 33
Şekil 7: Harran Üniversitesi için kullanılan HU kısaltması örneği 38
Şekil 8: Hacettepe Üniversitesi için kullanılan HU kısaltması örneği 38
Şekil 9: Macaristan'ın ülke kodu HU'nun adres alanında kullanımı 39
Şekil 10: Kahramanmaraş Sütçü İmam Üniversitesi kurumsal işbirliği ağı 41
Şekil 11: Standardize edilmiş Kahramanmaraş Sütçü İmam Üniversitesi verileri ile kurumsal işbirliği ağı 42
Şekil 12: Nooj yeni metin dosyası yaratma ekranı 44
Şekil 13: Türkiye adresli yayınlar için oluşturulan .not dosyası 44
Şekil 14: Nooj yeni gramer yaratma başlangıç ekranı 45
Şekil 15: Hacettepe teriminin doğru yazımlarını getirecek sonlu durum grafiği 46
Şekil 16: Harflerin eksik ya da fazla yazılması ile oluşan hataları saptayan grafik 47
Şekil 17: Hacettepe teriminin hatalı ve doğru girişlerinin saptanması için tasarlanan sonlu durum grafiği 48
Şekil 18: .nom dosyasının .not dosyası üzerinde uygulanması 48
Şekil 19: Tanımlanmış olası adresler ile eşleşen sözcüklerin gösterildiği bağımlı dizin (concordance) 49
Şekil 20: Hacettepe Üniversitesinin farklı kullanımlarını saptamak üzere geliştirilen sonlu durum grafiği 51
Şekil 21: Syntax ile oluşturulmuş grafiğin metin dosyası üzerine uygulanması 52
Şekil 22: Hacettepe Üniversitesi adresleri için yaratılan bağımlı dizin 53

TABLOLAR DİZİNİ

Tablo 1: En çok yayını olan ilk 20 üniversitenin hatalı adresten yayın kaybı 24
Tablo 2: 2009 yılı yayın istatistiklerinin karşılaştırma tablosu 27
Tablo 3: Sonlu durum grafiği ile erişilen terimler ve sıklıkları 49
Tablo 4: Dönüştürücünün saptayabildiği Hacettepe Üniversitesi adresleri 53
Tablo 5: Dönüştürücünün saptayamadığı Hacettepe Üniversitesi adresleri 54

1. BÖLÜM: GİRİŞ

1.1. KONUNUN ÖNEMİ

Günümüzde pek çok kurum veya kuruluş atıf dizinlerinde yer alan yayınlarının sayısı ile değerlendirilmektedir. Ancak niceliksel değerlendirmeye dayanan bu yayın sayısı saptama işi atıf dizinlerinde yer alan alanlarda yapılan hatalar nedeni ile sıkıntılar yaratmaktadır.

Atıf veri tabanlarından araştırmacıların literatür takiplerini kolaylaştırmak amacı ile yararlanıldığı gibi bu veri tabanları atıf analizi yapmak üzere de kullanılmaktadır. Günümüzde atıf analizi çalışmaları araştırmacıların entelektüel etkisi ve bilimsel çalışmaların kalitesini ölçmek amacı ile yapılmaktadır (Cole, 2000).

Atıf veri tabanlarının kullanımının yaygınlaşması ile birlikte içeriklerinde de artış yaşanmış ve bu durum tutarsızlık ile ilgili sorunları beraberinde getirmiştir (Galvez ve Moya-Anegón, 2006a). Bu tutarsızlıklardan biri yazarın bağlı olduğu kuruma ait adres bilgisinin (affiliation) yazımında yapılan heceleme, çeviri ve kısaltma gibi hatalardır (Galvez ve Moya-Anegón, 2007a). Yapılan bu adresleme hataları bazı problemleri de beraberinde getirmektedir. Adresleme hataları dolayısıyla üniversiteler ve ülkeler arasında bilimsel işbirliklerinin gösterimi ve değerlendirilmesinde problemler yaşanmakta, bilimsel alanların sınırlandırılması konusunda eksiklikler ortaya çıkmakta, en önemlisi performans değerlendirmelerinin doğru bir şekilde yapılamamasına neden olmaktadır (Galvez ve Moya-Anegón, 2006a).

Bilimsel çalışmalar, çalışmayı yayımlayan yazarın/yazarların adres bilgisini içermektedir. Makalenin başlığında, genellikle dipnot olarak kurumsal ve coğrafik adresler verilmektedir. Başlangıçta bu bilgiler, yazar ve okuru arasında iletişim sağlamak üzere yapılmışsa da, yayınların araştırma değerlendirmede kullanımıyla birlikte, söz konusu bilgilerin kullanım

amacında değişiklik görülmesine neden olmuştur. Çünkü her bir adres bilgisi bölümler, laboratuvarlar ve birimler için önemli ve yaşamsal hale gelmiştir (De Bruin ve Moed, 1992). Adres alanında yer alan bilgiler yazarın hangi adresi kullanmak istediği ile başlar ve yayıncının bunu nasıl biçimlendirdiği ile devam eder. Aynı kurumda ve hatta aynı bölümde çalışan kişiler bile aynı adresi farklı şekillerde gösterebilmektedirler. Bu da atıf dizinlerinde aynı adres için farklı yüzlerce kaydın oluşmasına neden olmaktadır (Moed, 2005: 183-184). Özellikle yayın üretimine bağlı olarak bütçelerin belirlendiği yapılarda, adres hatası nedeni ile görünürlüğü kaybolan çalışmalar kurumlar ya da araştırma gruplarının bütçe desteğini etkilemektedir.

Günümüzde Türkiye'de kurumlar atıf dizinlerinde yer alan yayınlarının sayıları üzerinden değerlendirilmektedir. Bu bakış açısı ile kimi kesimler tarafından en çok yayını olan üniversitenin en iyi üniversite olduğu düşünülmektedir. Bu durumda niteliksel bir değerlendirmeden çok niceliksel değerlendirmeler ön plana çıkmaktadır.

Pek çok performans değerlendirmesi nicel veriler üzerinden yapılmaktadır. Örneğin, yazarlara akademik yükselme alabilmeleri için atıf dizinlerinde belli sayıda yayın yapma zorunluluğu getirilmiştir (Öğretim, 2007). Bunun yanında TÜBİTAK (Türkiye Bilimsel ve Teknolojik Araştırma Kurumu), adres alanında TURKEY geçmeyen yayınları UBYTP (Uluslararası Bilimsel Yayınları Teşvik Programı) kapsamında desteklemeyeceğini belirtmiştir (ULAKBİM, 2010). Yine her yıl atıf dizinlerinde hangi Türk üniversitesinin kaçıncı sırada olduğu bilgisi YÖK (Yükseköğretim Kurulu) tarafından duyurulmaktadır (Yükseköğretim Kurulu, 2010). 2009 yılında ODTÜ (Orta Doğu Teknik Üniversitesi) Enformatik Enstitüsü bünyesinde kurulan URAP (University Ranking by Academic Performance) araştırma laboratuvarı da, içinde yayın sayısının da bulunduğu pek çok kriteri göz önüne alarak Türkiye'deki üniversitelerin sıralamalarını yayımlamaktadır (URAP, 2011).

Tüm bu uygulamalar atıf dizinlerinin kullanımlarının Türkiye'de hangi boyutta olduğunu göstermektedir. Ancak atıf dizinleri tarafından bir yayının adres bilgisi, yayında yer alan şekli kullanılarak elci sistem ile dizinlediğinden; dizinleme aşamasında çok sayıda hatanın yapılmış olması kaçınılmazdır. Bu nedenle de hatalı veriler üzerinden yapılmış analizlere dayanan bu niceliksel çalışmalar politika geliştirme ya da yönetimsel amaçlarla kullanıldığında dikkatle değerlendirilmelidir. Sayısal verilere dayanan ve bilimsel yayının kalitesini ölçmeyen bu gibi performans değerlendirmelerinde en azından kurumlara ait tüm yayınların saptanabilmesini sağlamak büyük önem arz etmektedir.

1.2. ARAŞTIRMANIN AMACI VE PROBLEMİ

Atıf dizinlerinde üniversite adreslerinin doğru bir şekilde yer alıyor olması pek çok nedenden dolayı önem taşımaktadır. Bir üniversitenin kendi yayınlarından bir kısmına adres alanında yapılmış hatalar nedeni ile erişemiyor olması başarı ölçütü olarak kabul edilen yayın sayısını etkilemektedir. Ancak daha önemlisi bu hatalar nedeni ile üniversiteler kendi perfomanslarını değerlendirme aşamasında sıkıntılar yaşamaktadır.

Bu çalışmanın temel amacı atıf dizinlerinde yer alan Türk üniversitelerinin adreslerinde yapılmış hataları saptayacak bir algoritma geliştirerek yayın kaybını en aza indirmeyi sağlamak ve gelecekte benzer hataların yapılmaması için öneriler getirmektir. Bu bağlamda öncelikle Türk üniversitelerinin adreslerinde yapılmış hataların türleri belirlenecek, bu hataların etkileri saptanacak ve bir sonlu durum algoritması geliştirilerek mevcut hataların kolaylıkla bulunması sağlanacaktır.

Türkiye'de atıf dizinlerinde adres standardizasyonu ya da birleştirimi (unification) ile ilgili az sayıda çalışma yapıldıysa da, bu çalışma sonlu durum tekniği kullanılarak yapılan ilk çalışma olacaktır. Bu nedenle çalışma pek çok üniversite kütüphanesine, bağlı oldukları üniversitenin yayınlarını saptarken yardımcı olacak nitelikte bir çözüm sunmaktadır.

1.3. ARAŞTIRMA SORULARI VE HİPOTEZLER

Bu çalışmada cevap aranacak temel araştırma sorusu "atıf dizinlerinde Türk üniversitelerine ait adres bilgilerinde yaşanan standardizasyon sorununun büyüklüğü nedir ve bu sorun nasıl çözümlenebilir?" olarak belirlenmiştir. Diğer araştırma soruları ise şunlardır:

- Atıf dizinlerinde üniversite adreslerinde yapılmış hataların türleri nelerdir?
- Bu hatalar ne gibi etkiler yaratmaktadır?
- Sonlu durum tekniği üniversite adreslerinin standardizasyon sorununa çözüm olabilir mi?

Araştırmanın temel hipotezi "sonlu durum tekniği aracılığı ile atıf dizinlerinde yapılmış hatalar saptanabilir" şeklinde belirlenmiştir. Bu ana hipotez bağlamında test edilecek alt hipotezler ise aşağıdaki gibidir:

- Atıf dizinlerinde yer alan Türk üniversitelerine ait adres girişlerinde bir standardizasyon sorunu bulunmaktadır.
- Adreslerde yapılan hataların büyüklüğü nedeni ile üniversitelere ait yayınların tamamının saptanması zorlaştığından, pek çok üniversitenin yayın sayısı sıralamasında baz olarak aldığı YÖK tarafından yapılan yıllık yayın sayısı sıralamasında hatalar vardır.
- Sonlu durum tekniği kullanılarak yaratılmış bir standart adres birleştirimi ile standardizasyon sorununun en aza indirilmesi mümkündür.

1.4. ARAŞTIRMANIN KAPSAMI

Bu araştırma *SCI (Science Citation Index), SSCI (Social Sciences Citation Index) ve A&HCI (Arts and Humanities Citation Index)* kapsamında dizinlenmekte olan 2009 yılına kadar yayımlanmış tüm Türkiye adresli yayınları kapsamaktadır. Bahsi geçen bu atıf dizinlerinde yapılan sorguda "Turkey", "Turkiye" ve "Turkei" anahtar sözcükleri kullanıldığından

Türkiye'den yazarlara sahip olan ancak adresinde Türkiye bilgisi yer almayan kayıtlar değerlendirme kapsamına alınamamıştır.

Araştırma evrenini Temmuz 2010 itibariyle YÖK'e bağlı olarak Türkiye'de eğitim öğretim faaliyeti sürdürmekte olan 146 üniversite oluşturmaktadır. Araştırmanın verileri *SCI, SSCI* ve *A&HCI* dizinlerinden sağlanmıştır.

Üniversite adreslerini saptamak amacıyla Nooj sonlu durum dönüştürücüsü kullanılarak bir hata saptayıcı algoritma tasarlanmıştır. Bu çalışmanın yöntemi hakkında detaylı bilgi Bölüm 3'te aktarılmaktadır.

1.5. ARAŞTIRMANIN DÜZENİ

Bu araştırma temel olarak beş bölüme ayrılmaktadır.

İlk bölümde atıf dizinlerinde standardizasyonun önemine değinilerek bu çalışmanın temel araştırma soruları, hipotezleri belirtilmekte ve kullanılacak yöntem hakkında kısa bir tanımlama yapılmaktadır.

İkinci bölümde üniversite adreslerindeki hataların tespitini sağlamak üzere kullanılacak sonlu durum tekniğinin literatürdeki kullanım alanları hakkında bilgi verilmekte ve kütüphanecilik alanında yapılmış bazı çalışmalardan örnekler sunulmaktadır.

Üçüncü bölümde çalışmada kullanılan yöntem ve teknikler detaylı olarak tanımlanmaktadır. Bu kısımda çalışma için kullanılan verilerin toplanmasından işlenmesine ve analizine kadar yapılan işlemler hakkında bilgi verilmektedir.

Çalışmanın dördüncü bölümünde üniversite adreslerinde saptanmış hatalar ve bu hataların etkileri sunulmakta ve yaratılan sonlu durum algoritması hakkında bilgi verilmektedir. Sonlu durum algoritmasının nasıl kullanılacağı ile ilgili bilgiler de yine bu kısımda yer almaktadır.

Son bölümde ise çalışma sonunda elde edilen bulgular değerlendirilerek adres standardizasyon sorunu için bazı öneriler getirilmektedir.

1.6. KAYNAKLAR

Çalışma ile ilişkili olan ve literatürde yer alan kaynaklara erişmek amacı ile basılı ve elektronik kaynaklardan yararlanılmıştır. Bu amaçla çalışma kapsamında sıklıkla başvurulan bilimsel dergiler ve bu dergilerin dizinlendiği veri tabanları;

- *Scientometrics* (SpringerLink),
- *Journal of the American Society for Information Science and Technology* (Wiley Online Library),
- *Computational Linguistics* (ACM Digital Library) ve
- *Journal of Documentation* (EmeraldInsight)'dır.

Yukarıda sıralanan temel dört dergi ve veri tabanının yanında *EbscoHost, Ebrary, ERIC, JSTOR, LISA (Library and Information Science Abstracts), OCLC ArticleFirst, ProQuest Dissertations and Thesis, Scopus* ve *WoS (Web of Science)* gibi veri tabanlarından da taramalar yapılarak kaynaklara erişim sağlanmıştır.

Tezin yazımında Hacettepe Üniversitesi Sosyal Bilimler Enstitüsü "Tez Yazım Yönergesi" kullanılmıştır.[1]

[1] http://www.sosyalbilimler.hacettepe.edu.tr/belgeler/ornek_tez_sablonu.doc

2. BÖLÜM: LİTERATÜR DEĞERLENDİRMESİ

2.1. GİRİŞ

Atıf dizinlerinde üniversite adreslerinin standardizasyonu büyük önem taşımaktadır. Bu çalışma kapsamında standardizasyon sorununa son yıllarda kullanım alanı gelişen sonlu durum tekniği kullanılarak çözüm bulunmaya çalışılmaktadır. Bu bölümde sonlu durum tekniğinin tanımı ve öğelerinin yanında, kullanıldığı farklı alanlardan örnekler de sunulmaktadır. Son olarak bu tekniğin kütüphanecilik ve bilgibilim literatüründe kullanıldığı çalışmalardan örnekler verilmektedir.

2.2. SONLU DURUM TEKNİĞİ (FINITE STATE TECHNIQUE)

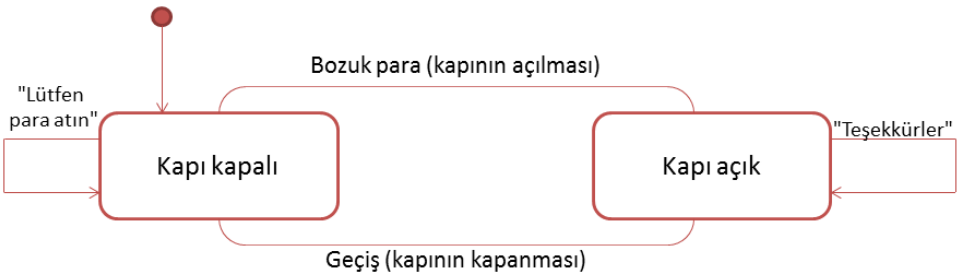
Sonlu durum algoritmaları bir dize (string) ya da tümcenin (sentence) kabul edilebilir düzenli bir dile sahip olup olmadığını belirlemek için kullanılan sistemlerdir. Algoritma soldan sağa her bir dizeyi okur ve geçişleri etiketler. Eğer bir geçiş giriş durumu ile aynı sembolle etiketlenirse, algoritma durum sonlanana kadar bir sonraki duruma geçer. Bu işlem son duruma kadar devam eder (Galvez ve Moya-Anegón, 2007a: 9).

Sonlu durum algoritmaları örüntü eşleştirme ve tanıma, konuşma işleme, el yazısı tanıma, optik karakter tanıma, şifreleme algoritması, veri sıkıştırma ve dizinleme de dâhil olmak üzere etki alanı geniş bir yelpazede kullanılmaktadır (Roche ve Schabes, 1997: 227). Aşağıda tekniğin önemli kullanım alanları sıralanmaktadır:

- Dijital devrelerin hareketlerinin tasarımı ve kontrolü için bir yazılım olarak,
- Girdi metinlerini anahtar sözcüklere, tanımlayıcılara ve noktalama işaretlerine ayıran tipik bir derleyicinin sözcüksel analizcisi (lexical analyzer) olarak,
- Web sayfalarında yer alan kelimeler, kelime grupları gibi büyük miktardaki metinsel bilginin taranabilmesi için bir yazılım olarak,

- Sonlu sayıda ve birbirinden farklı durumlara sahip uygulamalar için bir yazılım olarak kullanılmaktadır (Hopcroft, Motwani ve Ullman, 2000).

Bir sonlu durum algoritmasının çalışma mantığını metro turnikesine benzetmek olanaklıdır. Metro turnikelerinde sistem para atıldığında kapının açılması, ardından da kapanması şeklinde çalışmaktadır (bkz. Şekil 1).



Şekil 1: Metro turnikesi örneği (Scholl, 2008)

Şekil 1'de yer alan örnekte başlangıç durumunda metro turnikesinin kapısı kapalı durumdadır. Bozuk para atıldığında kapı açılır ve sistem "kapı açık" durumuna geçer. Sistem bu durumda iken "teşekkürler" uyarısı verilir ve kapı kapanır. Bir sonraki duruma, yani para atılmasına kadar kapı kapalı kalacaktır. Herhangi biri kapı kapalı durumda iken para atmaksızın geçiş yapmaya kalkarsa da bu kişiye "lütfen para atın" uyarısı belirecektir. Sonlu durum algoritmalarında da metro turnikesine benzer olarak son durum tamamlanana kadar işlem devam eder. Tanımlanan durumlara uyan sonuçlar saptanır.

Bir sonlu durum algoritması beş temel değişkenden ($\sum$, Q, q_o, F, δ) oluşur. Bu değişkenler;

- $\sum$ Giriş alfabesini
- Q Tüm durumların kümesini
- q_0 Başlangıç durumunu ($q_0 \in Q$)
- F Son durumu ya da durumları ($F \subseteq Q$)
- δ ise iki durum arasındaki geçişleri tanımlar (Roche ve Schabes, 1995: 236-237).

Sonlu durum dönüştürücüleri, sonlu durum algoritmalarının yaptığı bu işlemi otomatik olarak gerçekleştiren bilgisayar programlarıdır. Dönüştürücüler mevcut durumlara dayanarak sonuç dizesi (string) üretirler (Goldsmith, 1993). Bir kelime alırlar ve bu kelimenin mümkün olan bütün köklerini ve biçimlerini kontrol ederler. Eğer girdinin tüm kuralları kabul edilirse, girdi kabul edilir. Değilse girdinin doğruluğu kabul edilmez ya da kısmen doğru kabul edilir (Altıntaş, 2001).

Sonlu durum tekniği genellikle doğal dil işleme amacı ile kullanılan basit ve etkili bir model olarak tanımlanmaktadır. Bu teknik aracılığı ile sembolleştirmeler, sesbilimsel, biçimbilimsel ve sözdizimsel analizler ile dil modellemeleri kolaylıkla yapılabilmektedir (Roche ve Schabes, 1997). Ancak bu teknik zaman içinde sadece dilbilim alanında kullanılmakla kalmayıp bilimin çeşitli alanlarında kullanılmaya başlanmıştır.

2.2.1. Sonlu Durum Tekniği Kullanılan Çalışmalar

Literatürde sonlu durum tekniğinin mühendislik, dilbilim, tıp ve kütüphanecilik gibi farklı alanlarda kullanıldığı görülmüştür. Bu denli yaygın olarak kullanılmasının nedeni sonlu durum tekniğinin her alana uyarlanabilecek nitelikte olmasıdır.

Sonlu durum tekniği kullanılarak üretilen bilimsel çalışmaların büyük çoğunluğunun dillerin morfolojik yapılarının ortaya koyulması amacı ile yapıldığı görülmüştür. Morfolojik analizler analitik ve senteze yönelik olan yaklaşımlar ile yapılmaktadır. Daha çok bilinmeyen diller üzerinde uygulanan ve sözcükleri tamamen parçalamanın esas alındığı analitik yaklaşımın tersine senteze yönelik yaklaşımlar dilbilimsel bir teoriye dayanmakta ve bir dilin kurallarını esas almaktadır (Aronoff ve Fudeman, 2005: 13). Morfolojik araştırmaların gelişmesi ile dillerin bilgisayarlı hesaplamaları elzem hale gelmiş ve hesaplamalı morfoloji (computational morphology) kavramı ortaya çıkmıştır. Hesaplamalı morfoloji bir dilin yapısını anlamak üzere bir bilgisayar sistemine doğal bir dili oluşturma ve değerlendirme yeteneklerini kazandırmak olarak tanımlanmıştır. Böylece

bilgisayarların çeviri gibi dil ile ilgili karmaşık görevleri ya da gazeteden, dergiden gelen bilgileri işleme gibi konularda etkin kullanılması sağlanmıştır (Grishman, 1986: 1).

Sonlu durum algoritmalarının doğal dilleri, özellikle de İngilizceyi tanımlamada kullanılmasının imkânsız olduğu öne sürülmesine karşın (örneğin, Chomsky, 1964: 21), sonlu durum tekniğinin doğal dil işlemede sıklıkla kullanıldığı görülmüştür. Konuya olan ilginin giderek artmasının temel nedeni sonlu durum tekniğinin sunduğu hız ve yoğunluktur. Bu teknik üzerine çalışan pek çok araştırmacı doğal dil işleme amacı ile sonlu durumun kolaylıkla kullanılabileceğini belirtmişlerdir (Johnson, 1972; Kaplan ve Kay, 1994; Roche ve Schabes, 1995; Oflazer, 1996; Mohri, 1997).

Dillerin morfolojik yapılarını ortaya koymak amacı ile uluslararası alanda çok miktarda yayın yapılmıştır. Üzerinde morfolojik araştırma yapılan bu dillerden bazıları Arapça (Beesley, 1996), İngilizce (Karttunen ve Wittenburg, 1983) ve İrlandacadır (Ui Dhonnchadha, 2003). Türkçe'nin morfolojik yapısını anlamak üzere de pek çok kitap, proje, tez ve makale literatürde yer bulmuştur (Göksel ve Kerslake, 2005; Oflazer, 2002; Mersinli, 2010; Oflazer, 1994).

Dillerin morfolojik yapılarının sonlu durum tekniği aracılığı ile saptanmasının yanında bağ grameri (link grammar) çalışmaları da yapılmıştır. Sleator ve Temperley (1991) tarafından geliştirilen bağ gramerinde temel amaç herhangi bir dilin yapısına uygun bir algoritma geliştirerek sözcük temelli gramer yapısını ortaya koymaktır. Bu sistemde bir dilin sözdizim kuralları o dilde yer alan sözcükler ve bu sözcükler arasındaki ilişkiler belirtilerek tanımlanır. Ardından sisteme bir cümle verilir. Eğer bu cümle tanımlanan yapılara uyum sağlıyor ise sistem tarafından kabul edilir. Böylelikle soru cümleleri, edatlar, sıfat cümleleri gibi İngilizce'de sıklıkla kullanılan temel cümle yapıları ile oluşturulmuş

cümlelerin yapılarının belirlenmesi ve öğelerine ayrılması olanaklı hale getirilmiştir.

Sondan eklemeli diller için bir cümlenin içindeki sözcüklerin birbirleri ile sözdizimsel ilişkisinin saptandığı bir çalışmada yine sondan eklemeli bir dil olan Türkçe için bağ grameri geliştirilmiştir (İstek ve Çiçekli, 2007; İstek, 2006). Çalışma kapsamında Türkçe kelime grupları, tamlamalar, bağımlı cümlecikler, isim fiiller, isim ve fiil cümlecikleri, pozitif ve negatif düzgün cümleler, zamir düşmeleri, kısaltmalar ve sayılar gibi tüm yapılar tanımlanmıştır. Ancak devrik cümleler, deyimsel ifadeler ve noktalama işaretleri çalışma kapsamına alınmamıştır. Bu bağ grameri aracılığı ile Türkçe'nin morfolojik yapısı tamamen ortaya koyulmuştur. Sözcüksel bir değerlendirmenin yanı sıra cümle yapılarının saptanmasına da çalışılmıştır.

Yine dil ile ilgili çalışmaları geliştirmek amacı ile sonlu durum tekniği kullanılarak bir dili diğer bir dile dönüştüren çeviri programları tasarlanmıştır. Bu programlar yapıları birbirine yakın iki dilin birbirine çevrilmesini olanaklı kıldığı gibi bazı programlarla yapıları birbirinden tamamen farklı iki dilin birbirine çevrilmesi sağlanmıştır. Kırım Tatarcasının morfolojik analizinin yapıldığı bir çalışmada XEROX sonlu durum dönüştürücüsü kullanılarak yaratılan iki aşamalı bir morfoloji tasarlanmıştır (Altıntaş ve Çiçekli, 2001). Böylece Kazan Tatarcası, Kırgızca ve Kazakça gibi Kıpçak Türkçesi kökenine dayanan Türk dilleri için bir model tanımlanarak bu diller arasında çeviri yapmayı kolaylaştırmak hedeflenmiştir. Bunun yanında bazı farklılıklarına rağmen Kırım Tatarcası ile Türkçe arasında büyük benzerlikler olduğu sonucuna da ulaşılmıştır. Bu çalışma ile bağlantılı bir diğer çalışmada ise Türkçe ve Kırım Tatarcası arasında çeviri yapabilen bir sistem yaratılmıştır (Çiçekli, 2005). Bu sistemde sözcükler öncelikle köklerine ayrılmış ve kökler diğer dilin kökleri ile, sözcüklerin aldıkları ekler de diğer dildeki uygun ekler ile değiştirilmiştir. Örneğin, "akşam eve geleceğiz" şeklinde bir cümle bu çeviri sistemine girildiğinde cümle otomatik olarak "aqSam evge kelecekmiz" cümlesine çevrilmiştir. Literatürde çeviri programları çok miktarda

bulunmaktadır. İngilizceyi ve Uygurcayı Türkçeye çevirebilen programlar bu gibi çalışmalara örnek olarak gösterilebilir (Özbek ve Jonathan, 2006; Orhun, Adalı ve Tantuğ, 2008).

Sonlu durum tekniği yazılı metinlerde yapılmış hataları bulma ve düzeltme amacı ile de kullanılmıştır. Türkçe için sonlu durum tekniği kullanılarak tasarlanan ve temel amacı metin içindeki hataları bulmak olan bir çalışmada hatalı yazılmış sözcüklerin saptanmasına ve düzeltilmesine çalışılmıştır (Oflazer, 1996). Bu çalışma ile hatalı yazılmış "uygarlaştıramadıklarmıızdanmışsınızcasına" sözcüğü bile sisteme girildiğinde doğru biçimine ("uygarlaştıramadıklarımızdanmışsınızcasına") dönüştürülebilmiştir.

Literatürde sonlu durum tekniğinin büyük yoğunluklu metinlerin içinden bilgi aktarmak amacı ile de kullanıldığı görülmüştür. Bu gibi çalışmaların temel amacı doğal dil ile yazılmış metinlerin içinden istenilen bilgilerin sağlanmasıdır. Metin aktarımı amacı ile geliştirilmiş bir program olan FASTUS aracılığı ile sözcükler, kelime grupları, fiiller gibi dil ile ilgili öğeler tanımlanmış ve bu sayede metin içinden istenilen bilgilerin saptanarak aktarılması olanaklı hale getirilmiştir (Hobbs, Douglas, Bear, Israel, Kameyama, Stickel ve Tyson, 1996).

Sonlu durum tekniği kullanılarak okuma fonksiyonlu programların yaratıldığı görülmektedir. Bu programlar aracılığı ile yazılı metinler öncelikle sistem tarafından tanımlanmış, ardından da bu metinler sese dönüştürülmüştür. Bu gibi çalışmalar çocuklara hikaye okuma (Duchateau, Wigham, Demuynck ve Van Hamme, 2006) veya hasta bilgilerini okuma (Adalı ve Erenler, 2003) gibi amaçlarla tasarlanmıştır.

Sonlu durum tekniğinin dilbilim ve mühendislik alanlarının yanında tıp ve genetik alanlarında da kullanıldığı görülmüştür. Literatürde tıbbi dokümanların standardizasyonu amacı ile sonlu durum tekniği kullanılmıştır. Radyoloji raporlarının standart olmaması ve istenilen bilgilere erişimdeki güçlükler nedeni ile bu raporların standartlaştırılmasına

çalışılmış ve prototip bir Türkçe bilgi çıkarım sistemi geliştirilmiştir (Soysal, Çiçekli ve Baykal, 2009). Bu sistem ile doğal dilde "Karaciğer vertikal yüksekliği 14 cm'dir" şeklinde yazılmış bir rapor "Karaciğer.yükseklik = 14cm"ye dönüştürülmüştür. Bu çalışma aracılığı ile yapılandırılmış radyoloji raporları yaratılmış ve bu sayede tıbbın diğer alanları için de örnek oluşturulmuştur.

Genetik alanında ise gen adlarının sınıflandırılabilmesi amacı ile sonlu durum tekniği kullanılmıştır (Galvez ve Moya-Anegón, 2012). Çalışmada MEDLINE veri tabanından çekilen 2.572 kayıt işleme alınmış ve metin içeriklerinde yer alan genlere ait semboller değiştirilerek tekil birer sembole dönüştürülmüştür. Bu sayede gen adlarının standardize edilmesi sağlanmış ve metin içinden erişimi kolaylaştırılmıştır.

2.2.1.1. Kütüphanecilik ve Bilgibilim Alanında Sonlu Durum Tekniği Kullanılan Çalışmalar

Diğer alanlarda olduğu gibi sıklıkla rastlanmıyor olsa da literatürde kütüphanecilik ve bilgibilim alanında sonlu durum tekniği kullanılan çalışmalar bulunmaktadır. Bu çalışmaların bir kısmı bilgi erişim alanında tasarlanırken, bir kısmı da adres ve yazar adlarının standardize edilmesi amacı ile yapılmıştır. Bahsi geçen çalışmaların belirli bir yazar kitlesi tarafından hazırlandığı ve yazar çeşitliliğinin çok fazla olmadığı görülmüştür.

Sonlu durum tekniğinin bilgi erişim alanında kullanımı sözcüklerin ek-kök yapısı üzerinden yürütülmüştür. Bu çalışmalarda güdülen temel amaç sonlu durum tekniği kullanılarak bilgi erişimi daha etkili hale getirmek ya da mevcut durumu resmederek olumlu ve olumsuz yanları ortaya çıkarmaktır.

Galvez ve Moya-Anegón (2006b) bilgi erişim temelinde sonlu durum tekniğine dayanan ve anlamsal olarak birbirine benzeyen değişkenleri eşleştirmek ve gruplamak amacı ile kullanılan veri tümleştirme (conflation) tekniğinin doğruluğunu ölçmeye çalışmışlardır. Bu çalışmada tahmin

edilenin aksine veri tümleştirmenin ardından yapılan sorguların kalitesi değerlendirilmemiş, doğru olarak tümleştirilmiş sözcükler dikkate alınarak anma-duyarlık testleri yapılmıştır. Bu amaçla sonlu durum tekniği kullanılarak yaratılmış İspanyolca sözlüklerde yer alan sözcükler değerlendirilmiştir. Çalışma sonunda çeşitli biçimlerin analizinde bazı kısıtlamalar yaşansa da, bu teknik ile üretilen ve tümleştirilen sözcüklerin doğruluk oranınının yüksek olduğu sonucuna ulaşılmıştır. Bir başka çalışmada ise veri tümleştirme işinde dilbilimsel ve dilbilimsel olmayan yaklaşımlar karşılaştırılmış ve dilbilimsel yaklaşım çerçevesinde normalizasyon uygulamaları incelenmiştir (Galvez, Moya-Anegón ve Solana, 2005). Çalışmada terim çeşitliliğini standardize etmede sonlu durum tekniğinin önemli rol üstlenebileceği vurgulanmıştır.

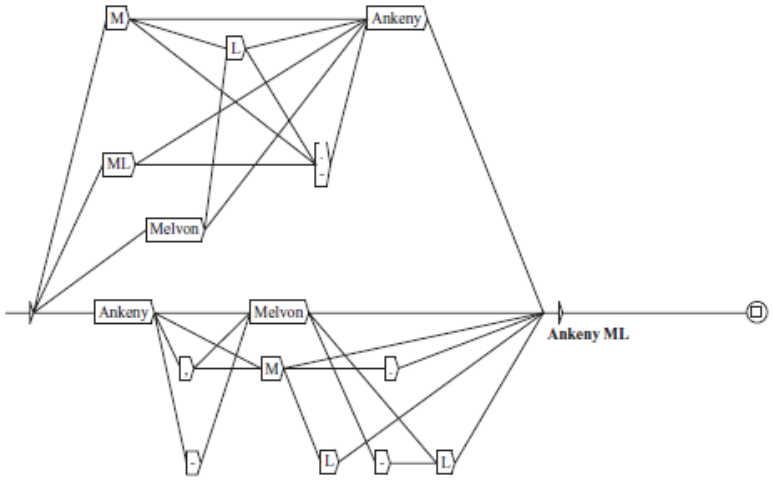
Benzer bir çalışmada Kettunen (2008) bilgi erişimde anahtar sözcüklerin dilbilimsel çekimlerinin getirdiği avantaj ve dezavantajları tartışmıştır. Çalışma kapsamında Türkçenin de dâhil olduğu 11 farklı dilde yazılmış metinler üzerinden bilgi erişim için kullanılan sözcüksel teknikler değerlendirilmiştir. Türkçe için en yararlı tekniğin gövdelemeye dayanan kök çözümleyici (lemmatizer) kullanmak olduğu ve kök çözümleyiciliğinde 50.000 ile 100.000 gibi büyük yoğunluklu dilbilimsel kurallarda en etkili yöntem olarak sonlu durum tekniğinin kullanılabileceği sonucuna ulaşılmıştır. Yine bu çalışmada kullanım kolaylığı, alan kazanımı ve erişim performansını artırması nedeni ile bilgi erişimde dilbilimsel yöntemlerin kullanımının önemi vurgulanmıştır.

Yazar adı ve kurum adreslerinin standardize edilmesi amacı ile de sonlu durum tekniği kullanılmıştır. Galvez ve Moya-Anegón (2006a)'un yaptıkları çalışmada, *SCI*'de dizinlenen ve adres alanında "University of Granada" ve "department" geçen sözcükler saptanmış ve analiz için kullanılmıştır. Bu çalışmada temel amaç üniversite adresinde yapılan heceleme hatalarını saptamak değil, adres biçiminin standart hale getirilmesini sağlamaktır. Standardizasyon işleminin ardından üniversite adresinin görünümünün *üniversite adı, fakülte/enstitü/araştırma grubu, bölüm, şehir, ülke* şeklinde

dizilimini sağlamak hedeflenmiştir. Çalışmada Nooj programının bir önceki sürümü olan INTEX sonlu durum dönüştürücüsü kullanılarak 719 kayıttan 631'inin standardize edildiği sonucuna ulaşılmıştır. Standardizasyonun doğruluğunu saptamak amacı ile de standardize edilmiş kayıtlar ile tüm kayıtların sayısı arasında anma ve duyarlık testleri gerçekleştirilmiştir.

Yapılan bir diğer çalışmada ise (Galvez ve Moya-Anegón, 2007a), *INSPEC, MEDLINE* ve *CAB Abstracts* veri tabanlarından "dept" ve "univ" sözcükleri kullanılarak yapılan taramalar ile elde edilen 3.916 kayıt dizilim biçimleri açısından değerlendirilmiş ve sonuçlar sunulmuştur. En etkili biçim standardizasyonu *INSPEC* veri tabanı içinde yer alan kayıtlar üzerinde gerçekleştirilmiştir.

Galvez ve Moya-Anegón tarafından hazırlanan ancak bu kez yazar adlarının standardizasyonu üzerinde durulan bir çalışmada (Galvez ve Moya-Anegón, 2007b) *LISA* ve *SCI* kullanılarak örnek bir yazara ait kayıtlar elde edilmiş ve sonlu durum tekniği uygulanmıştır. İlk aşamada örnek yazar olarak alınan "Melvon Ankeny" isminin farklı kullanımları saptanmış ve n-gram tekniği kullanılarak ağırlıklandırılmıştır. Ardından olası kullanımlar belirlenerek sonlu durum grafiği yaratılmıştır (bkz. Şekil 2). Çalışmanın ilerleyen kısımlarında ise yazar isimlerindeki hataların saptanması için genel bir formül tasarlanmış ve bu formülün adı girilen herhangi bir İspanyol yazar için de çalışması sağlanmıştır. Bu işlem *LISA* ve *SCI*'de dizinlenmiş 52 yazar ismi üzerinden yürütülmüş ve etkili sonuçlar alınmıştır. *LISA*'da yer alan 580 kaydın 574'ü tanımlanmış 570'i standardize edilmiştir. *SCI*'de yer alan 234 kayıt için de 220'si tanımlanmış ve 217'si standardize edilebilmiştir.



Şekil 2: Melvon Ankeny isminin farklı kullanımlarını saptamak üzere hazırlanmış bir sonlu durum grafiği (Galvez ve Moya-Anegón, 2007b)

Şekil 2'de yer alan grafik aracılığı ile yaratılan patika takip edilerek "Ankeny Melvon", "Melvon Ankeny", "ML, Ankeny" gibi farklı kullanımlara erişmek olanaklı hale getirilmiştir. Benzer bir grafik çalışmamız için de tasarlanmıştır. Bu bağlamda yürütülen tez çalışmasına en yakın yayının bu çalışma olduğunu söylemek yanlış olmayacatır.

Uluslararası literatürde sonlu durum tekniğinin adres standardizasyonuna uygulandığı çalışmalara bakıldığında konusal eğilimin adres biçimlerindeki standardizasyon sorununu saptamak ve çözüm getirmek olduğu ve harf/heceleme hatalarına ağırlık verilmediği görülmüştür. Ulusal kütüphanecilik ve bilgibilim literatüründe ise sonlu durum tekniği kullanılarak yapılmış çalışmalara rastlanmamıştır. Atıf dizinlerinde Türk üniversite adreslerinde yapılmış hataların sonlu durum tekniği kullanılarak tespit edilmesini amaçlayan bu çalışma ile harf ve heceleme hatalarının saptanması olanaklı hale gelecek ve daha önemlisi Türkiye kütüphanecilik ve bilgibilim literatürüne katkı olarak daha önce kullanılmamış bir tekniğin kullanılabilirliğini ortaya koyulacaktır.

2.2.1.2. Kütüphanecilik ve Bilgibilim Alanında Adres ve Yazar Adı Standardizasyonu Üzerine Yapılan Diğer Çalışmalar

Kütüphanecilik ve bilgibilim literatüründe adres ve yazar adı standardizasyonuna yönelik çalışmalar sonlu durum tekniği ile sınırlı kalmamıştır. Pek çok çalışma farklı yöntemler kullanarak bu gibi standardizasyon sorunlarını ortaya çıkarmaya ya da çözmeye odaklanmıştır.

Moed (2005) yaptığı çalışmada atıf dizinlerinde yazar ve kurum adları ile alt alan isimlerinde hatalar yapıldığını belirtmiştir. Bu çalışmada yazar adlarında var olan tutarsızlığın bir yazarın isminin birden farklı şekillerde kullanılması ve aynı adlı pek çok yazarın bilimsel çalışmalar yapması nedeni ile oluştuğu belirtilmiştir. Bunun yanında yazar isimlerinde çeviri yapıldığında (Çince ya da Korece gibi dillerden), yazarların soyisimleri değiştiğinde ya da takma isim kullandıklarında da yazar isimlerinde tutarsızlıklar oluştuğu vurgulanmıştır. Kurum adlarındaki tutarsızlığın temel nedeni yazarların diledikleri adres bilgisine çalışmalarında yer verebiliyor olmalarıdır. Şöyle ki, aynı kurumda çalışan iki kişi birbirinden farklı adres tanımlaması yapabilmekte ve bu durum çalışma yayımlandıktan sonra bile tesadüfen fark edilebilmektedir.

De Bruin ve Moed'in yaptığı bir çalışmada (1990) üniversite adreslerinin standart olmaması sonucu bilgi erişimde problemler yaşandığı ve bu problemlerin çözümü için doğru bir adres birleştiriminin yapılması gerekliliği vurgulanmıştır. Bu amaçla üniversite adresleri değerlendirilmiş, bulundukları ülkeler ve şehirler dikkate alınarak sınıflandırılmış, ardından da kurum ismine ya da kurum isminin kısaltmasına göre birleştirim işlemi yapılmıştır.

Hood ve Wilson (2003) yaptıkları araştırmada enformetrik çalışmalarda veri tabanı kullanımının sınırlılıklarını incelemişler ve elektronik veri tabanlarında var olan makro ve mikro düzeylerdeki tutarlılık sorunları ve sınırlılıklarını ortaya koymuşlardır. Çalışmada elektronik veri tabanlarında genellikle yazar ve dergi isimleri, tarihler ve kurum isimlerinde hata

yapıldığı sonucuna ulaşılmıştır. Kurum isimlerinde yapılmış hataların en temel nedeninin kurumların bölünmesi ya da birleşmesi sonucu isimlerinde değişiklik yapılması olduğu öne sürülmüştür. Kurum isimlerinin sabit olmaması ve sürekli değiştirilmesi ile ilgili Türkiye'de de sıkıntı yaşanmaktadır. Örneğin, Bilkent Üniversitesi olarak anılan kurumun adı 2010 yılında alınan bir karar ile İhsan Doğramacı Bilkent Üniversitesi adını almıştır. Yine en son alınan kararla 1992 yılından beri eğitim öğretimi sürdüren Zonguldak Karaelmas Üniversitesinin ismi Bülent Ecevit Üniversitesi olarak değiştirilmiştir (İlköğretim, 2012). İsim değişiklikleri yalnızca üniversite isimlerinin değiştirilmesi ile sınırlı kalmamıştır. Pek çok devlet kurumunun isminin yıllar içinde sürekli olarak değiştirildiği görülmektedir. İsmi sürekli değiştirilen kurumlardan biri olan T.C. Kültür ve Turizm Bakanlığının ismi kuruluşunda "Basın, Yayın ve Turizm Bakanlığı" iken "Turizm ve Tanıtma Bakanlığı" olarak değiştirilmiştir. Ardından iki ayrı bakanlığa bölünerek "Kültür Bakanlığı" ve "Turizm Bakanlığı" kurulmuştur. 2003 yılına gelindiğinde ise bu iki bakanlık birleştirilerek "Kültür ve Turizm Bakanlığı" ismini almıştır. Son düzenlemeler ile bakanlığın isminin "Turizm ve Tanıtma Bakanlığı" olarak yeniden yapılandırılması gündeme gelmiştir (Milliyet, 2012).

Stefaniak (1987) atıf dizinlerinde yayın yapan yazarların tamamının İngilizce konuşan insanlar olmaması nedeni ile kurum isimleri İngilizceye çevrilirken hata yapıldığını aktarmıştır.

Van Raan (2005) bibliyometrik araştırmalar ile üniversite sıralaması yapılırken karşılaşılan en önemli problemlerden birinin kurum adreslerindeki tutarsızlık olduğunu savunmuştur. Pek çok kişi üniversitelerin tüm yayınlarının atıf dizinlerinde "University X" şeklinde bulunduğunu düşünürken aslında bu varsayımın tamamen yanlış olduğunu öne sürmüştür. Bu nedenle üniversite sıralamaları yapılmadan önce birleştirim işleminin gerçekleştirilmesini gerekli görmüştür. Üniversite sıralamaları üzerine hazırlanan bir kitapta da kurumların ve yazarların yayın sayılarının ölçülmesinde adreslerde ve isimlerde tutarlılık olması halinde

etkili sonuçlara ulaşılabileceği belirtilmiştir (Toutkoushian ve Webber, 2011: 130).

Elektronik veri tabanlarında üniversite adreslerinin ya da yazar isimlerinin standardizasyonu üzerine yapılmış çalışmalarda sonlu durum tekniğinin yanında kümeleme (clustering) tekniğinin de kullanıldığı görülmüştür. Bu amaçla yaklaşık dizi eşleme algoritmaları (approximate string matching) kullanılmıştır. Literatürde yer alan bir çalışmada kurum isimleri kümeleme tekniği ile standardize edilerek bir yetke dizini oluşturulmuştur (French, Powell ve Schulman, 2000). Bu çalışmada kurum adreslerinde hatalara neden olan farklı yazımlar, heceleme hataları ve çeviri farklılıklarını saptayabilecek algoritmalar aracılığı ile nasıl yetke dizinleri oluşturulabileceği hakkında bilgi verilmiştir.

Uluslararası literatürde adres veya yazar adı standardizasyonu üzerine yapılmış başka çalışmalar yer alsa da (Cornell, 1982; Piternick, 1982, Williams ve Lannom, 1981; Ruiz-Pérez, López-Cózar ve Jiménez-Contreras, 2002), Türkiye'de bu konu üzerinde fazla çalışılmamıştır. Var olan tek birleştirim çalışmasının ULAKBİM (Ulusal Akademik Ağ ve Bilgi Merkezi) tarafından 2007 yılında çıkarılan *Türkiye Bilimsel Yayın Göstergeleri* adlı kitapta yapıldığını söylemek yanlış olmayacaktır (ULAKBİM, 2007). Bu kitapta Türkiye'de eğitim vermekte olan her bir üniversite için atıf dizinlerinde kullanılan adresler saptanmış ve *WoS* tarama formülleri olarak okuyucuya sunulmuştur. Kitapta belirlenen adresler ile bu çalışma kapsamında elde edilen adres girişleri karşılaştırılarak birleştirimin başarısı ortaya konulmuştur.

3. BÖLÜM: YÖNTEM VE TEKNİK

3.1. GİRİŞ

Bu araştırmada atıf dizinlerinde Türk üniversite adreslerinde yapılan hataları saptamak, bu hataların üniversite performans değerlendirmelerinde ortaya çıkardığı olumsuz etkileri belirlemek ve bu konuda çözüm önerileri sunmak amaçlanmaktadır. Bahsi geçen amaç doğrultusunda bu bölümde verilerin toplanması ve analize uygun hale getirilmesinden sonlu durum tekniğinin uygulanmasına kadar çalışmanın yöntemi ve tekniği konusunda bilgiler verilmektedir. Sonlu durum tekniğinin uygulanacağı dönüştürücü Nooj hakkında temel bilgiler de yine bu kısımda sunulmaktadır.

3.2. VERİLERİN TOPLANMASI VE ANALİZE UYGUN HALE GETİRİLMESİ

Araştırmada atıf dizinlerinde 2009 yılına kadar dizinlenmiş Türkiye adresli toplam 198.687 yayın indirilmiş ve incelenmiştir. Bu yayınlar üzerinde herhangi bir yayın türü (makale, bildiri özeti, kitap tanıtımı gibi) kısıtlaması yapılmamıştır.

Atıf dizinlerinde bazı yayınlarda adres bilgisi yer almamaktadır. *Middle Eastern Studies* ve *Athenaeum-Studi Periodici Di Letteratura E Storia Dell Antichitabazi* gibi dergiler adres bilgisi taşımayan yayınlara örnek olarak gösterilebilir. Bazı yayınlarda ise ülke bilgisinde hatalar yapılmıştır. Bu gibi durumlar nedeni ile adresinde Türkiye içermeyen ancak Türkiye'ye ait olan bazı yayınlar çalışma kapsamına alınamamıştır. Çalışmanın en önemli sınırlılığının bu durum olduğunu söylemek yanlış olmayacaktır.

İndirilen dosyalar içinden adreslerin belirlenebilmesi amacıyla *WoS*'un standart veri dosyasında yer alan ve yazar adreslerini içeren C1 (author address) alanı ile iletişim kurulacak yazarın adresini içeren RP (reprint address) alanında (Web of Science, 2009) yer alan adresler veri dosyasının

temelini oluşturmuştur. Yayınların tanımlanabilmesi amacı ile her kayda tekil birer numara verilmiştir.

Adres birleştirimi *WoS* adres alanlarında yer alan tüm adreslerin tek tek incelenmesinin ardından uygun adresler *institution* adlı başka bir sütuna yazılarak gerçekleştirilmiştir. Örneğin, C1 alanında "Istanbul Univ, Res & Applicat Ctr Biotechnol & Genet Engn, TR-34118 Istanbul, Turkey; Gaziosmanpasa Univ, Dept Med Biol, Fac Med, Tokat, Turkey" şeklinde yer alan bir kayıt "institution" alanına "ISTANBUL;GAZIOSMANPASA" olarak girilerek her bir kaydın hangi üniversiteye ya da üniversitelere ait olduğu saptanmıştır. Burada fakülte ya da bölüm bazlı bir değerlendirme yapılmaksızın sadece üniversite adları değerlendirmeye alınmıştır.

Bu aşamada C1 veya RP alanında birim adı verilmiş kayıtlar, o birimin bağlı olduğu kurumların ismi ile birleştirilmiştir. Örneğin, bir kaydın adresi "Marmara Research Center" ise bu kayıt "institution" alanına "TÜBİTAK" şeklinde girilmiştir. Benzer şekilde yüksekokul ya da fakülte ismi içeren kayıtlar da bağlı oldukları üniversite ismi ile tanımlanmıştır. Örneğin, "Çapa Tıp Fak" ya da "Cerrahpasa Sch Med" şeklinde verilmiş kayıtlar İstanbul Üniversitesi kayıtlarına dâhil edilmiştir.

Bazı adres girişlerinde kurum adının belirtilmemiş olması birtakım kayıtların çalışma kapsamından çıkarılmasına neden olmuştur. Örneğin, yazarın "dept plast & reconstruct surg, ankara, turkey" şeklinde adres belirttiği bir kaydın Ankara'da hangi üniversiteye ait olduğunu bulmak amacı ile yazar alanı incelenmiştir. Eğer yazar bilgisinden de kurum adına ulaşılamadıysa bu gibi kayıtlar çalışma kapsamından çıkarılmıştır. Benzer biçimde ev, iş yeri adresi gibi mesken adresinin adres alanında yer aldığı kayıtlar da değerlendirme kapsamına alınmamıştır. Ancak bu gibi durumlar az sayıda kayıt (647 kayıt / tüm yayınların %0,3'ü) için geçerli olduğundan çalışma için büyük bir problem oluşturmamaktadır.

Bazı kurum adreslerinde ise dizinleme hataları saptanmıştır (Univ Nigeria: Niğde Üniversitesi; Univ Centennial: Yüzüncü Yıl Üniversitesi gibi). Bu

gibi durumlarda kaydın aslında hangi üniversiteye ait olduğuna bakılarak veriler temizlenmiş ve düzeltilmiştir. Ayrıca adresinde Türkiye geçen ancak aslında Türkiye'ye ait olmayan kayıtlar da saptanarak çalışma kapsamından çıkarılmıştır. Bu gibi dizinleme hataları ile ilgili detaylı bilgi çalışmanın bulgular kısmında sunulmaktadır.

3.3. STANDARDİZASYON SORUNUNUN SAPTANMASI VE BÜYÜKLÜĞÜNÜN ÖLÇÜMÜ

Verilerin tamamının standart bir biçime dönüştürülmesinin ardından Türkiye adresli yayınların kurumlara dağılımı saptanmıştır. Bu işlemin ardından standardizasyon sorunun büyüklüğünün saptanması ve ölçümü için 4.000'den fazla yayın yapan ilk 20 üniversite seçilmiştir.

Örnekleme alınan bu 20 kurumun yayın sayılarından hareketle bu yayınların yüzde kaçında adres hatası yapıldığı ortaya koyulmuştur. Bu işlem uygulanırken üniversiteler için doğru biçimin "X Univ" ve "Univ X" olduğu varsayılmıştır ("Hacettepe Univ", "Univ Hacettepe" gibi). Bu durumda Hacettepe Üniversitesi yayınlarını tarayan bir kişi "Hacettepe" terimini sorgu için kullanabileceğini düşünebilir. Ancak Türkiye adresli yayınlar incelendiğinde, benzer isme sahip olan sokak veya cadde isimlerine de rastlanmıştır. Bu nedenle de bir kurumun doğru bir adres kaydında "Univ" kısaltmasının bulunması zorunludur. Örnek vermek gerekirse Dumlupınar Üniversitesi yayınlarını "Dumlupinar" sorgusu ile bulmaya çalışan bir kişinin eriştiği sonuçlar içinde "Dumlupınar Sokak", "Dumlupınar Mahallesi" gibi adresler de yer alacaktır. Doğal olarak bu şekilde tarama yapılması doğru sonuçlar sağlamayacaktır. Diğer taraftan üniversiteler için doğru adres girişleri belirlenirken eğer kurum isminin İngilizce bir karşılığı var ise (ODTU, Middle East Technical University gibi) sorguya eklenmiştir. Her bir kurum için doğru kabul edilen adresler Ek 1'de detaylandırılmıştır.

Standardizasyon sorununun büyüklüğünün gösterilmesi amacı ile de bibliyometrik haritalamadan yararlanılmıştır. Bu haritalar aracılığıyla standardize edilmiş ve edilmemiş adresler kullanılarak yaratılmış haritalar

arasındaki farkın gösterilmesi amaçlanmıştır. Standardize edilmiş adresler kullanılarak yaratılmış işbirliği haritalarının daha doğru ve etkili sonuçlar ortaya çıkaracağı öngörülmektedir.

3.4. SONLU DURUM DÖNÜŞTÜRÜCÜSÜNÜN UYGULANMASI

Bu çalışmada sonlu durum dönüştürücüsü kullanılarak büyük yoğunluklu adres bilgisi içeren bir metin dosyası içinden bir kurum için kullanılmış tüm adres girişlerine ulaşılması hedeflenmektedir. Bu amaçla en yüksek yayın sayısına sahip Hacettepe Üniversitesi üzerinden işlem gerçekleştirilmiştir.

Çalışma kapsamında sonlu durum dönüştürücüsü olarak 2002 yılında Max Silbertzein tarafından geliştirilen Nooj doğal dil işleme aracı kullanılmıştır. Nooj dönüştürücüsü geniş kapsamlı sözlükleri ve gramer yapılarını çalıştırabilen açık kaynak kodlu bir yazılımdır ve Türkçe de dâhil olmak üzere pek çok dili kapsar (Nooj, 2012). Literatür kısmında da değinildiği üzere daha önce yapılan az sayıdaki adres standardizasyon çalışması Nooj'un bir önceki versiyonu olan INTEX programı kullanılarak geliştirilmiştir. Ancak zaman içinde doğal dil işleme açısından INTEX programı kullanışlılığını yitirdiğinden daha üst düzey bir sistem olan Nooj dönüştürücüsü tasarlanmıştır (Silbertzein, 2002). Çalışma kapsamında Nooj programı ile ilgili bilgi almak ve oluşan sıkıntıları çözmek amacı ile Nooj-Info e-posta grubuna[2] üye olunmuş ve yazılımın yaratıcısı Max Silbertzein ile sık sık iletişime geçilmiştir.

Nooj programının çalıştırılması için öncelikle .txt dosyalarının Nooj'un metin biçimi olan .not biçimine dönüştürülmesi gerekmiştir (Silbertzein, 2003). Ardından bu metin dosyaları üzerinden işlem yapabilmek amacı ile gramer grafikleri yaratılmıştır. Yaratılan bu gramer grafikleri metin dosyası üzerinde uygulandığında gramerde belirtilen kurallara uyan sonuçlar listelenebilmiştir.

Nooj programının nasıl çalıştığı ile ilgili detaylı bilgi bulgular kısmında uygulamalı bir şekilde anlatılmaktadır.

---

[2] http://tech.groups.yahoo.com/group/nooj-info/

4. BÖLÜM: BULGULAR VE DEĞERLENDİRME

4.1. GİRİŞ

Çalışmanın bu kısmında birinci bölümde tanımlanan araştırma sorularını cevaplamak ve hipotezleri test etmek amacı ile atıf dizinlerinde yer alan Türkiye adresli yayınlarda yapılan hataların saptanmasına, söz konusu hataların etkilerinin belirlenmesine ve sonlu durum tekniği kullanılarak bu duruma bir çözüm bulunmasına çalışılmıştır.

Çalışma kapsamında atıf dizinlerinde yayımlanmış toplam 198.687 yayın incelenmiştir. İnceleme sonucunda en çok yayını olan ilk 20 üniversitenin hatalı adresten dolayı yayın kaybı Tablo 1'de gösterilmektedir.

Tablo 1: En çok yayını olan ilk 20 üniversitenin hatalı adresten yayın kaybı

Üniversite	Toplam yayın sayısı	Hatalı (N)	Hatalı (%)
Hacettepe Üniversitesi	19.166	340	1,8
İstanbul Üniversitesi	16.390	1.691	10,3
Ankara Üniversitesi	13.275	224	1,7
ODTÜ	11.201	102	0,9
Ege Üniversitesi	9.428	654	6,9
Gazi Üniversitesi	9.281	85	0,9
İstanbul Teknik Üniversitesi	8.613	215	2,5
Dokuz Eylül Üniversitesi	6.069	210	3,5
Atatürk Üniversitesi	5.816	46	0,8
GATA	5.300	639	12,1
Marmara Üniversitesi	5.136	136	2,6
Çukurova Üniversitesi	4.953	60	1,2
Erciyes Üniversitesi	4.795	48	1,0
Ondokuz Mayıs Üniversitesi	4.537	173	3,8
Başkent Üniversitesi	4.418	32	0,7
Selçuk Üniversitesi	4.380	101	2,3
Boğaziçi Üniversitesi	4.102	57	1,4
Fırat Üniversitesi	4.091	86	2,1
Karadeniz Teknik Üniversitesi	4.014	395	9,8
Bilkent Üniversitesi	3.950	11	0,3

Tablo 1'de de görüldüğü üzere üniversite adreslerinde yapılan hataların büyüklükleri üniversitelere göre farklılık göstermektedir. En çok yayın yapan Hacettepe Üniversitesine ait yayınlarda %1,8'lik bir hata oranı görülse de, ikinci sırada yer alan İstanbul Üniversitesi yayınlarının %10'undan fazlasında adres hatası yapıldığı saptanmıştır. Hacettepe Üniversitesi yayınlarında yapılan adres hatası oranının düşük olmasının başlıca nedeni olarak Hacettepe Üniversitesinin son yıllarda akademik çalışmalar yapan personeline bir adres listesi sunması ve yayınlarında bu adresleri kullanmalarını tavsiye etmesi gösterilebilir (Hacettepe Üniversitesi Kütüphaneleri, 2012).

İstanbul Üniversitesinin YÖK tarafından yıllık olarak açıklanan üniversite yayın sayısı sıralamalarında (Yükseköğretim Kurulu, 2010) her yıl birinci olarak gösterilmesine rağmen bu çalışma kapsamında yapılan detaylı incelemede ikinci sırada yer alıyor olması ve bunun yanında yayınlarının %10'undan fazlasında adres hatası bulunması İstanbul Üniversitesi yayınları üzerinde daha detaylı çalışma gereğini ortaya koymuştur. Zira atıf dizininde 2 Nisan 2012 tarihinde "Univ Istanbul" sözcüğü kullanılarak yapılan taramada 4.383 adet İstanbul Üniversitesine ait olmayan yayın bulunmuştur. Sonuçlar içinde İstanbul Teknik Üniversitesine ait 2.000'den fazla yayın bulunurken, Koç, Sabancı, Marmara ve Yıldız Teknik Üniversitesi gibi İstanbul'da yer alan üniversitelerin yayınları da bu listede yer almıştır. Bunun nedeni bu üniversitelerin adres kayıtlarında "Koc Univ, Istanbul" şeklinde kullanımların olmasıdır. Her ne kadar tırnak içinde tarama yapılıyor olsa da, atıf dizini arada virgül olmasına rağmen bu adres biçimini de getirmiştir. Bu durum hem İstanbul Üniversitesi Kütüphanesi hem de bu gibi sıralamaları yapan ve bu sıralamaları dikkate alarak politika geliştiren kişi veya kurumların daha dikkatli davranmaları gerektiği konusunda ipucu vermiştir.

Çalışmada Türkiye'de en çok yayın yapan ilk 20 kurum arasında en yüksek oranda hataya sahip kurumun GATA (Gülhane Askeri Tıp Akademisi) olduğu sonucuna ulaşılmıştır. Şimdiye kadar yapılmış akademik yayın sayısı

dikkate alınarak hazırlanan pek çok üniversite sıralamasında GATA genellikle alt basamaklarda yer almıştır (Yükseköğretim Kurulu, 2010; ULAKBİM, 2007: 168). Bunun nedeni bu gibi çalışmalarda GATA'ya ait yayınların pek çoğunun saptanamamış olmasıdır.

Hata oranı %10'a yakın olan Karadeniz Teknik Üniversitesi yayınlarının pek çoğunda Karadeniz Teknik Üniversitesi adının İngilizce yazımının kullanıldığı, bunun yanında üniversite adının kısaltmasının da adreslerde sıkça yer aldığı görülmüştür. Ancak bu üniversite ismi için "Black Sea Technical University" gibi İngilizce bir çeviri bulunmamakta üniversite isminin İngilizce karşılığı "Karadeniz Technical University" olarak kullanılmaktadır (http://obs.ktu.edu.tr/en/). Kısaltma kullanılarak verilen adresler ise sadece Karadeniz Teknik Üniversitesi yayınlarında değil, diğer tüm üniversitelerin yayınlarına erişimde sıkıntı yaşatmaktadır. Konu ile ilgili detaylı örnekler çalışmanın ilerleyen bölümlerinde sunulmuştur.

Tablo 1'deki listede yer almayan ancak hata oranının çok daha yüksek olduğu üniversiteler de bulunmuştur. Çalışma kapsamında yapılan değerlendirmede Kahramanmaraş Sütçü İmam Üniversitesi yayınlarının %25'inden fazlasının, Yüzüncü Yıl Üniversitesinin ise yayınlarının %15'inin hatalı adresten dolayı erişilemez olduğu görülmüştür.

Bu denli farklı hataların ortaya çıkması hataların yapılma nedenlerinin belirlenmesi gerekliliğini doğurmaktadır. Eğer hatalar üniversite isminin uzunluğu, kısalığı, Türkçe karakter içermesi gibi bir nedene bağlı ise bu neden ortaya çıkarıldığında ilgili kişilerin daha dikkatli davranmaları sağlanarak sorunun en aza indirilebileceği düşünülebilir. Bu bağlamda hatalı adreslerin hangi nedenlere bağlı olduğunu saptamak amacıyla yayın sayısı en yüksek olan ilk 40 üniversite için korelasyon testi uygulanmıştır. Ancak, korelasyon testi sonucunda Türk üniversite adreslerinde var olan adres hataları ile üniversite isminde yer alan karakter sayısı (r=0,064, p=0,699), sözcük sayısı (r=0,040, p=0,810) veya Türkçe karakter sayısı (r=0,066, p=0,692) arasında anlamlı bir ilişki olmadığı görülmüştür. Hatalı adres

sayısı ile ilişkili tek değişkenin yayın sayısı olduğu saptanmıştır (r=0,585, p<0,001). Ancak hatalı yayın sayısı ile toplam yayın sayısı arasında pozitif bir korelasyonun varlığının bulunması doğal ve kaçınılmazdır. Türk üniversite adreslerinde yapılmış hataların belirli bir mantık çerçevesinde açıklanamaması nedeni ile adreslerde var olan standardizasyon sorununu azaltmak hayli güç olacaktır.

Üniversitelere ait kayıtlardaki hata oranları dikkate alındığında yapılan veya ileride yapılacak olan her türlü yayın sayısı sıralamasında hatalar bulunabileceği ve bu nedenle de bu sıralamalara eleştirel yaklaşılması gereği ortaya çıkmıştır. Bu bağlamda yayın sayısı sıralamalarında otorite kabul edilen ve URAP gibi pek çok üniversite sıralamasında baz alınan YÖK üniversite yayın sayısı istatistikleri, çalışma kapsamında birleştirim işlemi sonrasında belirlenen üniversite yayın sayıları ile karşılaştırılmıştır. Sonuçta çalışmamızda elde edilen üniversite yayın sayılarının YÖK listesinden çok daha fazla olduğu sonucuna ulaşılmıştır. 2009 yılı yayın sayısı karşılaştırmaları Tablo 2'de verilmektedir.

Tablo 2: 2009 yılı yayın istatistiklerinin karşılaştırma tablosu

Üniversite adı	Yayın sayısı (YÖK)	Gerçek yayın sayısı	YÖK sıralaması	Gerçek sıralama
İstanbul Üniv.	1.338	1.690	1	1
Hacettepe Üniv.	1.182	1.561	2	2
Ankara Üniv.	1.170	1.355	3	3
Gazi Üniv.	1.066	1.287	4	4
Ege Üniv.	972	1.212	5	5
ODTÜ	830	971	6	6
Atatürk Üniv.	721	833	7	7
İstanbul Teknik Üniv.	623	732	8	10
Selçuk Üniv.	609	730	9	11
Ondokuz Mayıs Üniv.	603	697	10	12
Erciyes Üniv.	595	658	11	15
Dokuz Eylül Üniv.	593	746	12	9
Uludağ Üniv.	521	571	13	17
GATA	494	763	14	8
Marmara Üniv.	475	675	15	13
Çukurova Üniv.	474	576	16	16
Başkent Üniv.	472	661	17	14
Fırat Üniv.	455	507	18	19
Süleyman Demirel Üniv.	442	498	19	21
Karadeniz Teknik Üniv.	437	507	20	20

Tablo 2'den de anlaşılacağı üzere hatalı yayın adresi nedeni ile en çok yayın yapan ilk yedi üniversitenin sıralamasında değişiklik yaşanmasa da bu üniversitelerin belirlenen yayın sayıları arasında büyük farklılıklar olduğu görülmektedir. 8. sıra ve devamından itibaren yayın sayılarının yanında sıralamalarda da değişiklikler olduğu saptanmıştır. Örneğin, aslında yayın sayısı açısından 17. sırada yer alması gereken Uludağ Üniversitesi YÖK listesinde 13. sırada yer bulmuştur. Bunun gibi GATA'nın da, 14. sırada görünmesine rağmen, gerçek yerinin 8. sıra olduğu ortadadır. İlk 20 üniversite dışında da sıralamalarda büyük farklılıklar olduğu görülmüştür. Yayın sayısı azaldıkça üniversiteler arasındaki sıralamalarda daha büyük hatalar saptanmıştır.

Bu çalışma ile atıf dizinlerinde yapılan hatalar sınıflandırılarak hata türleri ortaya koyulmaya çalışılmıştır. Hataların neler olduğunun bilinmesi durumunda yazarlar, yayıncılar ya da dizinleme yapan kişilerin adres belirtirlerken daha dikkatli davranmaları sağlanabilecektir. Ancak ne kadar dikkatli davranılırsa davranılsın yapılan yayın sayısı listelerinin tam doğru sonucu yansıtmayacağı aşikârdır.

4.2. ATIF DİZİNLERİNDE ÜNİVERSİTE ADRESLERİNDE YAPILAN HATALAR

Verilerin incelenmesi ve birleştiriminin ardından üniversite adreslerinde yapılan hataların çeşitlilik gösterdiği saptanmıştır. Bu hataların yazar, editör ya da dizinlemeden kaynaklandığı düşünülebilir. Saptanan temel hata türleri aşağıda listelenmiştir.

4.2.1. Harf ve Heceleme Hataları

Çalışma kapsamında veriler incelenirken sıklıkla rastlanan hata çeşidinin üniversite isimlerinin yazımı sırasında yapılan klavye kaynaklı hatalar olduğu saptanmıştır. Damerau yaptığı çalışmada (1964) heceleme hatalarının %80'inden fazlasının harflerin fazla, yanlış ya da eksik yazılması ile klavyede birbirine bitişik olan harflerin yerlerinin değiştirilmesi nedenlerine bağlı olduğunu aktarmıştır. Üniversite adresleri üzerinde

yürütülen bu çalışmada da hataların büyük çoğunluğunun bu gibi harf ve heceleme yanlışlarından kaynaklandığı görülmüştür. “Hacetteppe Univ”, “Hacattepe Univ”, “Maramara Univ”, “Egge Univ”, “Inonoii Univ”, “Dukuz Eylul Univ” gibi adresler bu tip hatalara örnek olarak gösterilebilir. Ancak bu gibi hataların yalnızca yazar kaynaklı olduğunun düşünülmemesi gerekmektedir. Söz konusu hatalar dergi editörleri ya da dizinleme işinden sorumlu veri tabanı çalışanları kaynaklı da olabilmektedir.

4.2.2. Dizinleme Hataları

WoS’un dizinleme mantığı; kaynakların OCR (Optical character recognition) ile sayısallaştırılmasının ardından kayıtların elci sistem aracılığı ile veri tabanına girilmesi esasına dayanmaktadır (Thomson Reuters, 2009). Ancak *WoS*’un tanımlayıcı metinlerinde yazar adreslerinin nasıl seçildiği ve veri tabanına girildiği ile ilgili bilgi verilmemektedir. Thomson Reuters’ın teknik destek ekibi ile yapılan e-posta görüşmesinde adreslerin veri tabanına girilmesinde temel kuralın adreslerin orijinal metinde geçtiği şekli ile dizinlenmesi olduğu vurgulanmıştır.[3] Bunun yanında bazı sözcük ve sözcük grupları için belirlenen bir kısaltma listesinin varlığı konusuna da değinilmiştir.[4] *WoS*’un bu dizinleme mantığı sonucunda dizinleme sırasında insan hatalarının ve kaynakların sayısallaştırılması sırasında gerçekleşen bilgisayar kaynaklı hataların yapılması kaçınılmaz olacaktır.

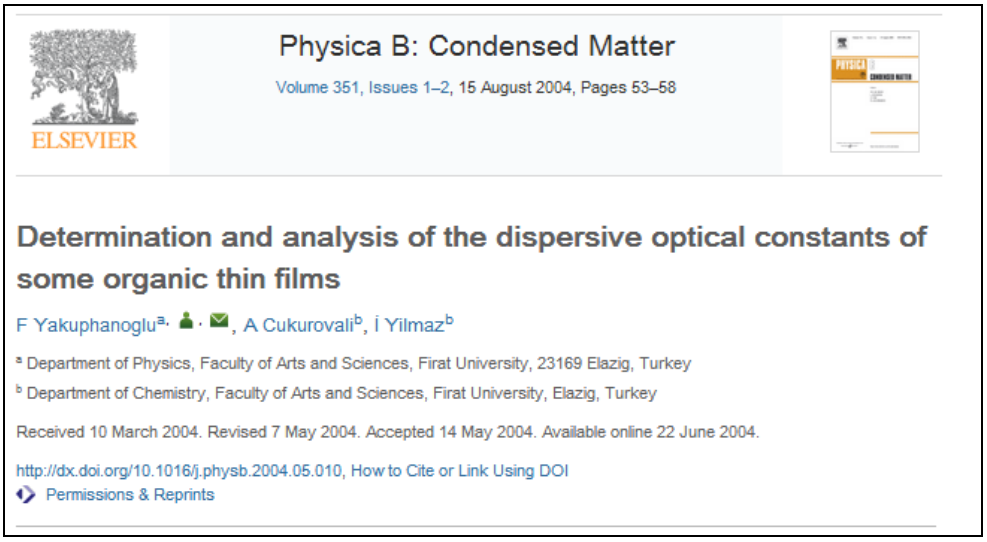
Sıkça yapılan ve oldukça ilginç olan dizinleme hatalarından birinin Türk üniversitelerin isimlerinin yerine kullanılan yabancı üniversite isimleri olduğu saptanmıştır. Bu kayıtlarda Yüzüncü Yıl Üniversitesi yerine “Centennial Univ” (4 kayıt); Niğde Üniversitesi yerine “Univ Nigeria” (6 kayıt); Fırat Üniversitesi yerine “Univ Florence” (12 kayıt), Başkent Üniversitesi yerine “Univ Basel” (1 kayıt), Gazi Üniversitesi yerine “Graz Univ” (1 kayıt) ve Kocaeli, Koç ve Afyon Kocatepe Üniversitelerinin yerine

---

[3] Thomson Reuters teknik destek ekibi ile yapılan 11.05.2012 tarihli e-posta görüşmesi

[4] *WoS* kısaltma listesi; http://images.webofknowledge.com/WOKRS56B5/help/WOS/hp_address_abbreviations.html

“Kochi Univ” (23 kayıt) adresleri kullanılmıştır. Bu tür hataların yapıldığı kayıtlar arasından bir örnek orijinal metinlerde yer alan adresleri ile birlikte Şekil 3 ve 4’te sunulmuştur. Thomson Reuters’ın orijinal metinde geçen adresin aynen alınması konusunda verdiği bilginin aksine bahsi geçen kayıtların orijinal metinlerinin büyük çoğunluğunda adresler doğru olarak girilmiştir. Bu gibi dizinleme hatalarında dizinlemeyi yapan kişinin adresleri açılır kutu aracılığı ile seçtiği ve bu aşamada hata yaptığı tahmin edilebilir. Ancak bu gibi bir uygulamanın varlığı konusunda herhangi bir bilgi verilmemiştir. Bu tip dizinleme hataları ile bahsi geçen üniversitelerin tüm yayınlarına erişim problemli hale gelirken, diğer taraftan bu hatalar nedeni ile Türkiye’nin işbirliği içinde olduğu üniversiteler saptanırken yanlış yorumlamaların yapılmasına neden olmaktadır. Örneğin, Türkiye’nin Nijerya Üniversitesi ile ortak yayını bulunmazken bu gibi hatalı kayıtlar nedeni ile Türkiye’nin Nijerya Üniversitesi ile işbirliği bulunduğu sonucuna ulaşılmaktadır.


ELSEVIER

Physica B: Condensed Matter

Volume 351, Issues 1–2, 15 August 2004, Pages 53–58

Determination and analysis of the dispersive optical constants of some organic thin films

F Yakuphanoglu[a], A Cukurovali[b], İ Yilmaz[b]

[a] Department of Physics, Faculty of Arts and Sciences, Firat University, 23169 Elazig, Turkey

[b] Department of Chemistry, Faculty of Arts and Sciences, Firat University, Elazig, Turkey

Received 10 March 2004. Revised 7 May 2004. Accepted 14 May 2004. Available online 22 June 2004.

http://dx.doi.org/10.1016/j.physb.2004.05.010, How to Cite or Link Using DOI

Permissions & Reprints

Şekil 3: Fırat Üniversitesi adresli yayının orijinal metninde geçen adresler[5]

[5] Yakuphanoğlu, F., Çukurovalı, A. ve Yılmaz, L. (2004). Determination and analysis of the dispersive optical constants of some organic thin films. *Physica B-Condensed Matter*, 351(1-2), 53-58.

Determination and analysis of the dispersive optical constants of some organic thin films

Author(s): Yakuphanoglu, F (Yakuphanoglu, F); Cukurovali, A (Cukurovali, A); Yilmaz, I (Yilmaz, I)

Source: PHYSICA B-CONDENSED MATTER **Volume:** 351 **Issue:** 1-2 **Pages:** 53-58 **DOI:** 10.1016/j.physb.2004.05.010 **Published:** AUG 15 2004

Times Cited: 33 (from Web of Science)

Cited References: 19 [view related records] Citation Map

Abstract: The optical constants (refractive index n, extinction coefficient k and dielectric constant epsilon) of some organic thin films were determined using reflectance and transmittance spectra. The real part and imaginary parts of the complex dielectric constant were calculated. The analysis of the optical absorption data revealed both optical direct and indirect transition with band gaps of 2.492-2.926 and 2.07-2.13 eV, respectively. The refractive index dispersion obeyed the Single oscillator of Wemple-Didomenico model and dispersion parameters and high-frequency dielectric constant were determined. (C) 2004 Elsevier B.V. All rights reserved.

Accession Number: WOS:000223439000008

Document Type: Article

Language: English

Author Keywords: optical constants; optical band gap; thin films

KeyWords Plus: SCHIFF-BASE LIGAND; METAL-COMPLEXES; ABSORPTION; BEHAVIOR; GLASSES; CO(II); NI(II)

Reprint Address: Yakuphanoglu, F (reprint author), Univ Florence, Fac Arts & Sci, Dept Phys, TR-23169 Elazig, Turkey

Addresses:
1. Univ Florence, Fac Arts & Sci, Dept Phys, TR-23169 Elazig, Turkey
2. Univ Florence, Fac Arts & Sci, Dept Chem, Elazig, Turkey

Şekil 4: Şekil 3'te yer alan yayının *WoS* kaydı

Türkiye'deki üniversitelerin isimleri yerine diğer ülkelerdeki üniversitelerin isimleri kullanıldığında atıf dizini üzerinden bu kayıtlara erişimde göze çarpan bir diğer unsur da sorguya eklenen sözcüklerin çok dikkatli seçilmesi gerektiğidir. Örneğin, üniversite adresinin "Univ Florence" olarak verildiğini bilen Fırat Üniversitesi kendi yayınlarına erişmek için sorgu sözcükleri arasına "Univ Florence" terimini aldığı takdirde İtalya'da bulunan üniversitenin kayıtlarına da kendi yayınıymış gibi erişecektir. "NOT Italy" gibi bir sözcük grubu ile içinde İtalya geçen kayıtları kapsam dışında bıraktığı takdirde ise kurumunun İtalya ile işbirliği yaparak yaptığı yayınları değerlendirme kapsamına alamayacaktır. Bunun yanında İtalya sorgu içinden çıkarılmış olsa da erişilen sorgu sonuçları içinde "Istanbul Bilim **Univ, Florence** Nightingale Hosp" kayıtları da yer alacaktır. Bu nedenle bir üniversite yerine tamamen farklı bir üniversitenin isminin yazılmış olması çok büyük sorunları da beraberinde getirmektedir. Eğer atıf dizini üzerinden tarama yapılarak bir kurumun tüm yayınlarına erişmek isteniyorsa bu gibi kayıtların değerlendirme kapsamından çıkarılması araştırmanın gidişatı açısından daha doğru bir çözüm olacaktır. Ancak hata yapmamak için izlenecek daha etkili yol atıf dizininden verilerin indirilmesinin ardından gerekli değerlendirmelerin yapılmasıdır. Çünkü veri tabanı üzerinden yapılan sorgular ile bir kurumun yayınlarının tümüne

erişmek uzun ve yorucu bir işlem gerektirmektedir. Bu gibi bir işlem ardından bile doğru sonuçlara ulaşılmasının zor olduğu görülmektedir.

Daha önce de belirtildiği gibi 2007 yılında çıkarılan *Türkiye Bilimsel Yayın Göstergeleri* kitabında da atıf dizini üzerinde üniversite adreslerini tarayabilmek için anahtar sözcükler belirlenmiş ve sunulmuştur (ULAKBİM, 2007). Ancak her bir kurum için belirlenen girişler daha detaylı incelendiğinde elde edilecek sonuçları büyük ölçüde etkileyecek hatalar görülmüştür. Bu anahtar sözcükler içinde aşağıda belirtilen örnekler gibi pek çok hataya rastlanmıştır.

- İstanbul Üniversitesi anahtar sözcükleri arasında “Istanbul Mem Hosp” (Memorial Hastanesi) ve “Tech Univ Istanbul” (İstanbul Teknik Üniversitesi),
- Marmara Üniversitesi anahtar sözcükleri arasında “Marmara Res Ctr” ve “Marmara Res Ins” (TÜBİTAK Marmara Araştırma Merkezi),
- Hacettepe Üniversitesi anahtar sözcükleri arasında “S Demirel Univ” (Süleyman Demirel Üniversitesi),
- Sakarya Üniversitesi anahtar sözcükleri arasında “Sakarya Mahallesi”,
- Gazi Üniversitesi anahtar sözcükleri içinde “AU Vet Fak” (Ankara Üniversitesi).

Hazırlanan bu kitap alanda büyük bir boşluğu doldurmuş olmasına rağmen bu gibi hataların yapılmış olması anahtar sözcük belirleyerek veri tabanı üzerinden yapılan taramaların yanlış sonuçlar doğurabileceğini ispatlamıştır.

Atıf dizinlerinde yapılan bir diğer dizinleme hatası türünün de harflerin başka harfler ile değiştirilmesi sonucu oluştuğu saptanmıştır. Yapılan araştırmada atıf dizinlerinde yer alan bazı kayıtlarda “m” harfinin yerine “rn”; “in” harflerinin yerine “m”, “i” harfinin yerine “1” ve “l” harfinin yerine “i” harflerinin kullanıldığı belirlenmiştir. “Parnukkale Univ”, "Dokuz Eylui Univ”, “F1rat Univ” ve “Dumlupmar Univ” bu tür hatalara örnek

olarak gösterilebilir. Sayılan bu sorunların daha önce de değinildiği gibi kaynakların sayısallaştırılması sırasında yapılan OCR hatalarına bağlanması mümkündür. OCR hatalarının büyüklüğü göz önüne alındığında sayısallaştırılmış kaynakların yeniden gözden geçirilmesi gerekliliği ortaya çıkmaktadır.

Saptanan en ilginç dizinleme hatalarından biri soy ismin üniversite adresi yerine kullanılmasıdır. Yazarlarından ikisi "Uslu, G." ve "Uslu, M." olan ve orijinal metninde doğru adresi içeren bir yayında adres "Uslu Univ" olarak verilmiştir (bkz. Şekil 5 ve 6).

Acne: prevalence, perceptions and effects on psychological health among adolescents in Aydin, Turkey

Author(s): Uslu, G (Uslu, G.)[1]; Sendur, N (Sendur, N.)[1]; Uslu, M (Uslu, M.)[1]; Savk, E (Savk, E.)[1]; Karaman, G (Karaman, G.)[1]; Eskin, M (Eskin, M.)[2]

Source: JOURNAL OF THE EUROPEAN ACADEMY OF DERMATOLOGY AND VENEREOLOGY **Volume:** 22 **Issue:** 4 **Pages:** 462-469 **DOI:** 10.1111/j.14 3083.2007.02497.x **Published:** APR 2008

Times Cited: 11 (from Web of Science)

Cited References: 50 [view related records] Citation Map

Abstract: Background Acne vulgaris is a common disease among adolescents and known to have adverse effects on psychological status.

Methods A cross-sectional study was conducted on 600 participants by means of a questionnaire designed for this study in high school students. In addi objective evaluation of acne in participants was undertaken. The questionnaire consisted of questions about acne, General Health Questionnaire (GHQ) Rosenberg Self-Esteem Scale (RSES).

Results Five hundred and sixty-three questionnaires out of 600 were answered and 550 adolescents who gave permission for examination were evaluat study population consisted of 303 girls and 260 boys between the ages 13 and 19, and the mean age was 15.24 +/- 1.05 years. Acne prevalence was 63. 29.2% non-inflammatory and 34.4% inflammatory acne. It was more prevalent and severe in boys than in girls. Not the objective but the subjective severit and opinion that one could benefit from acne treatment was found to be related to anxiety, depression and self-esteem. Factors implicated among cause were food, bad skin hygiene and hormones in decreasing frequency. Forty-eight per cent of adolescents expect a maximum duration of 4 weeks for treatn

Conclusion Despite the high prevalence of acne, there is still much deficiency of knowledge and wrong beliefs about acne. This indicates that there is an for education about etiopathogenesis, potential complications and importance of effective treatment for acne. Effective treatment may make significant co for the mental health of adolescent and as well as adult populations.

Accession Number: WOS:000254276000010

Document Type: Article

Language: English

Author Keywords: acne; adolescents; prevalence; psychological health

KeyWords Plus: QUALITY-OF-LIFE; DERMATOLOGICAL PATIENTS; SUICIDAL IDEATION; EMOTIONAL IMPACT; SEBACEOUS GLANDS; SKIN-DISEASE; ACNE; VULGARIS; DEPRESSION; ISOTRETINOIN

Reprint Address: Uslu, M (reprint author), Uslu Univ, Fac Med, Dept Dermatol, TR-09100 Aydin, Turkey

Şekil 5: "Uslu Univ" adresli yayının *WoS* kaydı

ORIGINAL ARTICLE

Acne: prevalence, perceptions and effect health among adolescents in Aydin, Turk

G Uslu,† N Şendur,† M Uslu,*† E Şavk,† G Karaman,† M Eskin‡

†Adnan Menderes University Faculty of Medicine, Department of Dermatology, ‡Department o

Keywords
acne, adolescents, prevalence, psychological health

Abstract

Background Acne vulgaris is

Şekil 6: Şekil 6'da verilen kaydın orijinal metni

Şekil 6'da da görüldüğü üzere kaynağın orijinal metninde yazar adresi "Adnan Menderes University Faculty of Medicine, Department of Dermatology" olarak verilmiş olmasına rağmen kayıt *WoS*'da "Uslu Univ" şeklinde dizinlenmiştir. Bunun olası nedeni yazar ismi ile kurum isminin birbirine karıştırılması olarak görülebilir. Burada önemli olan nokta bu gibi hataların saptanmasının ancak ve ancak tüm kayıtların tek tek incelenmesi ile mümkün olabileceği gerçeğidir. Başka herhangi bir şekilde bu kaydın bulunması imkânsızdır. Kaydın bulunmasına engel olacak kadar büyük bu gibi hatalara sıklıkla rastlanmasa da diğer üniversite adreslerinde de benzer yanlışlar görülmektedir.

4.2.3. Üniversite İsimlerinin İngilizce Çevirilerinde Yapılan Hatalar

Türkiye'de pek çok üniversite hem yurt içi hem de yurt dışında aynı isimle anılmaktadır. Birkaç üniversite dışında (Orta Doğu Teknik Üniversitesi - Middle East Technical University; İstanbul Teknik Üniversitesi - Istanbul Technical University; Gülhane Askeri Tıp Akademisi - Gulhane Military Medical Academy gibi) üniversite isimlerinde çeviri yapılmamaktadır. Atıf dizinlerinde sıklıkla yapılan hatalardan biri Türkçe üniversite isimlerinin İngilizceye çevrilmesi olarak saptanmıştır. Örneğin, Boğaziçi Üniversitesi atıf dizinlerinde kimi zaman "Bosphorus Univ" olarak anılıyor olmasına rağmen üniversitenin İngilizce ismi de Boğaziçi University'dir (http://boun.edu.tr/). Benzer şekilde Dokuz Eylül Üniversitesi yerine "Nine September Univ" kullanılmıştır.

Bu durum dünyada başka bir "Bosphorus Univ" ya da "Nine September Univ" olmadığından bahsi geçen üniversiteler için büyük bir problem teşkil etmemektedir. Ancak, tüm üniversiteler için bunu söylemek zordur. Ege Üniversitesinin pek çok yayınında üniversitenin isminin "Aegean Univ" olarak geçtiği görülmüştür. Yunanistan'da da "University of Aegean" adında bir üniversite bulunmaktadır (http://www3.aegean.gr/) ve bu üniversitenin atıf dizinlerinde yayınları yer almaktadır. Bunun yanında Akdeniz Üniversitesinin kendi adresi yerine kullandığı "Mediterranian

Univ" adresi Fransa'da bir üniversite ile (http://www.univmed.fr/) ve Trakya Üniversitesinin kendi adresi yerine kullandığı "Trakia Univ" adresi Bulgaristan'daki bir üniversite ile (http://www.uni-sz.bg/engl) çakışmaktadır. Bu nedenle Türkiye adresli yayınların bulunmasında bu sorgu sözcüklerini sorguya dâhil etmek elde edilecek sonuçları artırmasının yanında yapılacak bibliyometrik çalışmaların da hatalı olmasına neden olacaktır. Dâhil edilmemesi durumunda ise üniversitenin yayınlarında çok yüksek miktarda eksiklik çıkacaktır. Örneğin, Ege Üniversitesinin yayınlarının pek çoğunda üniversite adresinin "Univ Aegean" olarak verilmiş olması bu üniversitenin yayınlarının ya kaybolmasına ya da Ege Üniversitesine ait olmayan kayıtların da Ege Üniversitesininmiş gibi sayılmasına neden olabilecektir.

4.2.4. Diğer Hatalar

Atıf dizinlerinde yukarıda bahsedilen ve sıklıkla karşılaşılan hataların yanında daha farklı hatalara da rastlanmıştır. Bu gibi hataların örnekleri aşağıda verilmiştir.

Bazı kayıtlarda üniversitelerin isimlerinin yerine bu üniversitelerin bulunduğu illerin isimlerinin kullanıldığı görülmüştür. Bu gibi kayıtlara Çukurova Üniversitesi yerine kullanılan Adana Üniversitesi; Ege ve Dokuz Eylül Üniversiteleri yerine kullanılan İzmir Üniversitesi ve Selçuk Üniversitesi yerine kullanılan Konya Üniversitesini örnek olarak göstermek mümkündür. Bu durum adı verilen ilde bu adı taşıyan başka bir üniversite bulunmaması halinde büyük bir sorun oluşturmayacaktır. Ancak eğer o ilde bu adı taşıyan başka bir üniversite var ise problem yaşanacaktır. Örneğin, 2007 yılından önce Dokuz Eylül ve Ege Üniversitelerinin isimleri yerine İzmir Üniversitesi kullanılırken, 2007 yılında İzmir Üniversitesinin kurulması ile birlikte atıf dizininde bu konuda karışıklık yaşandığı görülmüştür. Bu durum kayıtların dikkatli bir şekilde incelenmesi gerekliliğini bir kez daha gündeme getirmektedir.

Bazı yayınlarda üniversite isimlerinin yerlerine çok çeşitli ve niçin kullanıldığı anlaşılamayan ilginç isimler kullanıldığı da görülmüştür. Örneğin, ODTÜ'nün yerine "Tech Univ Angora", "Tech Univ Ankara" gibi kullanımların olduğu saptanmıştır.

Adres alanında Türkiye bilgisi olmayan ancak Türkiye'de yer alan üniversitelerin adreslerini taşıyan kayıtlar da bulunmuştur. Bu gibi kayıtlara örnek olarak Hacettepe Üniversitesi yayınının İsrail (Hacettepe Univ, Inst Child Health, DNA Cell Bank & Gene Res Lab, Jerusalem, Israel) ve Gazi Üniversitesi yayınının Mısır adresi (Gazi Univ, Fac Pharm, Dept Pharmacognosy, Giza, Egypt) taşıması gösterilebilir. Bu iki örnek için de orijinal metin incelenmiş ve orijinal metinde hata olmadığı, hatanın dizinlemeden kaynaklandığı görülmüştür. Bu gibi hataların saptanması çok zordur. Çünkü bu çalışmaların adres alanlarında Türkiye geçmediği için çok daha detaylı değerlendirmelerin yapılması zorunludur. ULAKBİM'in kitabında (ULAKBİM, 2007) tüm üniversite adresi formüllerine "Turkey OR Turkiye" eklenmiş ve bu kapsama Türkiye adresli olmayan yayınlar alınmamıştır. Zaten TÜBİTAK da adresinde Türkiye geçmeyen yayınlara UBYTP kapsamında teşvik vermemektedir (ULAKBİM, 2010: 2).

Yukarıda sayılan durumun tam tersi olarak adres alanında Türkiye olan ancak Türkiye'de eğitim veren üniversitelere ait olmayan yayınlar da bulunmuştur. "Ruppin Acad Coll, Dept Psychol, Emek, Hefer, Turkey; Tabriz Univ Med Sci, Div Nephrol, Eman Hosp, Tabriz, Turkey; Democritus Univ Thrace, Fac Med, Dept Pediat, TR-22030 Edirine, Turkey" gibi kayıtlar bu tür hatalara örnek olarak gösterilebilir. UBYTP kapsamında Türk veya yabancı uyruklu araştırmacılara destek verilebilmektedir. Bu mantık ile düşünüldüğünde adres alanında "Turkey" geçen bu çalışmaların yazarları Türkiye'de çalışmıyor ve Türk uyruğu taşımıyor olmalarına rağmen UBYTP kapsamında destek hakkına sahip görünmektedir. Bu gibi sebeplerden ötürü teşvik programının esaslarının yeniden değerlendirilmesi gereklidir.

Üniversite adreslerinde yapılan dizinleme, heceleme ve diğer hataların yanında adreslerin erişilir olmasını engelleyen bir diğer etmenin de bu adreslerde bulunan standart yapı sorunu olduğunu söylemek yanlış olmayacaktır. Bu nedenle standart yapı sorunu üzerinde de çalışılması gerekmektedir.

4.3. ÜNİVERSİTE ADRESLERİNDE STANDART YAPI SORUNU

Üniversite adreslerinde yapılan yazım hatalarının yanında bu adreslerde standart bir yapının olmadığı da saptanmıştır. Standart yapı ile kastedilen üniversite adresleri verilirken temel bir sıralama ya da kullanımın olmayışıdır. Galvez ve Moya-Anegón (2007a: 8) adres standardizasyonu üzerine yaptığı çalışmada bir adres kaydının üniversite, fakülte, bölüm ismi, posta kodu, şehir ve ülke bilgisini içermesi gerektiğini belirtmiştir. Ancak Türkiye adresli yayınların genelinde bu yapı görülmemiştir.

Bu tip standart dışı kullanımların en sık rastlanan örneği “Med Sch” ve “Sch Med” kullanımları olmuştur. Üniversitelerin tıp fakülteleri “Univ” terimini kullanmak yerine bu iki sözcüğü kullanmışlardır (“Ankara Sch Med”, “Hacettepe Med Sch” gibi).

Bir diğer standart olmayan kullanım şekli üniversite isimlerinin kısaltılarak verilmesidir. Ancak bu yapıda da bariz bir karmaşa olduğu görülmüştür. Örneğin, Harran ve Hacettepe Üniversitelerinin ikisinin de “HU” kısaltmasını kullandıkları saptanmıştır. Bu gibi kullanımlar oldukça yaygındır. Bu kullanımlar iki ya da daha fazla Türk üniversitesinin yayınlarının birbiri ile karıştırılmasının yanında çok daha karmaşık durumlara neden olabilmektedir. Örneğin, *WoS*’da adres alanında “HU” ve “Turkey” arama sözcükleri kullanılarak yapılan bir taramada Şekil 7, 8 ve 9’daki gibi örneklere rastlanmıştır.

Synthesis, spectroscopy, X-ray structure and redox behaviors of 4-(N-R-salicylideneimine)-2,6-diphenylphenols

Author(s): Kasumov, VT (Kasumov, Veli T.)[1]; Turkmen, H (Tuerkmen, Hasan)[1]; Ucar, I (Ucar, Ibrahim)[2]; Bulut, A (Bulut, Ahmet)[2]; Yayli, N (Yayli, Nurettin)[3]

Source: SPECTROCHIMICA ACTA PART A-MOLECULAR AND BIOMOLECULAR SPECTROSCOPY **Volume:** 70 **Issue:** 1 **Pages:** 60-68 **DOI:** 10.1016/j.saa.2007.07.021 **Published:** JUN 2008

Times Cited: 2 (from Web of Science)

Cited References: 50 [view related records] Citation Map

Abstract: A series of sterically hindered 4- (N-R-salicylaidimine)-2,6-diphenylphenols (X), where R = H (1), 3-CH3 (2), 5-CH3 (3), 3-OCH3 (4), 4-OCH3 (5), 5-OCH3 (6), 3-Bu-t (7), 5-Bu-t (8), 3,5-Bu-t(2)(9) and 5,6-benzo(10), were synthesized and their structure as well as redox behavior studied by analytical, spectroscopic [H-1, (C-13) NMR, IR, UV-vis and mass spectrometry] and cyclic voltammetric (CV) techniques. Single crystal X-ray diffraction studies of 7 evidenced its existence as non-planar enol-imine tautomer structure, in which the phenol ring of the molecule is twisted around C-N single bond by 21.5(2)degrees. The packing structure of 7 is stabilized by C-H center dot center dot center dot pi(Ph) and O center dot center dot center dot O and C center dot center dot center dot O intermolecular short contact interactions. The CV of X display rate is dependent on irreversible and quasi-reversible redox waves in the anodic and cathodic regions due to oxidation and reduction of phenolic and iminic groups, respectively. As evidenced by ESR and UV-vis study, chemical oxidation of X by PbO2 and (NH4)(2)Ce(NO3)(6) in MeCN and CHCl3 generates stable phenoxyl radicals [(g similar to 2.005 and lambda similar to 450 nm (1600-8200 M-1 cm(-1))]. (C) 2007 Elsevier B.V. All rights reserved.

Accession Number: WOS:000256556000009

Document Type: Article

Language: English

Author Keywords: 2,6-diphenylphenol-salicylaldimines; redox behavior; X-ray crystal structure

KeyWords Plus: SCHIFF-BASES; GALACTOSE-OXIDASE; CRYSTAL-STRUCTURE; ELECTROCHEMICAL CHARACTERIZATION; PROTON-TRANSFER; SOLID-STATE; COMPLEXES; REDUCTION; RADICALS; SALICYLIDENEANILINE

Reprint Address: Kasumov, VT (reprint author), HU, Fac Arts & Sci, Dept Chem, Sanliurfa, Turkey

Addresses:
1. HU, Fac Arts & Sci, Dept Chem, Sanliurfa, Turkey
2. OMU, Fac Arts & Sci, Dept Phys, Kurupelit, Samsun, Turkey
3. KTU, Fac Arts & Sci, Dept Chem, Ankara, Turkey

Şekil 7: Harran Üniversitesi için kullanılan HU kısaltması örneği

Teachers' proficiency perceptions of about the measurement and evaluation techniques and the problems they confront

Author(s): Gelbal, S (Gelbal, Selahattin)[1]; Kelecioglu, H (Kelecioglu, Huelya)[1]

Source: HACETTEPE UNIVERSITESI EGITIM FAKULTESI DERGISI-HACETTEPE UNIVERSITY JOURNAL OF EDUCATION **Issue:** 33 **Pages:** 135-145 **Published:** 2007

Times Cited: 16 (from Web of Science)

Cited References: 12 [view related records] Citation Map

Abstract: In this study, the teachers' opinions about measurement and evaluation methods which they use in their classrooms were tried to be described. The survey developed for this purpose was administred to 242 classroom teachers and branch teachers teaching 1-6. grades in primary schools in centeral province of Ankara. The teachers preferred traditional measurement methods at which they considered themselves more adaquate to determine students' achivement. The most important problem which the teachers meet at using measurement instruments is crowded classrooms and insufficient course duration. A great part of teachers' opinion that has showed parallelism that the primary school programme. Taking this finding and other findings of the study into consideration, it is understood that teachers need education on using and preparing measurement methods.

Accession Number: WOS:000256878200013

Document Type: Article

Language: Unspecified

Author Keywords: measurement instruments; teacher proficiencies; primary school program; measurement and evaluation

Reprint Address: Gelbal, S (reprint author), HU Egitim, Fak Egitim Bil Bol Egitimde Olcme Degerlendirm An, Ankara, Turkey

Addresses:
1. HU Egitim, Fak Egitim Bil Bol Egitimde Olcme Degerlendirm An, Ankara, Turkey

Şekil 8: Hacettepe Üniversitesi için kullanılan HU kısaltması örneği

Primes in tuples II

Author(s): Goldston, DA (Goldston, Daniel A.)[1]; Pintz, J (Pintz, Janos)[2]; Yildirim, CY (Yildirim, Cem Yalcin)[3]

Source: ACTA MATHEMATICA **Volume:** 204 **Issue:** 1 **Pages:** 1-47 **DOI:** 10.1007/s11511-010-0044-9 **Published:** MAR 2010

Times Cited: 0 (from Web of Science)

Cited References: 20 [view related records] Citation Map

Accession Number: WOS:000274962700001

Document Type: Article

Language: English

KeyWords Plus: NUMBERS

Reprint Address: Goldston, DA (reprint author), San Jose State Univ, Dept Math, 1 Washington Sq, San Jose, CA 95192 USA

Addresses:
1. San Jose State Univ, Dept Math, San Jose, CA 95192 USA
2. Hungarian Acad Sci, Alfred Renyi Inst Math, HU-1053 Budapest, Hungary
3. Bogazici Univ, Dept Math, TR-34342 Istanbul, Turkey

Şekil 9: Macaristan'ın ülke kodu HU'nun adres alanında kullanımı

Bir üniversitenin yayınlarının isminin kısaltması kullanılarak taranması, eğer o kurumun kendine has bir kısaltması yoksa (ODTÜ gibi), doğru sonuçlar getirmeyecektir. “HU” anahtar sözcüğü kullanılarak yapılan bir taramanın Harran Üniversitesi, Hacettepe Üniversitesi, Haliç Üniversitesi, Hakkari Üniversitesi ve Hitit Üniversitesi gibi üniversitelere ait yayınları getirmesi kaçınılmazdır. Ayrıca bu şekilde yapılacak bir tarama ile Macaristan ülke kodunun HU olması nedeni ile Macaristan’a ait yayınlara erişilecek olmanın yanında daha pek çok farklı kuruma ait kayda da erişmek mümkündür (“ICO Badalona, HU Germans, Barcelona, Spain”, “HU Bellvitge, Lhospitalet De Llobregat, Spain”, “HU Vaudois, Lausanne, Switzerland” gibi). Tüm bu nedenlerden dolayı yazarların adres verirken kısaltma kullanmaması büyük önem taşımaktadır.

İl adı içeren üniversitelerin isimlerinde standart olmayan bir kullanım görülmüştür. Bazı kayıtlarda il adı üniversite ismiymiş gibi verilirken bazı kayıtlarda il adları geçmemiştir. Örneğin, Kahramanmaraş Sütçü İmam Üniversitesi adresi pek çok farklı şekilde verilmiştir. Bunlar: “Kahramanmaraş Univ”, “Sutcu Imam Univ”, “Kahramanmaras Imam Univ”, “Imam Univ”, “Sutcu Univ” vb.

Son olarak, birleşik ismi olan üniversitelerde bu isimlerin birleşik veya ayrı yazımı konusunda çeşitli kullanımlar görülmüştür. Örneğin, Osmangazi ve

Gaziosmanpaşa Üniversitelerinde bu duruma ek olarak özel bir sorun gözlemlenmiştir. Bu iki üniversitenin isimleri ayrı yazıldığı takdirde arama sonuçları içinde Gazi Üniversitesi yayınlarının da yer alması kaçınılmaz olacaktır. Bu durum bahsi geçen üniversitelerin yayınlarına erişimde yine problemler yaratmaktadır.

Adreslerde yapılan hatalar ve standardizasyon sorunu bilimsel arenada üniversitelerin görünürlüğünü azaltmanın yanında pek çok farklı problemi de beraberinde getirmektedir. Bu nedenle sorun ortaya koyulurken yarattığı etkileri de bilmek konunun öneminin daha net bir şekilde anlaşılmasına yardımcı olacaktır.

4.4. ADRESLERDE YAPILMIŞ HATALAR VE STANDARDİZASYON SORUNLARININ ETKİLERİ

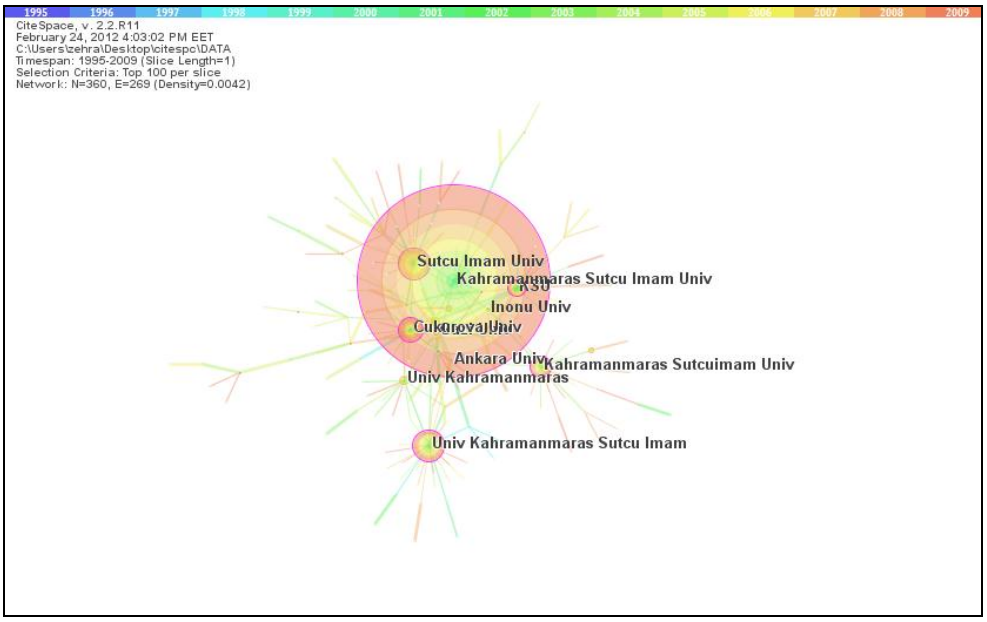
Atıf dizinlerinde üniversite adreslerinde yapılmış hatalar pek çok problemi de beraberinde getirmektedir. En önemli problem üniversitelerin görünürlüğünün azalması olarak düşünülse de, daha farklı alanlarda da problemler yarattığı aşikârdır.

Daha önceki kısımda da belirtildiği üzere teşvik sistemi üzerinde (UBYTP) yayınlarda verilen adreslerin önemli bir etkisi olduğu görülmüştür. Bir yazar adres alanında Türkiye sözcüğünü kullanmadıysa TÜBİTAK'tan bilimsel yayın teşviki sağlayamamaktadır. Ancak tek sorun bilimin teşvik edilmesi ile ilgili programlarda yaşanan sıkıntılar olarak görülmemelidir. Kurumların veya ülkelerin bilimsel performanslarını değerlendirmek amacı ile kullanılan bibliyometrik çalışmaların hatalı olması da adres standardizasyonunu önemli hale getirmektedir.

Bibliyometri ile ilgili bilimsel araştırmalarda kurum adresleri sıklıkla kullanılmaktadır. Ancak kurum adreslerinde yapılmış hatalar bibliyometrik çalışmaların da hatalı olmasına neden olabilmektedir. Bu çalışma kapsamında üniversite adreslerinde yapılan hataların bilimsel araştırmaları ne ölçüde etkileyebileceği de araştırılmıştır.

Bibliyometrik çalışmalar ile kurumlar arasında işbirliği ağları yaratılmakta ve bu ağlar aracılığı ile hangi kurumun diğer hangi kurumlarla birlikte yayın yaptığı, aralarındaki ilişkiler ve yayın yılları da göz önüne alınarak saptanmaktadır. Bu gibi çalışmalarda verilerin doğru ve güvenilir olmasının yanında standart bir biçimde olması da önemlidir. Standart olmayan kurum isimleri kullanılarak yaratılmış işbirliği haritaları hem görsel olarak çok şey anlatmamakta hem de bağlantıları net bir şekilde ortaya koyamamaktadır.

Bu çalışma kapsamında yayın adreslerinde en çok hata yapılan üniversite olan Kahramanmaraş Sütçü İmam Üniversitesinin kurumsal işbirliği ağı veriler standart biçime dönüştürülmeden önce ve standartlaştırıldıktan sonra haritalandırılmıştır. Veriler standart biçime dönüştürülmeden önce yaratılan ağ haritası Şekil 10'da verilmektedir.

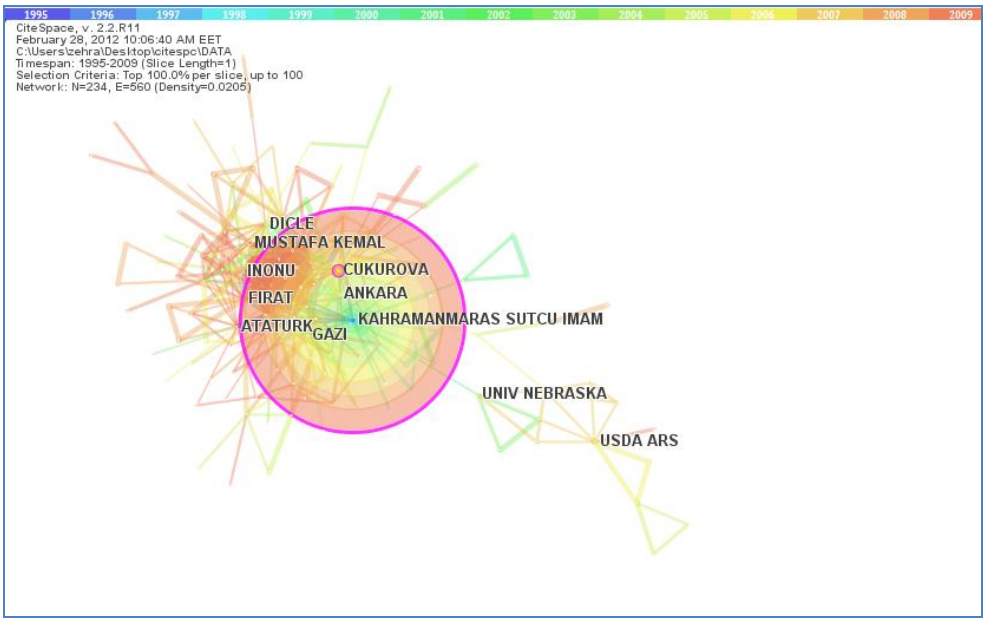


Şekil 10: Kahramanmaraş Sütçü İmam Üniversitesi kurumsal işbirliği ağı

Bu gibi ağlar kurumların işbirliği içinde oldukları diğer kurumları göstermek açısından önemlidir. Şekil 10'daki haritada Kahramanmaraş Sütçü İmam Üniversitesinin birlikte yayın yaptığı üniversiteler gösterilmektedir. Ancak dikkatli bakıldığında haritadaki düğümlerin (node) pek çoğunun Kahramanmaraş Sütçü İmam Üniversitesinin isminin farklı kullanımlarından oluştuğu görülmüştür. "Sutcu Imam Univ",

“Kahramanmaras Sutcu Imam Univ”, “KSU”, “Kahramanmaras Sutcuimam Univ”, “Univ Kahramanmaras Sutcu Imam” gibi kullanımlar sağlıklı bir bibliyometrik harita yaratmayı ve bu haritayı yorumlamayı zorlaştırmaktadır.

Standardizasyonun önemini göstermek amacı ile Kahramanmaraş Sütçü İmam Üniversitesine ait adresler tek biçime dönüştürülerek harita yeniden oluşturulmuştur. Bu haritada konunun önemi daha açık bir şekilde anlaşılmaktadır (bkz. Şekil 11).



Şekil 11: Standardize edilmiş Kahramanmaraş Sütçü İmam Üniversitesi verileri ile kurumsal işbirliği ağı

Şekil 11’de de görüldüğü üzere standardize edilmiş adresler sayesinde üniversiteler arasındaki işbirliği daha net bir şekilde görüntülenebilmiş ve bağlantılar ortaya çıkarılabilmiştir. Bir önceki haritada bağlantısı görüntülenemeyen pek çok düğüm bu standardizasyon işleminin ardından Şekil 11’de yer alan haritada ortaya çıkarılabilmiştir.

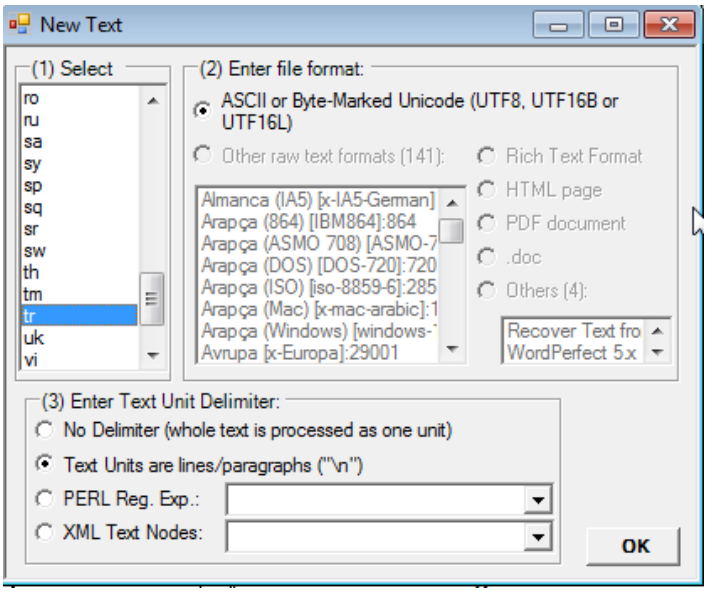
Şekil 10 ve 11’de gösterilen haritalardan da anlaşılacağı gibi üniversite adreslerinin standardizasyonu sadece üniversite yayın sıralamaları açısından değil; bibliyometrik çalışmaların doğruluğu, güvenilirliği ve anlaşılırlığı açısından da büyük önem taşımaktadır. Ancak üniversite adreslerinin

standardizasyonun elci sistem ile yapılması hem yoğun zaman gerektirmesi açısından hem de verilerin doğru bir şekilde standardize edilmesinin zorluğu açısından sıkıntılar doğurmaktadır. Bu nedenle de üniversite isimlerinin standardizasyonunu sağlamak için otomatik bir sistem geliştirilmesi önemlidir. Bu çalışma kapsamında standardizasyon sorununu çözebilecek bir algoritma tasarlanmasına çalışılmıştır.

4.5. ÜNİVERSİTE ADRESLERİ STANDARDİZASYON SORUNU İÇİN ÇÖZÜM ÖNERİLERİ

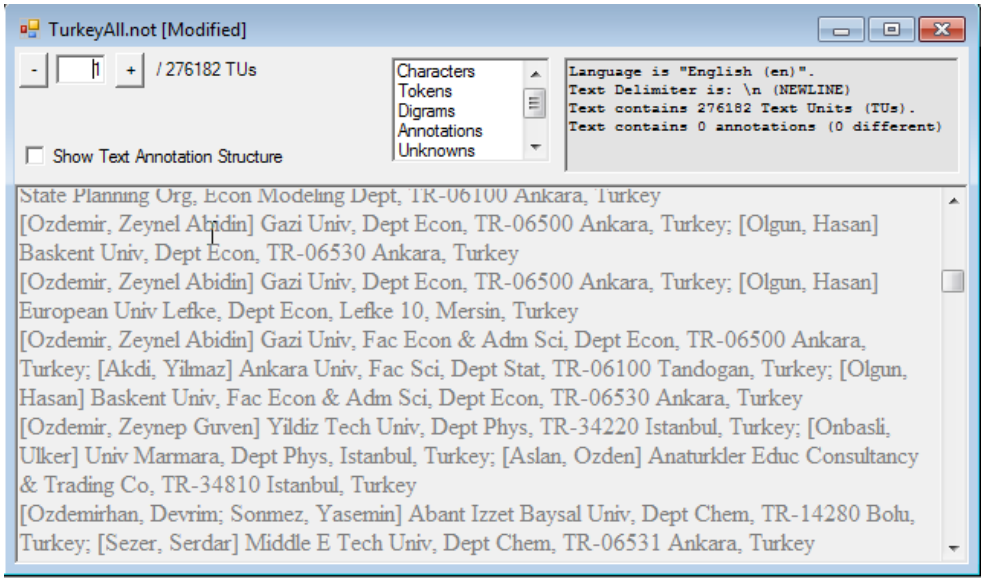
Üniversite adreslerinde yapılan hataların çeşitlilik gösterdiği, bu durumun hem üniversitelerin yayın sayılarının tam olarak belirlenememesine hem de bibliyometrik çalışmaların doğru sonuçlar ortaya çıkaramamasına neden olduğu daha önceki bölümlerde sıralanan bulgular aracılığı ile sunulmuştur. Çalışmanın bu kısmında Nooj programı kullanılarak, sonlu durum tekniği ile yaratılmış bir adres dönüştürücü ile standardizasyon sorununa çözüm getirilmeye çalışılmıştır.

İlk aşamada Türkiye adresli tüm yayınların adres alanlarında (C1 ve RP) yer alan bilgiler kullanılarak bir .not (metin) dosyası oluşturulmuştur. Bu dosyanın oluşturulması için Nooj programının File >> New >> Text yolu kullanılmıştır. Metin dosyası oluşturmak üzere açılan pencerede metin dosyasında kullanılacak dili, dosya formatını ve dosya ayırım biçimini seçmek mümkündür. Nooj programı içinde bir Türkçe Dil paketi yer almaktadır. Ancak atıf dizininde Türkçe karakterlere yer verilmemiş olması nedeni ile Türkçe dil paketi çalışma kapsamında kullanılmamıştır.



Şekil 12: Nooj yeni metin dosyası yaratma ekranı

Şekil 12'de yer alan ekranda gerekli seçim işlemlerinin gerçekleştirilmesinin ardından boş bir metin dosyası ekrana gelmektedir. Bu ekrana Türkiye adresli 198.687 kaydın C1 ve RP alanlarında yer alan adresler kopyala yapıştır yöntemi ile yapıştırılmıştır. Kopyalama işleminden hemen önce tekrar eden adresler ayrıştırılmış, bu sayede işlemlerin programı yormadan daha hızlı bir şekilde yapılması sağlanmıştır. Türkiye adresli tüm yayınların adreslerini içeren .not dosyası Şekil 13'teki gibi oluşturulmuştur.



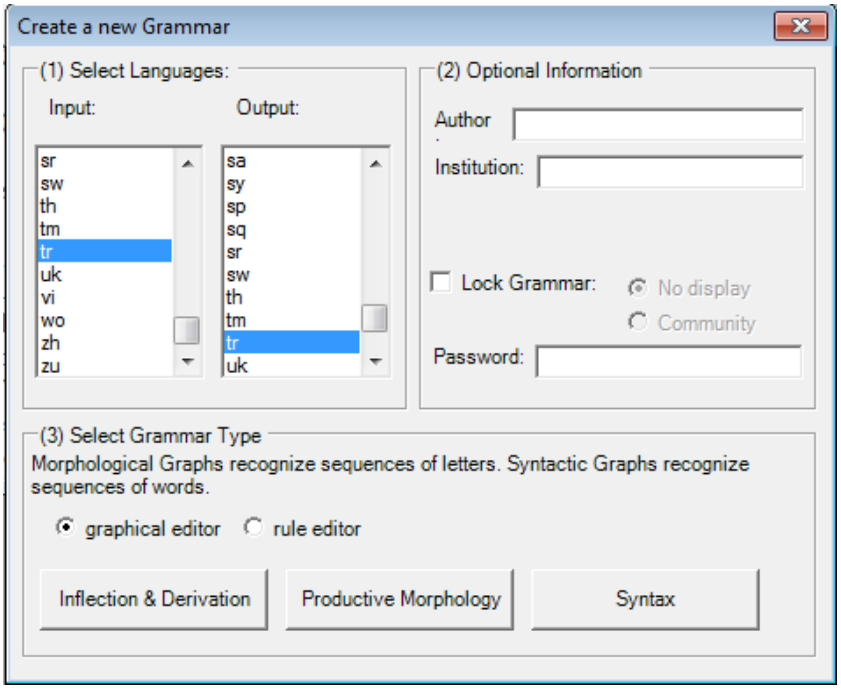
Şekil 13: Türkiye adresli yayınlar için oluşturulan .not dosyası

Türkiye adresli yayınların yer aldığı .not dosyası üzerinden herhangi bir üniversite için hatalı ve doğru adreslerin tespitinin sağlanması amaçlanmıştır. Bu amaç doğrultusunda adreslerin tespiti için atıf dizinleri kapsamında en çok yayını olan Hacettepe Üniversitesi örneği seçilmiştir. Hacettepe Üniversitesi adresleri için iki aşamalı bir yapı oluşturulmuş ve bu yapı sonucunda hatalı adreslerin tespitini yapabilmek hedeflenmiştir.

4.5.1. 1. Aşama: Hatalı Adres Girişlerinin Tespiti

Hacettepe Üniversitesi için hatalı adres girişlerinin saptanması amacı ile Nooj programı kullanılarak bir sonlu durum grafiği çizilmiştir. Bu grafiğin çizilmesindeki amaç olası adres çeşitlerinin sonlu durum dönüştürücüsüne tanımlanarak dönüştürücünün bu adresleri bulabilmesine olanak sağlamaktır.

Nooj içinde morfolojik gramer olarak tanımlanan bu grafikleri yaratabilmek için File >> New >> Grammar patikası takip edilmiştir (bkz. Şekil 14).

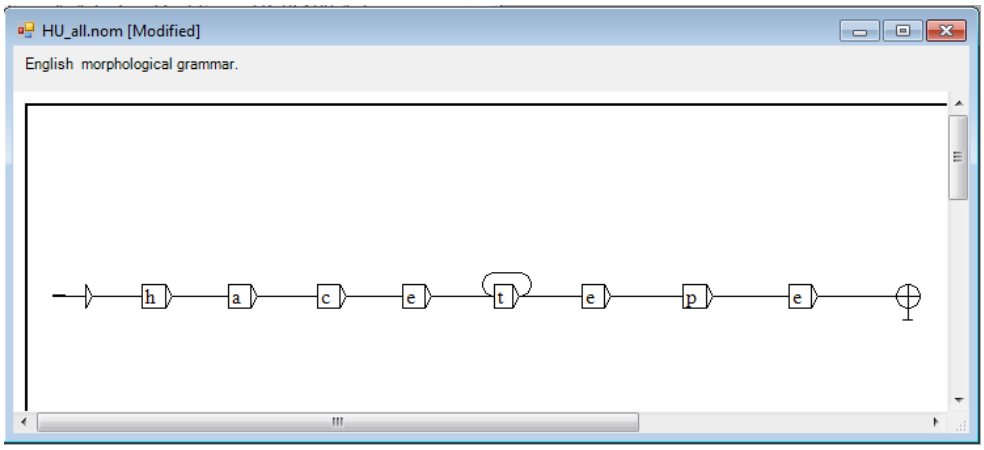


Şekil 14: Nooj yeni gramer yaratma başlangıç ekranı

Bu grameri yaratmadaki temel amaç bir sözcüğün (Hacettepe) farklı kullanımlarını (hatalı - doğru) saptamak olduğundan gramer tipi *productive*

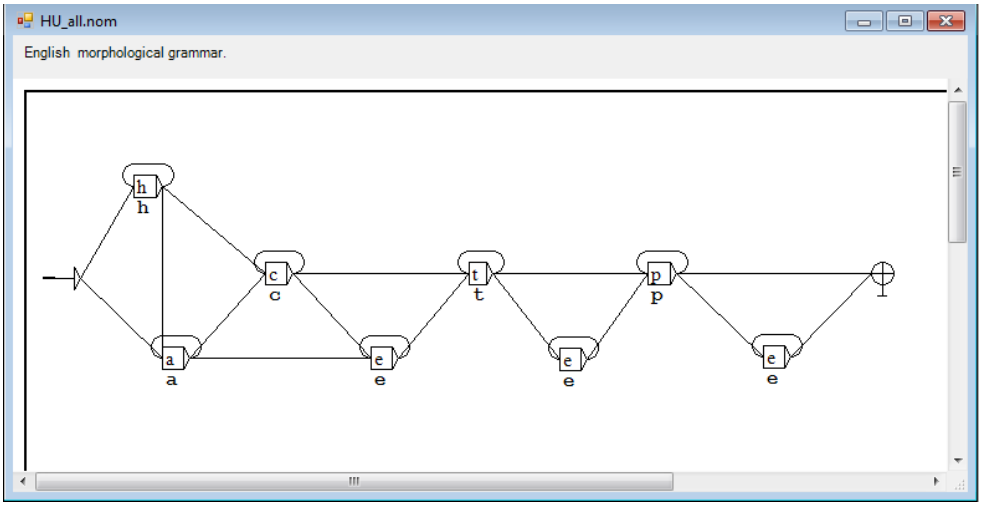
morphology olarak belirlenmiştir. Aksi takdirde yaratılan gramer dosyasını oluşturulan metin dosyası üzerine uygulamak mümkün olmayacaktır.

Productive morphology düğmesine tıklandığında açılan grafik editörü arayüzü ile ilk aşamada Hacettepe isminin doğru kodlaması yapılmıştır. Doğru kayıtları getirecek bu grafik Şekil 15'teki gibi tanımlanmıştır.



Şekil 15: Hacettepe teriminin doğru yazımlarını getirecek sonlu durum grafiği

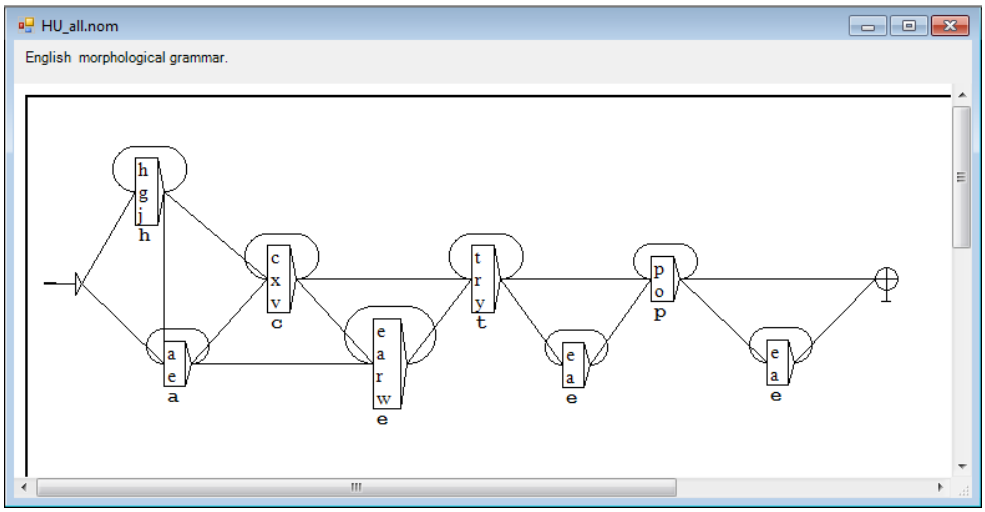
Doğru yazımları getirecek sonlu durum grafiğinde algoritma harfleri tarayacak ve en son duruma gelene kadar tanımlanan duruma uyan tüm kayıtları getirecektir. "t" harfindeki döngünün nedeni Hacettepe teriminde bulunan "t" harfindeki ikilemedir. Algoritma "t" harfine geldiğinde geri dönecek ve harfin tekrarlanması sağlanacaktır. Ancak hatalı kayıtların da bulunabilmesi amacı ile grafiğin biraz daha kapsamlı hale getirilmesi gerekli görülmüştür. Daha önceki kısımlarda da belirtildiği gibi hataların nedenlerinin çoğunlukla harflerin eksik ya da fazla yazılması, harflerin birbiri yerine yazılması veya klavyede yakın olan harflerin yazılması gibi durumlar olduğu literatürde izlenmiştir. Bu nedenle de sonlu durum haritasına olası durumlar eklenerek hatalı kayıtlar saptanmaya çalışılmıştır. Harflerin eksik veya fazla yazılması ya da harflerin yerlerinin değiştirilmesi sonucunda oluşan hatalı girişlerin saptanması amacı ile grafik Şekil 16'daki gibi yeniden düzenlenmiştir.



Şekil 16: Harflerin eksik ya da fazla yazılması ile oluşan hataları saptayan grafik

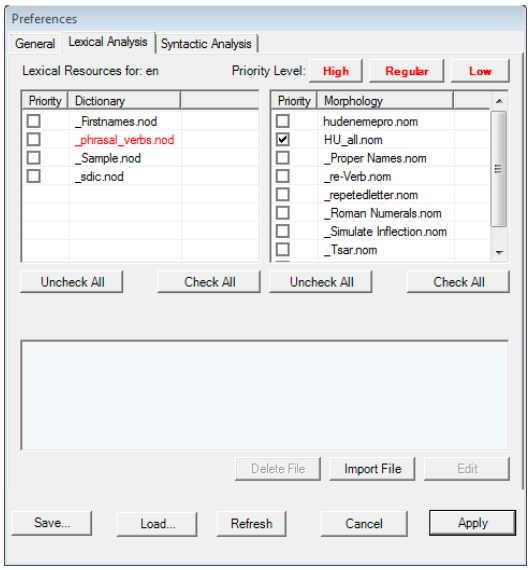
Şekil 16'da yer alan grafikte her bir harfin etrafındaki döngü ile harflerin fazla yazılması ile oluşan hataların saptanması amaçlanmıştır. Bu sayede bir kayıtta adres "hhaacceettteeppee" şeklinde verilmiş olsa bile grafik bunu saptayabilecektir. Harflerin eksik yazımını saptamak için de harfler arasında köprüler kurulmuştur. Örneğin, c ile t harfleri arasındaki köprü aracılığı ile "hacttepe" şeklinde yazılmış bir adres girişi kolaylıkla belirlenebilecektir.

Başlangıç durumu olarak ilk aşamada "h" harfi tanımlanmış olmasına rağmen ilk harfin yazılmadığı kayıtları da belirleyebilmek amacı ile sonlu durum grafiğinin başlangıç durumlarını "h" ve "a" harfleri oluşturmuştur. Klavyede yakın olan harflerin tanımlanması aşamasında her bir durumda yer alan harfin sağında ya da solunda bulunan harf o duruma eklenmiştir. *WoS*'un Türkçe karakter içermiyor olması nedeni ile Türkçe klavyede "p" harfinin yanında yer alan "ğ" harfi durumlara eklenmemiştir. Bu durumda İngilizce klavye de kontrol edilmiş "p" harfinin yanında "[" işareti bulunduğundan "p" durumuna yalnızca "o" harfi eklenmiştir. Sesli harflerde ise klavye dizilimi dikkate alınmamış diğer sesli harfler olasılık olarak tanımlanmıştır. Örneğin, "a" harfini içeren duruma "e", "e" harfini içeren duruma da "a" harfi eklenerek sesli harflerde yapılmış muhtemel bir hatanın saptanabilmesi hedeflenmiştir. Hatalı kayıtları bulması amacı ile tasarlanan gramer yapısının son şekli Şekil 17'de tanımlanmıştır.



Şekil 17: Hacettepe teriminin hatalı ve doğru girişlerinin saptanması için tasarlanan sonlu durum grafiği

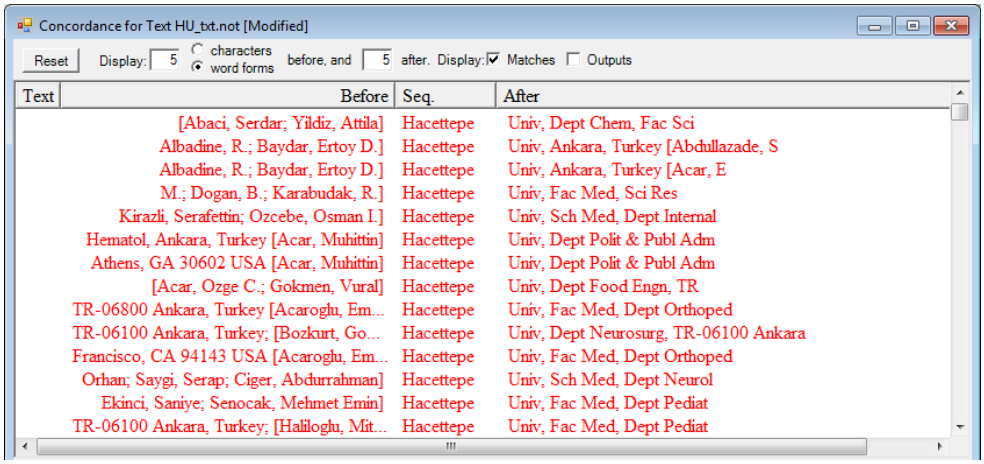
Şekil 17'de tasarlanan sonlu durum grafiğinin tüm adresleri içeren .not dosyasına uygulanması ile Hacettepe Üniversitesine ait hatalı girişlerin saptanması olanaklı hale gelmiştir. Bu uygulamayı gerçekleştirmek için öncelikle Nooj>> Info>> Preferences>> Lexical Analysis sekmesi altından gelecek ekranda yaratılan grafik işaretlenerek analizlerde bu grafiğin kullanılacağını tanımlamak gerekmektedir (bkz. Şekil 18).



Şekil 18: .nom dosyasının .not dosyası üzerinde uygulanması

Gerekli işaretlemenin yapılmasının ardından adres girişlerini içeren .not dosyası açılarak uygulama gerçekleştirilmiştir. Bunu gerçekleştirmek amacı ile Nooj >> Text >> Linguistic Analysis bağlantısı kullanılmıştır.

Morfolojik çözümleme işleminin dosya üzerinde gerçekleştirilmesinin ardından sonlu durum grafiği ile tanımlanan adresler ile eşleşen girişlerin görüntülenmesi amacı ile *unambiguous words* bağlantısı tıklanmıştır. Bu bağlantı tıklandığında açılan ekran Şekil 19'da gösterilmektedir.



Şekil 19: Tanımlanmış olası adresler ile eşleşen sözcüklerin gösterildiği bağımlı dizin (concordance)

Ortaya çıkan bağımlı dizin ile .not dosyasında yer alan kayıtların içinde yer alan 27.725 sözcüğün sonlu durum grafiği ile tanımlanan adreslere uyum sağladığı görülmüştür. Bu liste sayısal olarak değerlendirmek istendiğinde liste üzerinde sağ tıklanarak *statistical analysis* bağlantısını tıklamak yeterli olacaktır. Sonlu durum grafiği ile erişilen sözcüklerin listesi ve sıklıkları Tablo 3'te verilmiştir.

Tablo 3: Sonlu durum grafiği ile erişilen terimler ve sıklıkları

Sıra	Terim	Sıklık	Sıra	Terim	Sıklık
1	Hacettepe	21.024	9	Hacateppe	2
2	HACETTEPE	4.658	10	HACCETEPE	2
3	HACETEPPE	9	11	HACCETTEPE	2
4	Hacattepe	7	12	HACETTPE	2
5	HECETTEPE	5	13	Hacttepe	1
6	HACETTEPPE	4	14	Haccattepe	1
7	Hacetepe	3	15	Haccettepe	1
8	Hecettepe	3	16	HACETEPE	1

Çizilen sonlu durum grafiği ile .not dosyası üzerinde yer alan "Hacettepe", "Hacattepe", "Haccattepe", "HACETEPE", "HACCETEPE", "HACCETTEPE", "HECETTEPE", "HACETEPPE", "HACETTPE",

“HACETTEPPE”, “Hacateppe”, “Hacttepe” gibi yazımlar bulunabilmiştir. Nooj programı küçük harfle yaratılmış bir grafik ile büyük harfle yazılmış girişleri de saptayabilmektedir. Bu nedenle de *WoS*’da 1960’lı yıllardan 1997’ye kadar C1 ve RP alanlarında büyük harf kullanılması nedeni ile var olan büyük harfli kullanımlar da yaratılan grafik aracılığı ile saptanmıştır. Ancak çalışmanın ilerleyen bölümlerinde yer alan tablolarda ikilemeleri önlemek amacı ile büyük ve küçük harfli kullanımlar ayrı ayrı belirtilmemiş, tüm adresler küçük harfler ile sunulmuştur. Bunun yanında Hacettepe Üniversitesi yayınlarının sayısının 19.166 olmasına rağmen sıklıklar toplamında bu sayının daha yüksek olmasının nedeni bir yayında birden fazla Hacettepe Üniversitesi adresi bulunabilmesi ve “Hacettepe Univ”, “Hacettepe Med Sch” şeklinde kullanımlar olmasıdır. Bunun yanında daha önce Tablo 1’de de belirtildiği gibi Hacettepe Üniversitesi yayınlarında toplam 340 hata yapılmış olmasına rağmen Tablo 3’te hatalı yayınların sıklıklarının 43 hata ile sınırlı kalmasının temel nedeni ise hatalı yazımların farklı kombinasyonlar ile gerçekleşmiş olmasıdır. Çalışma sonunda verilen Tablo 4’te tüm hatalı kayıtların geçiş sıklığı konusunda bilgi verilmektedir.

.not dosyası üzerindeki *tokens* bağlantısı tıklatıldığında grafiğin tanımlayamadığı ancak Hacettepe’yi temsil eden diğer sözcükler görülmüştür. Bu sözcükler “Halettepe”, “Hakettepe”, “Hacehepe”, “Hacette” ve “HACETIEPE”dir. Ancak bu sözcüklerdeki hatalar yapısal olarak herhangi sabit bir nedene dayandırılamadığı için sonlu durum grafiğini daha karmaşık hale getirecektir. Aksi takdirde alfabedeki tüm harflerin her bir durum için tanımlanması gerekecektir ki, bu da doğru sonuçlara ulaşılmasını engelleyecektir.

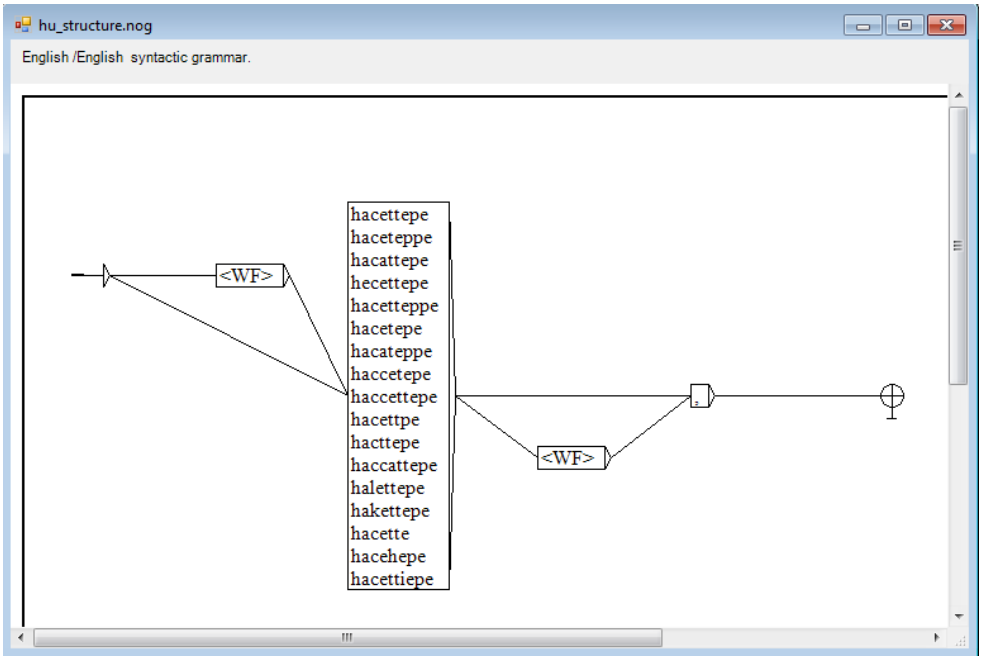
2. Aşama: Adres Yapısının Tespiti

Hacettepe adının yazımı ile ilgili hataların bulunmasını sağlayan gramerin oluşturulmasının ardından Hacettepe Üniversitesine ait adreslerin yapısının tespit edilmesi için de bir gramer yaratılmıştır. Bu gramer ile “Hacettepe

Univ" veya "Univ Hacettepe" kullanımları dışında kalan adresler da saptanabilecek ve bu sayede yayın kaybı en aza indirilebilecektir.

Standart yapının tespiti için daha önce yaratılan .not dosyası üzerinden işlemler gerçekleştirilmiştir. Programın gramer oluşturma aşamasında sunduğu iki seçenekten biri olan ve ilk aşamada yapılan işlemlerde tercih edilen *productive morphology* bir sözcüğü ek-kök olarak incelerken, *syntax* söz dizisi ya da cümleleri yapısal olarak incelemektedir. Bu nedenle de adreslerin yapısı incelenirken *syntax* tercih edilmiştir.

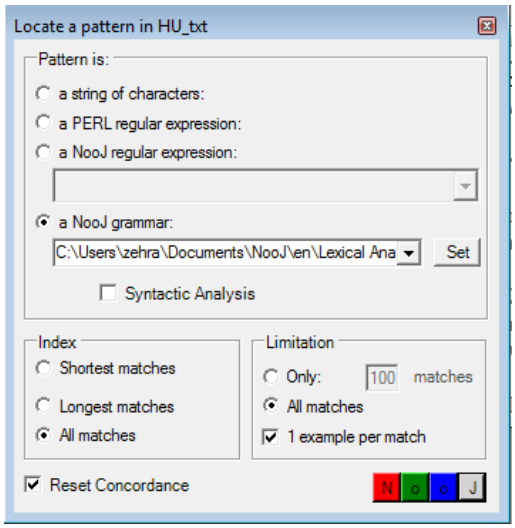
Hacettepe Üniversitesi adreslerinin standart yapısını tespit etmek amacı ile hazırlanan bu grafikte temel amaç adres alanlarında Hacettepe Üniversitesinin hangi terimler ile birleştirilerek kullanıldığını saptamak olmuştur. *WoS*'da "Univ" kısaltması kullanılarak yapılan taramalar sonucunda bazı kayıtların bu kısaltmayı içermemeleri nedeni ile gözden kaçtığı görülmüştür. Bu nedenle de Şekil 20'deki gibi bir sonlu durum grafiği ile farklı kullanımları saptamak hedeflenmiştir.



Şekil 20: Hacettepe Üniversitesinin farklı kullanımlarını saptamak üzere geliştirilen sonlu durum grafiği

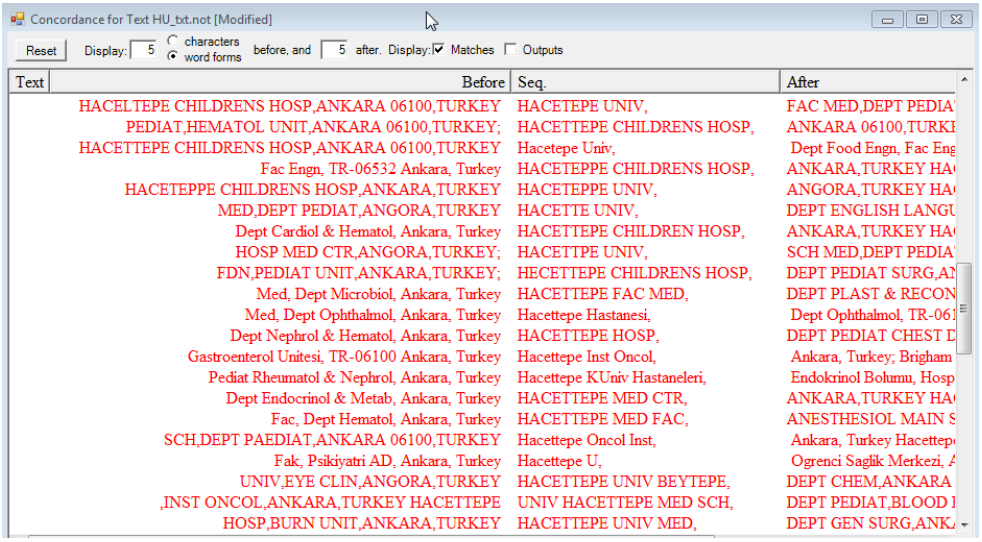
Şekil 20'de yer alan grafikte Hacettepe terimi için ilk bölümde tespit edilen farklı girişler kullanılarak yapının tespitine çalışılmıştır. Grafiğin başında ve sonunda yer alan <WF> işleçleri Hacettepe teriminden önce ya da sonra kullanılan herhangi bir terimi temsil etmektedir. *WoS* adres kayıtlarının bölüm ve fakülte isimlerinden virgül ile ayrılması nedeni ile grafik virgül ile sonlandırılmıştır.

Hacettepe Üniversitesinin farklı kullanımlarını saptamak amacı ile geliştirilen grafiğin yaratılmasında kullanılan yöntem *syntax* olduğundan bu grafiğin .not dosyasına uygulanması bir önceki uygulamadan biraz daha farklıdır. Bu grafiği metin dosyasına uygulamak için Text >> Locate bağlantısı altında çıkan alana grafik tanımlanmış ve sonuçlar listelenmiştir (bkz. Şekil 21).



Şekil 21: Syntax ile oluşturulmuş grafiğin metin dosyası üzerine uygulanması

Sonlu durum grafiğinin metin dosyasına uygulanması gramer dosyasının ve gerekli diğer seçimlerin yapılmasının ardından N simgesine tıklanması ile mümkün olmuştur. Simgeye tıklandığında açılan bağımlı dizin Şekil 22'deki gibidir.



Şekil 22: Hacettepe Üniversitesi adresleri için yaratılan bağımlı dizin

Hacettepe Üniversitesi adreslerinin verilmesinde “Univ Hacettepe” ve “Hacettepe Univ” kullanımları haricinde 47 farklı söz dizisinin kullanıldığı saptanmıştır. Bu adres yapıları Tablo 4’te gösterildiği gibidir.

Tablo 4: Dönüştürücünün saptayabildiği Hacettepe Üniversitesi adresleri

Adres	Sıklık	Adres	Sıklık
Hacettepe Univ / Univ Hacettepe	18.826	Haccetepe Fac Med	1
Hacettepe Childrens Hosp	84	Haccettepe Univ Hosp	1
Hacettepe Med Sch	60	Hacettepe Med Acad	1
Hacettepe U	33	Hacettepe Kuniv Hastaneleri	1
Hacettepe Sch Med	28	Hacehepe Univ	1
Hacettepe Med Fac	23	Haceteppe Childrens Hosp	1
Hacettepe Fac Med	12	Hacette Unıv	1
Hacettepe Hastaneleri	9	Hacettepe Technopolis	1
Hacettepe Hosp	9	Sociales Hacettepe	1
Hacettepe Hastanesi	7	Hacettepe Adult Hosp	1
Hacettepe Oncol Inst	5	Hacettepe Child Hosp	1
Haceteppe Univ	5	Hacettepe Cocuk Hastabanesi	1
Hecettepe Univ	5	Hacttepe Univ	1
Hacettepe Med Ctr	4	Hakettepe Childrens Hosp	1
Hacettepe Tip Fak	3	Halettepe Univ	1
Hacattepe Univ	3	Univ Haceteppe	1
Hacettepe Children Hosp	3	Hacettepe Cocuk Hastahanesi	1
Hecettepe Univ	3	Hacettepe Cocuk Hastanesi	1
Hacetepe Univ	2	Hacettepe Cocuk Hastenesi	1
Hacetteppe Univ	2	Hacettepe Eriskin Hastanesi	1
Hacettpe Univ	2	Hecettepe Childrens Hosp	1
Hosp Hacettepe	2	Klinikum Hacettepe	1
Hacattepe Univ Hosp	1	Hacettepe Inst Oncol	1
Haccattepe Univ	1	Unit Hacettepe	1

Program Hacettepe Üniversitesi yayınlarında kullanılmış olan adresler içinden 10 adresi saptayamamıştır. Bu adresler yalnızca birer yayında

kullanılmıştır. Dönüştürücünün saptayamadığı adreslerin listesi Tablo 5'te verilmiştir.

Tablo 5: Dönüştürücünün saptayamadığı Hacettepe Üniversitesi adresleri

Laacettepe Univ	HUTF Plast Cerrahi ABD
Ihsan Dogramaci Childrens Hosp	HU Yuksek Lisans Ogrencisi
Ibsan Dogramaci Childrens Hosp	HU Biol Dept
Ihsan Dooramaci Children Hosp	HU Egitim
HU Pharm Fac	HUTF

2007 yılında ULAKBİM'in yaptığı çalışmada Hacettepe Üniversitesi için 69 farklı adres girişi belirlenmiştir (ULAKBİM, 2007: 354-355). Ancak bu adreslerin 21'inde "Hacettepe Univ" veya "Univ Hacettepe" sorguları kullanılmıştır (Hacettepe Univ Hastaneleri, Hacettepe Univ Hastanesi, Hacettepe Univ Med gibi). Bu durumda bahsi geçen 21 kayıt yerine ilk iki sorguda kullanılan "Hacettepe Univ" ve "Univ Hacettepe" terimlerinin kullanılması yeterli olacakken olası anahtar sözcük sayısı artırılmıştır. Buna rağmen belirlenen adres girişlerinden 2007 yılından önce yayınlanmış kayıtlarda yer alan adreslerin 11'i ULAKBİM'in sağladığı listede yer almamıştır. Bunun yanında programın saptayamadığı "Ihsan Dogramaci Childrens Hosp" ve "HU Biol Dept" adresleri ULAKBİM listesinde yer almıştır.

Bu çalışma ile atıf dizinlerine hatalı girilmiş üniversite adreslerinin saptanması ile ilgili bir algoritma geliştirilmiş ve bu algoritma aracılığı ile hatalı adresten kaynaklanan yayın kaybı en aza indirilmeye çalışılmıştır. Sonuç olarak Hacettepe Üniversitesi için kullanılan adreslerden 49 tanesi saptanabilirken, 10 tanesinin saptanması mümkün olmamıştır. Ancak önemli olan sıklıkla kullanılan adres türlerinin saptanabilmiş olmasıdır. Çünkü saptanamayan 10 adres yalnızca birer kayıtta yer almıştır.

Atıf dizinlerinde üniversite adreslerinde yapılan hataların ve standardizasyon sorununun saptanması ile ilgili yapılan bu çalışma ile üniversite ismi üzerinden muhtemel adres girişlerini belirlemek olanaklı hale gelmiştir. Ancak Ege Üniversitesi ve Gazi Üniversitesi gibi üniversite

isimleri üzerine bu tekniğin uygulanmasının zor olduğu ortadadır. Bu teknik kullanılarak Gazi Üniversitesi adresleri bulunmaya çalışıldığında içinde Gazi geçen tüm kayıtlar getirilecek ve değerlendirilmesi gereken kayıt sayısı artacaktır. Bu nedenle de bu tekniğin adı kısa olmayan ve ayırt edici bir isme sahip üniversiteler için daha doğru sonuçlar vermesi beklenmelidir.

Diğer taraftan Galvez ve Moya-Anegón'un (2007b) tek bir yazar ismi örneğinden hareketle tüm yazar isimlerini saptayabilecek bir sonlu durum grafiği aracılığıyla isim saptayıcı yaratma fikrini getirdikleri çalışma üniversite adreslerine de uygulanabilir görülmüştür. Bu aşamada adreslerde yapılmış dizinleme ya da heceleme hataları bir tarafa bırakıldığında adreslerde bulunan standart yapı sorununun tamamen ortaya koyulması halinde benzer bir çalışma yürütülebilir. Örnek vermek gerekirse Türkiye'deki tıp fakültelerinin C1 veya RP alanlarında "univ" kısaltması kullanmak yerine genellikle "med sch" veya "sch med" kısaltmalarını kullandıkları saptanmıştır. Bu açıdan düşünüldüğünde yaratılan sonlu durum grafiğine tüm olasılıklar eklendiğinde herhangi bir üniversite için çalışabilecek bir sistem tasarlanmış olacaktır. Yürütülen bu çalışma genel bir adres saptayıcı yaratılması süreci ile devam ettirilebilir görünmektedir. Ancak Türk üniversitelerinin kullandıkları adres girişlerinde çeşitliliğin çok fazla olması nedeni ile bu çalışmanın uzun zaman alabileceği öngörülebilir.

4.6. TARTIŞMA VE YORUM

Bu çalışmada atıf dizinlerinde üniversite adreslerinde yapılan hatalar ve standardizasyon sorununun net bir şekilde ortaya koyulması, bu sorunun yarattığı etkilerin belirlenmesi ve hatalı adreslerin tespitini sağlayacak bir algoritma tasarlanması amaçlanmıştır. Bu kapsamda temel hipotez; "sonlu durum tekniği aracılığı ile atıf dizinlerinde yapılmış hatalar saptanabilir" şeklinde belirlenmiştir. Çalışma sonunda hipotez kanıtlanmış ve sonlu durum tekniğinin tüm üniversiteler için olmasa da, pek çok üniversite için hatalı adres girişlerini saptamaya yardımcı olabileceği anlaşılmıştır.

Yazar, editör ya da atıf dizini sağlayan kuruluş, yayının hazırlanmasından dizinlenmesine kadar geçen sürede herhangi bir noktada hata yapmış olabilir. Söz konusu hatanın hangi aşamada ve neden yapıldığının saptanması işi önemlidir. Çünkü hataların kaynağı öğrenildiğinde önlem alınması da kolaylaşacaktır. Ancak, atıf dizinlerinde üniversite adreslerinde yapılmış hataların üniversite isminde yer alan karakter sayısı, Türkçe karakter sayısı ya da sözcük sayısı gibi belirli bir nedenle doğrudan ilişkisi olmadığı sonucuna ulaşılmıştır. Bu nedenle de hata yapılmaması için önlem alınmasının hayli zor olduğu görülmektedir.

Atıf dizinlerinde yapılan hataların daha çok harf, heceleme ve dizinleme hataları ya da Türkçe kurum isimlerinin İngilizceye çevrilmesinde yaşanan problemler nedeni ile oluştuğu sonucuna ulaşılmıştır. Ayrıca kaynaklar sayısallaştırılırken gerçekleşen bilgisayar kaynaklı hataların sayısının da azımsanmayacak kadar yüksek olduğu görülmüştür.

Türk üniversiteleri adreslerinde var olan bu standardizasyon sorunu nedeni ile YÖK tarafından yayın sayısına göre yapılan üniversite sıralamalarının yanlış olduğu ve ayrıca kurumların uluslararası görünürlüğünün de standardizasyon sorunu dolayısıyla azaldığı ortaya koyulmuştur. Bunun yanında hatalı adresler bibliyometrik çalışmaların doğruluğu için de tehlike oluşturmaktadır. Doğru ve tutarlı olmayan veriler nedeni ile kurumsal hiçbir değerlendirme istenilen sonucu vermeyecektir.

Ayrıca çalışmada sonlu durum tekniği kullanılarak yaratılmış bir algoritma aracılığı ile standardizasyon sorununun en aza indirilebileceği saptanmıştır. Sonlu durum tekniğinin sağladığı avantajlar kullanılarak hem üniversite adreslerinde yapılmış harf ve heceleme hataları belirlenebilmekte hem de var olan standart yapı sorunu en aza indirilebilmektedir. Sonlu durum tekniğinin adres ve yazar adı standardizasyonunda etkili bir şekilde kullanılabileceği ortaya çıkmıştır.

Araştırma bulguları sonuç ve öneriler kısmında değerlendirilmiş ve kişi ve kurumlar için öneriler sunulmuştur.

5. BÖLÜM: SONUÇ VE ÖNERİLER

5.1. SONUÇ

Son yıllarda kurumların bilimsel araştırma sayılarına göre değerlendirilmesi yaygın olarak görülmekte ve yayınların niteliğinden çok niceliği ön plana çıkarılmaktadır. Bu nedenle de yayın sayısı dikkate alınarak değerlendirme yapılan akademik çevrelerce "çok yayını olan iyidir" görüşü hakim olmaya başlamıştır. Aynı akademik çevreler atıf dizinleri kapsamındaki yayınları diğer yayınlarla farklı yere koymaktadırlar. Bu durum atıf dizinlerinin ortaya çıkış amacının dışında kullanılmaya başlandığını işaret etmektedir.

Bu bağlamda bilimsel yayınların varlığının tespit edildiği platformlardan biri olan atıf dizinlerinde verilerin elci sistemle veri tabanına kaydediliyor olması pek çok problemi de beraberinde getirmekte ve doğru bir değerlendirme yapmayı zorlaştırmaktadır. Yayınların dizinlenmesi aşamasında yapılan hatalar nedeni ile atıf dizinlerinden verilerin doğrudan çekilmesiyle yapılmış veya yapılacak olan üniversite yayın sıralamaları konusundaki kuşkuları da beraberinde getirmektedir.

Hatalı adreslerden dolayı etkilenen niceliksel performans değerlendirmeleri bir tarafa bırakıldığında asıl büyük problemin bu gibi hatalı adresler sonucunda Türkiye'de hizmet veren üniversitelerin uluslararası platformda görünürlüklerinin azalması olduğunu söylemek yanlış olmayacaktır. Görünürlüğün yeterli etkiyi sağlayamaması Türkiye'deki üniversiteler için işbirliği fırsatlarını kaçırma anlamına dahi gelebilir. Akademik çevrede son derece önemli olan kurumlar arası işbirliklerinin artırılabilmesi amacı ile atıf dizinlerinde adres standardizasyonunun sağlanması önemlidir. Atıf dizinlerinde yapılan hataların en aza indirilmesi ile üniversiteler; işbirliği fırsatlarını kaçırmamanın yanında, halihazırda birlikte çalıştıkları kurumları da daha net bir şekilde ortaya koyabileceklerdir. İşbirliği ağlarının yaratılmasında standardize edilmiş verilerin kullanılması büyük önem taşımaktadır.

Son yıllarda Türkiye'de de üzerinde araştırma yapılan ve sonuçları kamuoyunda büyük etkiler yaratabilen bibliyometrik çalışmalarda adres standardizasyonu sağlanmadan elde edilen bilgiler değerlendirildiğinde hem çalışmalar doğru bilgileri yansıtmayacak hem de kamuoyu yanıltılmış olacaktır. Bibliyometrik araştırmaların doğru bilgileri yansıtabilmesi açısından adres standardizasyonunun düzgün bir şekilde yapılmış olması en önemli koşullardan biri olarak görülmektedir.

Bu çalışma kapsamında atıf dizinlerinde yer alan üniversite adreslerinde yapılmış hatalar saptanmış ve bu saptama işini kolaylaştıracak bir algoritma tasarlanarak standardizasyon sorununa farklı bir bakış açısı getirilmiştir. Literatürde atıf dizinlerinde adres standardizasyonu ile ilgili olarak az sayıda yayında kullanılan sonlu durum tekniğinden yararlanılarak yaratılan bu algoritma ile örnek olarak alınan Hacettepe Üniversitesi adreslerinde yapılmış hatalara ulaşılmıştır.

Çalışmanın temel hipotezi "sonlu durum tekniği aracılığı ile atıf dizinlerinde yapılmış hatalar saptanabilir" çalışma sonunda Hacettepe Üniversitesi için kanıtlanmıştır. Bu teknik birçok üniversite (Hacettepe, Uludağ, Atatürk gibi) için düzgün bir şekilde çalışmaktadır, ancak daha önce de belirtildiği gibi ismi çok kısa üniversiteler için benzer bir çalışmanın etkili bir sonuç ortaya çıkaramayacağı öngörülmektedir.

Sonlu durum tekniği ile yaratılmış adres saptayıcı algoritma aracılığı ile şimdiye kadar hiç belirlenememiş adres kayıtları da saptanabilmektedir. Bu nedenle de üniversite kütüphanelerinin yıllık yayın performanslarını raporlamaları esnasında bu teknik kullanılarak en az çaba ile en etkili sonuçlara ulaşmaları sağlanabilir. Ayrıca ilerleyen çalışmalar ile kapsam genişletilerek Türkiye'deki tüm üniversite adreslerinde bulunan standardizasyon sorununu saptayacak tek bir algoritmanın geliştirilmesi de mümkündür. Uluslararası literatürde yazar adları için benzer bir çalışma yapılmıştır.

Bu çalışmanın bir diğer önemi de adreslerde yapılan hataları ve bu hataların büyüklüklerini ortaya koyarak şimdiye kadar yapılmış veya ileride yapılacak olan üniversite yayın sayısı sıralamalarına yönelik nasıl davranılması gerektiğini ortaya koymuş olmasıdır. Atıf dizinlerinde bu denli hata yapılabiliyorken üniversite yayın sayısına dayanılarak geliştirilen politikaların da hatalı olabileceği aşikârdır. Kaldı ki üniversitelerin performanslarının niceliksel değerlendirmeler ile belirlenmesinin ne derece doğru olduğu da araştırılması gereken bir diğer konudur. Niceliksel değerlendirmelere alternatif yapıların ortaya koyulacağı çalışmalar gelecekteki araştırmaların temasını oluşturabilir.

5.2. ÖNERİLER

Atıf dizinlerinde halihazırda yapılmış hataların saptanması, ileride yapılacak hataların en aza indirilmesi ve bu hataların değerlendirilmesi ile ilgili olarak her bir kişi ya da kurumların üzerine düşen görevlerin bilincinde olması gerekmektedir. Bu çalışmada yazarlar, kütüphaneciler, politika geliştirenler ve yöneticiler gibi atıf dizinlerinde bulunan yayınlar üzerinde çalışanlar için bazı öneriler getirilerek sorunun nasıl üstesinden gelinebileceği konusunda bilgi verilmiştir.

Üniversite kütüphanelerinin hizmet verdikleri kurumun yayın istatistiklerini raporlarken adreslerde var olan bu sorunun farkında olmaları gerekmektedir. Çünkü üniversite adreslerinde tahmin edilebileceğinden daha büyük hatalar bulunabilmekte ve bu durum kurum performansının iyi resmedilememesine neden olabilmektedir. Bu durumda kütüphanelerin yapmaları gereken ilk hareket raporlama esnasında daha dikkatli davranmak olacaktır. Bu amaçla anahtar sözcükler geliştirerek yapılmış çevrimiçi taramalardan çok verilerin bilgisayara kaydedilerek değerlendirilmesi daha etkili olacaktır. Bu sayede tarama ile gözden kaçırılabilecek ya da sonuçlara fazladan eklenebilecek kayıtların engellenmesi olanaklı hale gelecek ve kütüphanecinin tüm yayınlara hakim olması sağlanacaktır.

Kütüphaneler kendi kurumlarına bağlı olarak çalışan yazarlara kurumun olası adres girişlerini sağlayarak gelecekte yapılacak yayınlarda verilebilecek hatalı adreslerin önüne geçebilirler. Hacettepe Üniversitesi bu eylemi gerçekleştirmiş ve bu uygulamayı hayata geçirebilmiştir. Bu sayede yayın sayısı çok yüksek olmasına rağmen yayınlarında yapılmış hataların oranının düşük kalmasını sağlamıştır. Ayrıca akademik çalışma yapan yazarların doğru adres vermeye teşvik edilmesi ile de yazarların bilinçlenmesi sağlanabilecektir.

Yazarlar adres verirken dikkatli davranmalıdırlar. Atıf dizinlerinde yapılan hatalar eğer dizinleme kaynaklı bir hata ise düzeltilebilmektedir.[6] Bu durumda yazarların yapması gereken şey dizinleme kaynaklı hatayı gösterip, bu sorunun giderilmesini takip etmektir. Yazarların adreslerinde atıf dizinlerinden kaynaklanan hatalar yapıldığını tespit etmesi durumunda bu hatanın düzeltilmesi amacı ile Thomson Reuters yayın düzeltme formunu[7] doldurmaları yeterli olacaktır.

Dergi **editör**lerinin adres standardizasyondaki rolü ise dergilerine gelen ve yayımlayacakları makalelerin denetimini düzgün bir şekilde yapmak ve hatalar yapılmasına fırsat bırakmamaktır. Editoryal değerlendirmesi düzgün bir şekilde yapılmış yayınlarda yazar kaynaklı hata bulunma ihtimali azalacaktır.

Adres standardizasyonu sorununun giderilmesinde **atıf dizini sağlayıcılarına** da büyük görevler düşmektedir. Çünkü bu gibi hatalar atıf dizinlerine güveni azaltmakta ve bu da sağlayıcıların itibar kaybetmelerine neden olmaktadır. Atıf dizini sağlayıcıları elci sistem ile verilerin veri tabanına aktarılması işini otomatik bir sisteme geçirmelidirler. Çünkü insan kaynaklı hataların saptanması bilgisayar kaynaklı hataların saptanmasından daha zordur. Bu nedenle de atıf dizinlerinde dizinleme sistemi öncelikle değerlendirilmelidir.

---

[6] Thomson Reuters ürün uzmanı Guillaume Rivalle ile 18 Ağustos 2010 tarihinde yapılan görüşme.
[7] http://ip-science.thomsonreuters.com/techsupport/datachange/

Her kurum için kurumsal bir kimlik numarası (InstitutionalID) belirlenmesi ve kurumlara ait yayınların bu numaralar aracılığı ile dizinlenmesi atıf dizini sağlayıcılarının yapması gereken bir numaralı görev olarak düşünülmelidir. Benzer bir çalışma *Scopus* veri tabanı için yapılmaktadır. *Scopus* veri tabanında "adres tanımlayıcı" (Affiliation Identifier) olarak adlandırılan ve her bir adrese tekil birer numara verme esasına dayanan sistem (SciVerse Scopus, 2012) adres standardizasyon sorununu azaltmada önemli rol üstlenmektedir. *WoS'*da ise yazar adı standardizasyonunu sağlamak üzere ResearcherID gibi uygulamalar yapılsa da, kurum isimleri için benzer bir çalışma yapılmamıştır. Thomson Reuters'in kurumsal kimlik numarası ile ilgili güncel bir projesi olmadığı ancak ResearcherID içine gömülü bir alan yaratılarak yazarların kurum bilgilerini bu kısma eklemelerinin sağlanabileceği konusunda bilgi olsa da, yakın zamanda böyle bir çalışma yapılmayacağı aktarılmıştır.[8] Ancak böyle bir sistem adres standardizasyon sorununu en aza indirebilecektir. Yapılan diğer tüm uygulamalar geçici çözümler sunacaktır.

Diğer taraftan atıf dizinlerinde yer alan bilgileri elektronik ortama aktaracak personelin kurumlar hakkında bilgi sahibi olması da önemli hale gelmektedir. Örneğin, Türkiye'deki kurumları tam olarak bilmeyen (çoğunlukla Türk olmayan) bir dizinlemeci Hacettepe Üniversitesi ya da İstanbul Üniversitesinden haberdar olabilir ancak daha küçük üniversiteler onun için bir anlam ifade etmeyebilir. Bu durumda sayısallaştırma aşamasında yapılmış bir hata kontrolden geçerken saptanamayacaktır. Bu nedenle atıf dizinleri için dizinleme yapacak personelin işinin uzmanı olması gerekmektedir. Ayrıca atıf dizinlerinin arama özellikleri geliştirilmeli ve daha detaylı aramalar yapmaya izin vermelidir. Bu sayede istenilen kayıtlara daha hızlı ve etkili bir şekilde ulaşılabilecektir.

Olaya farklı bir boyuttan bakılacak olursa, **yöneticiler ve politika geliştirenlerin** atıf dizinlerinde hatalar olabileceğini çok iyi kavramış

---

[8] Thomson Reuters ürün uzmanı Guillaume Rivalle ile 18 Ağustos 2010 tarihinde yapılan görüşme.

olmaları gerekmektedir. Üniversitelerin yayın sıralaması ve dolayısıyla da kimin hangi sırada olduğu hem kurumların hem de kamuoyunun merak ettiği hususlardan biridir. Öğrenciler üniversite tercihi yaparken hangi üniversitenin daha fazla yayını olduğuna bakabilmekte ve TÜBİTAK uluslararası bilimsel yayınları desteklerken yayın adreslerini dikkate alabilmektedir. Bu durumda yöneticiler ve politika geliştirenlere düşen en önemli görev kendilerinden hizmet bekleyen kişilere durumu net bir şekilde anlatmalarıdır. Çünkü toplumda bu şekilde bir bilinç oluştuğunda yönetici ve politika geliştiricilerin işi de kolaylaşacaktır. Türkiye'de halihazırda yayınların niteliğinden çok niceliğine önem veriliyor olması birkaç yıl içerisinde işe yaramayan bir yayın yığınının oluşmasına ve herhangi bir bilimsel ilerleme kaydedilmemesine neden olacaktır. Bu nedenle politika oluşturanların yayınların niteliği saptamaya yönelik çalışmalar yapmaları gerekmektedir.

Yöneticiler ve politika geliştirenlerin dikkat etmesi gereken bir diğer husus da kurum isimlerinin seçilmesidir. Çünkü bir kurumun ismi onun kimliğidir ve kurum var olduğu sürece bu isimle anılacağını gösterir. Bu nedenle isim seçilirken dikkatli davranılması gerekmektedir. Ancak, daha önce de değinildiği üzere Türkiye'de kurum isimlerinin sürekli değiştirildiği gözlenmektedir. Bu durum adres standardizasyonu sorununun boyutlarını da değiştirmektedir. Kurum isimlerinin bu kadar hızlı ve kolayca değiştirilmesi sonucunda pek çok problem ortaya çıkmaktadır. Bu nedenle kurumların kimliklerini yansıtan isimlerinin sürekli değiştirilmesinden kaçınılması gerekmektedir. Ayrıca her bir kurumun henüz kuruluş aşamasındayken yabancı dillerdeki karşılıklarının da belirlenmiş olması önemlidir. Bu sayede yayın yapan akademisyenler yayınlarında kurumlarının İngilizce karşılığını kullanmak istediklerinde sıkıntı çekmeyeceklerdir.

Tüm bunların yanında eğer yazarlara bilimsel yayın yaptıkları için teşvik verilecekse, şartların daha dikkatli belirlenmesi, adreslerde ve yazar isimlerinde yapılmış hataların göz önünde bulundurulması gerekmektedir.

Özetle adres standardizasyonu konusu pek çok kesimden kişi veya kurumla doğrudan ya da dolaylı olarak ilişkilidir. Sorunun en aza indirgenebilmesi için adres standardizasyonunun sağlanması konusunda herkes üzerine düşen dikkat ve özeni göstermeli ve konunun öneminin farkına varmalıdır.

KAYNAKÇA

Adalı, Ş. ve Erenler, Y. (2003). Türkçe için okuma fonksiyonlu otomatik metin oluşturma sistemi. *Elektrik-Elektronik ve Bilgisayar Mühendisliği 10. Ulusal Kongresi*. İstanbul, *Bildiriler* içinde s. 484-487.

Altıntaş, K. (2001). *Turkish to Crimean Tatar machine translation system*. Yayımlanmamış yüksek lisans tezi, Bilkent Üniversitesi, Ankara.

Altıntaş, K. ve Çiçekli, İ. (2001). A morphological analyser for Crimean Tatar. *10th Turkish Symposium on Artifical Intelligence and Neural Networks (TAINN'2011). KKTC, Proceedings* içinde s. 180-189.

Aronoff, M. ve Fudeman, K.A. (2005). *What is morphology?* Malden: Blackwell Publishing.

Beesley, K.R. (1996). Arabic finite-state morphological analysis and generation. *COLING '96*. Kopenhag, *Proceedings of the 16th Conference on Computational Linguistics* içinde s. 89-94. Stroudsburg, PA: Association for Computational Linguistics.

Chomsky, N. (1964). *Syntatic structures*. The Hague: Mouton de Gruyter.

Cole, J. R. (2000). A short history of the use of citations as a measure of the impact of scientific and scholarly work. *The Web of Knowledge: a festschrift in honor of Eugene Garfield* içinde s. 2w81-298. New Jersey: Information Today.

Cornell, L.L. (1982). Duplication of Japanesse names: a problem in citations and bibliographies. *Journal of the American Society for Information Science*, 33(2), 102-104.

Çiçekli, İ. (2005). Türkçe ve Kırım Tatarcası arasında bir çeviri sistemi. *Bilgisayar Destekli Dilbilim Çalıştayı.Ankara,Bildiriler* içinde s. 123-132. Ankara: TDK.

Damerau, F.J. (1964). A technique for computer detection and correction of spelling errors. *Communications of the ACM*, 7(3), 171-176.

De Bruin, R.E. ve Moed, H.F. (1990). The unification of addresses in scientific publications. *Informetrics*, (89-90), 65-78.

De Bruin, R.E. ve Moed, H.F. (1992). Delimitation of scientific subfields using cognitive words from corporate addresses in scientific publications. *Scientometrics*, 26(1), 65-80.

Duchateau, J., Wigham, M., Demuynck, K. ve Van Hamme, H. (2006). A flexible recognizer architecture in a reading tutor for children. *ITRW on Speech Recognition and Intrinsic Variation. Toulouse, Proceedings* içinde s.59-64. Toulouse: ICASSP.

French, J. C., Powell, A. L. ve Schulman, E. (2000). Using clustering strategies for creating authority files. *Journal of the American Society for Information Science*, 51(8), 774-786.

Galvez, C., Moya-Anegón, F. ve Solana, V.H. (2005). Term conflation methods in information retrieval: non-linguistic and linguistic approach. *Journal of Documentation*, 61(4), 520-547.

Galvez, C. ve Moya-Anegón, F. (2006a). The unification of institutional addresses applying parametrized finite-state graphs (P-FSG). *Scientometrics*, 69(2), 323-345.

Galvez, C. ve Moya-Anegón, F. (2006b). An evaluation of conflation accuracy using finite-state transducers. *Journal of Documentation*, 62(3), 328-349.

Galvez, C. ve Moya-Anegón, F. (2007a). Standardizing formats of corporate source data. *Scientometrics*, 70(1), 3-26.

Galvez, C. ve Moya-Anegón, F. (2007b). Approximate personal name-matching through finite-state graphs. *Journal of the American Society for Information Science and Technology*, 58(13), 1-17.

Galvez, C. ve Moya-Anegón, F. (2012). A dictionary-based approach to normalizing gene names in one domain of knowledge from the biomedical literature. *Journal of Documentation*, 68(1), 5-30.

Goldsmith, J.A. (1993). *The last phonological rule: reflections on constraints and derivations*. Chicago: University of Chicago Press.

Göksel, A. ve Kerslake, C. (2005). *Turkish: a comprehensive grammar.* New York: Taylor and Francis Group.

Grishman, R. (1986). *Computational linguistics: an introduction*. Cambridge: Cambridge University Press.

Hacettepe Üniversitesi Kütüphaneleri. (2012). İndekslerde HÜ-TR: bilimsel yayınlarda adres bilgisi. 1 Nisan 2012 tarihinde http://library.hacettepe.edu.tr/index.php?kid=105&sid=111 adresinden erişildi.

Hobbs, J. R., Douglas, A., Bear, J., Israel, D., Kameyama, M., Stickel, M. ve Tyson, M. (1996). FASTUS: a cascaded finite state transducer for extracting information from natural language text. *Finite State Devices for Natural Language Processing* içinde (s. 383-406). Editörler: E. Roche ve Y. Schabes. Cambridge: MIT Press.

Hood, W.W. ve Wilson, C.S. (2003). Informetric studies using databases: opportunities and challenges. *Scientometrics*, 58(3), 587-608.

Hopcroft, J.E., Motwani, R. ve Ullman, J.D. (2000). *Introduction to automata theory, languages and computation*. Stanford: Addison Wesley.

İlköğretim ve Eğitim Kanunu ile Bazı Kanunlarla Değişiklik Yapılmasına Dair Kanun, T.C. Resmi Gazete. (28261, 11 Nisan 2012). 12 Nisan 2012 tarihinde http://www.resmigazete.gov.tr/main.aspx?home=http://www.resmigazete.gov.tr/eskiler/2012/04/20120411.htm&main=http://www.resmigazete.gov.tr/eskiler/2012/04/20120411.htm adresinden erişildi.

İstek, Ö. (2006). *A link grammar for Turkish*. Yayımlanmamış yüksek lisans tezi, Bilkent Üniversitesi, Ankara.

İstek, Ö. ve Çiçekli, İ. (2007). A link grammar for an agglutinative language. Recent Advances in Natural Language Processing (RANLP 2007). Borovets, *Proceedings* içinde s. 285-290. New York: Springer.

Johnson, C.D. (1972). *Formal aspects of phonological description (monographs on linguistic analysis)*. The Hague: Mouton De Gruyter.

Kaplan, R.M. ve Kay, M. (1994). Regular models of phonological rule systems. *Computational Linguistics*, 20, 301-378.

Karttunen, L. ve Wittenburg, K. (1983). Two level morphological analysis of English. *Texas Linguistic Forum*, 22, 217-228.

Kettunen, K. (2008). Reductive and generative approaches to management of morphological variation of keywords in monolingual information retrieval: an overview. *Journal of Documentation*, 65(2), 267-290.

Mersinli, Ü. (2010). *Design of a finite state transducer for parts of speech tagging of Turkish.* Yayımlanmamış yüksek lisans tezi, Mersin Üniversitesi.

Milliyet. (2012). Kültür ve Turizm Bakanlığı ayrılıyor. 26 Nisan 2012 tarihinde http://siyaset.milliyet.com.tr/kultur-ve-turizm-bakanligi-ayriliyor/siyaset/siyasetdetay/02.04.2012/1522846/default.htm adresinden erişildi.

Moed, H. F. (2005). *Citation analysis in research evaluation*. Dordrecht: Springer.

Mohri, M. (1997). Finite-state transducers in language and speech processing. *Computational Linguistics*, 23(2): 269-311.

Nooj. (2012). Nooj Introduction. 30 Mart 2012 tarihinde http://www.nooj4nlp.net/pages/introduction.html adresinden erişildi.

Oflazer, K. (1994). Two-level description of Turkish morphology. *Literary and Linguistic Computing*, 9(2), 137-148.

Oflazer, K. (1996). Error-tolerant finite-state recognition with applications to morphological analysis and spelling correction. *Computational Linguistics*, 22(1), 73-89.

Oflazer, K. (2002). *Türkçe için sonlu durumlu "hafif" doğal dil çözümleyicisi ve bilgi çıkarımı uygulamasının gerçekleştirilmesi.* (TÜBİTAK Elektrik Elektronik ve Enformatik Araştırma Grubu: 199E027). Bilkent Üniversitesi ve Sabancı Üniversitesi. Ankara – İstanbul, 2002.

Orhun, M., Adalı, E., Tantuğ, A.C. (2008). Uygur dili ve makineli çeviri. *Akademik Bilişim 2008, 30 Ocak - 1 Şubat 2008, Çanakkale Onsekiz Mart Üniversitesi, Çanakkale.* 26 Nisan 2012 tarihinde http://ab.org.tr/ab08/kitap/Bildiriler/Orhon_Adali_Tantug_AB08.pdf adresinden erişildi.

Öğretim Üyeliğine Yükseltilme ve Atanma Yönetmeliğinde Değişiklik Yapılmasına Dair Yönetmelik, T.C. Resmi Gazete (26533, 26 Mayıs 2007). 26 Nisan 2011 tarihinde http://www.resmigazete.gov.tr/main.aspx?home=http://www.resmigazete.gov.tr/eskiler/2007/05/20070526.htm&main=http://www.resmigazete.gov.tr/eskiler/2007/05/20070526.htm adresinden erişildi.

Özbek, G. ve Jonathan, S. (2006). A suite of tools for augmenting English-to-Turkish statistical machine translation. *Natural Language Processing Final Project*. 25 Mart 2012 tarihinde http://infolab.stanford.edu/~jonsid/turkalator.pdf adresinden erişildi.

Piternick, A.B. (1982). Standardization of journal titles in databases. *Journal of the American Society for Information Science*, 33(2), 105.

Roche, E. ve Schabes, Y. (1995). Deterministic part-of-speech tagging with finite-state transducers. *Computational Linguistics*, 21(2), 227-253.

Roche, E. ve Schabes, Y. (1997). *Finite-state language processing (language, speech and communication).* Cambridge: The MIT Press.

Ruiz-Peérez, R., López-Cózar, E.D., Jimėnez-Contreras, E. (2002). Spanish personal name variations in national and international biomedical databases: implications for information retrieval and bibliometric studies. *Journal of the Medical Library Association*, 90(4), 411-430.

Scholl, M.H. (2008). (Some) Formal foundations of modelling dynamics. 27 Mart 2012 tarihinde http://www.informatik.uni-konstanz.de/fileadmin/dbis/im08/4-Dynamik-B-4.pdf adresinden erişildi.

SciVerse Scopus. (2012). Affiliation identifier. 26 Nisan 2012 tarihinde http://help.scopus.com/flare/schelp_Left.htm#CSHID=h_affilsrchtips.htm|StartTopic=Content%2Fh_affilsrchtips.htm|SkinName=svs_SC adresinden erişildi.

Silbertzein, M. (2002). NooJ: an object-oriented approach. *5th Intex Workshop, Mayıs 2002, Marseille*. 26 Nisan 2012 tarihinde http://www.nooj4nlp.net/NooJArchitecture.pdf adresinden erişildi.

Silbertzein, M. (2003). Nooj Manual. 30 Mart 2012 tarihinde http://www.nooj4nlp.net/NooJManual.pdf adresinden erişildi.

Sleator, D.D.K. ve Temperley, D. (1991). Parsing English with a link grammar. *Third International Workshop on Parsing Technologies.* 20 Mart 2012 tarihinde http://www.link.cs.cmu.edu/link/ftp-site/link-grammar/LG-IWPT93.pdf adresinden erişildi.

Soysal, E., Çiçekli, İ. ve Baykal, N. (2009). Radyoloji raporları için Türkçe bilgi çıkarım sistemi. *6. Ulusal Tıp Bilişimi Kongresi*. *Antalya, Bildiriler*. 26 Nisan 2012 tarihinde http://web.cs.hacettepe.edu.tr/~ilyas/PDF/2009_UlusalTipBilisimKongresi.pdf adresinden erişildi.

Stefaniak, B. (1987). Use of bibliographic data bases for scientometric studies. *Scientometrics*, 12(3-4), 149-161.

Thomson Reuters. (2009). Web of Science 8.0. 2 Nisan 2012 tarihinde http://thomsonreuters.com/content/science/pdf/ssr/training/wos_workbook_en.pdf adresinden erişildi.

Toutkoushian, R.K. ve Webber, K. (2011). Measuring the research performance of post-secondary institutions. *University Rankings: theoretical basis, methodology and impacts on global higher education* içinde s. 123-144. New York: Springer.

Ui Dhonnchadha, E. (2003). Finite state morphology and Irish. *10th Conference of the European Chapter of the Association for Computational Linguistics*. Budapeşte, *Proceedings of the Workshop on Finite-State Methods in Natural Language Processing* içinde s.43-49. East Stroudsburg: Association for Computational Linguistics.

ULAKBİM. (2007). *Türkiye bilimsel yayın göstergeleri*. Ed. İ.H. Demirel, C. Saraç ve E.A. Gürses. Ankara: ULAKBİM.

ULAKBİM. (2010). TÜBİTAK Uluslararası bilimsel yayınları teşvik programı uygulama esasları. 3 Nisan 2012 tarihinde http://www.ulakbim.gov.tr/cabim/ubyt/esaslar_v2.pdf adresinden erişildi.

URAP (University Ranking by Academic Performance). (2011). Genel Bilgi. 13 Nisan 2012 tarihinde http://tr.urapcenter.org/2011_2/2011_7.php adresinden erişildi.

Web of Science. (2009). Web of Science field tags (articles and conference proceedings). 26 Nisan 2012 tarihinde http://images.webofknowledge.com/WOK46/help/WOS/h_advanced_fieldtags.html adresinden erişildi.

Williams, M.E. ve Lannom, L. (1981). Lack of standardization of the journal title element in databases. *Journal of the American Society for Information Science*, 32(3), 229-233.

Van Raan, A.F.J. (2005). Fatal attraction: conceptual and methodological problems in the ranking of universities by bibliometric methods. *Scientometrics*, 62(1), 133-143.

Yükseköğretim Kurulu. (2010). 2010 yılı yayın istatistikleri. 2 Nisan 2012 tarihinde http://www.yok.gov.tr/index.php?option=com_docman&task=doc_download&gid=728 adresinden erişildi.

EK

Ek 1. En çok yayını olan ilk 20 üniversite için doğru kabul edilen adres girişleri

Üniversite	Doğru kabul edilen adresler
Hacettepe Üniversitesi	Hacettepe Univ
	Univ Hacettepe
İstanbul Üniversitesi	Istanbul Univ
	Univ Istanbul
Ankara Üniversitesi	Ankara Univ
	Univ Ankara
ODTÜ	Middle E Tech Univ
	Middle East Tech Univ
	Orta Doğu Teknik Univ
	Middle East Technol Univ
	Orta Dogu Tekn Univ
	METU
	ODTU
Ege Üniversitesi	Ege Univ
	Univ Ege
Gazi Üniversitesi	Gazi Univ
	Univ Gazi
İstanbul Teknik Üniversitesi	Istanbul Tech Univ
	Istanbul Technical Univ
	Tech Univ Istanbul
	Istanbul Teknik Univ
	Istanbul Tekn Univ
Dokuz Eylül Üniversitesi	Dokuz Eylul Univ
	Univ Dokuz Eylul
Atatürk Üniversitesi	Ataturk Univ
	Univ Ataturk

GATA	GATA
	Gulhane Mil Med Acad
	GMMA
	Gulhane Askeri Tip Akademisi
Marmara Üniversitesi	Marmara Univ
	Univ Marmara
	Cukurova Univ
Çukurova Üniversitesi	Univ Cukurova
Erciyes Üniversitesi	Erciyes Univ
	Univ Erciyes
Ondokuz Mayıs Üniversitesi	Ondokuz Mayis Univ
	Univ Ondokuz Mayis
Başkent Üniversitesi	Baskent Univ
	Univ Baskent
Selçuk Üniversitesi	Selcuk Univ
	Univ Selcuk
Boğaziçi Üniversitesi	Bogazici Univ
	Univ Bogazici
Fırat Üniversitesi	Firat Univ
	Univ Firat
Karadeniz Teknik Üniversitesi	Karadeniz Tech Univ
	Tech Univ Karadeniz
	Karadeniz Tekn Univ
	Karadeniz Teknik Univ
	Karadeniz Technical Univ
Bilkent Üniversitesi	Bilkent Univ
	Univ Bilkent