

Teknoloji Alanında Faaliyet Gösteren Şirketlerin Finansal Form Okunabilirliğinin Metin Madenciliği Teknikleriyle Ölçülmesi

Measuring Financial Form Readability of Technology Companies with Text Mining Techniques

M.Fevzi Esen¹ , Tutku Tuncalı Yaman² 



ÖZ

Finansal piyasalarda işlem gören şirketlerin yükümlülükleri kapsamında bulunan ve işletme operasyonları ile işletmelerin finansal durumlarına ilişkin kapsamlı bilgi sağlayan kamuya açıklama formlarının (Disclosure forms) yatırımcılar tarafından anlaşılabilirliği için, karmaşık olmayan ve anlaşılabilir bir dil kullanılması gerekmektedir. Metinlerin söz dizimsel uzunluğu ve karmaşıklığının tespitinde kullanılan okunabilirlik indeksleri, uygulamalı dilbilim, metin ve söylem analizi ile doğal dil işleme gibi alanlarla yakından ilişkili olup, veri kaynağı olarak metin üzerinden anlamlı bilgi edinilmesini amaçlayan metin madenciliği fonksiyonlarından biridir. Bu çalışmada, 2010 – 2019 yılları arasında New York Borsası'nda (NYSE) işlem gören ve teknoloji alanında faaliyet gösteren 144 şirkete ilişkin farklı türlerde toplam 23.103 adet kamuya açıklama formu kullanılarak, her bir form türüne ilişkin okunabilirlik skorları hesaplanmıştır. Alternatif okunabilirlik indeksleriyle elde edilen skorlar karşılaştırılarak şirket ölçeklerine göre ortalama skorlar değerlendirilmiştir. Çalışma sonucunda, dosya boyutu haricindeki indekslerin birbirine yakın sonuçlar verdiği tespit edilmiş olup, incelenen kamuya açıklama formlarının, lisansüstü eğitime sahip yatırımcıların seviyesinden de yüksek skorda, “çok zor” okunabilirlik düzeyinde olduğu; orta ve büyük ölçekteki teknoloji şirketlerince raporlanan formların ise daha düşük okunabilirlik düzeyine sahip olduğu sonucuna ulaşılmıştır.

Anahtar kelimeler: Okunabilirlik İndeksi, Finansal Piyasalar, Metin Madenciliği

ABSTRACT

In financial markets, companies are required to disclose a comprehensive review of their business operations and financial positions via filing forms. For these forms to be understood by investors, it is necessary to use plain language without complexity. Readability indexes are used to determine the syntactic length and complexity of the texts, and they are closely related to fields such as applied linguistics, text analysis, and natural language processing. These indexes are one of the text mining functions, which aims to obtain meaningful information from texts as a primary data source. In this study, we calculated the readability scores of a total of 23,103 public disclosure forms of 144 technology companies on the New York Stock Exchange (NYSE) between 2010 - 2019. The scores were evaluated by comparing alternative readability indexes. As a result of the study, it was determined that readability indexes, except for file size, provide similar results, and it was concluded that public disclosure forms are “very difficult” to understand and require higher educational levels than a graduate education. It was also shown that the forms of medium and large-sized technology companies have lower readability levels than other technology companies.

Keywords: Readability Index, Financial Markets, Text Mining

¹Dr. Öğr. Üyesi, Sağlık Bilimleri Üniversitesi, İstanbul, Türkiye

²Dr. Öğr. Üyesi, Beykent Üniversitesi, İktisadi ve İdari Bilimler Fakültesi, Lojistik Yönetimi, İstanbul, Türkiye

ORCID: M.F.E. 0000-0001-7823-0883;
T.T.Y. 0000-0001-8742-2625

Corresponding author:
M. Fevzi ESEN
Sağlık Bilimleri Üniversitesi, İstanbul, Türkiye
E-mail address: fevzi.esen@sbu.edu.tr

Submitted: 24.10.2020
Revision Requested: 13.11.2020
Last Revision Received: 16.12.2020
Accepted: 16.12.2020

Citation: Esen, M. F. ve Tuncalı Yaman, T. (2020). Teknoloji alanında faaliyet gösteren şirketlerin finansal form okunabilirliğinin metin madenciliği teknikleriyle ölçülmesi. *Acta Infologica*, 4(2), 77-97.
<https://doi.org/10.26650/acin.815945>

1. GİRİŞ

Günümüzde yazılı ve görsel veri kaynaklarındaki artış, pek çok türde ve uzantıda enformasyonun üretilmesine sebep olmaktadır. Yazılı kaynakların hem basılı hem de elektronik formatta mevcut olmasından dolayı, insanlar pek çok kararı söz konusu kaynaklardan edindikleri bilgiye dayanarak vermektedir. Bu bağlamda, farklı konularda karar alırken gerekli bilginin edinilmesinde yararlanılan metinlerin kolay okunabilirliği, diğer bir ifadeyle, anlaşılabilirliği konusu dikkat çekmektedir. Özellikle Amerika Birleşik Devletleri'nde (ABD) uzun yıllardır ele alınan okunabilirlik düzeyi ölçümlemesi sadece istatistiksel araçlarla metnin karmaşıklığının ölçülmesinden ibaret olmayıp, hazırlanan metnin okunabilirlik düzeyinin; araştırmacılar, eğitimciler, dil bilimciler, yönetim, sağlık ve teknik alanlarda bilgilendirici doküman hazırlayanlar gibi tüm uzmanları bağlayıcı bir nitelik taşımaktadır. Teorik olarak uygulamalı dil bilimi, doğal dil işleme algoritmaları veyahut metin madenciliği alanında ele alınan okunabilirlik kavramı, pratikte her alanda doğru anlaşılabilme için doğru yazılı materyalin seçilmesinde, farklı alanlarda uzman olan veya olmayan kişilere bilgi aktarılmasında ve yazılı teknik metinler için bir standart oluşturulmasında kullanılmaktadır (Bailin ve Grafstein, 2016).

Finansal piyasalarda yatırımcıların özel durumlar hakkında bilgilendirilmesine yönelik yapılan kamuya açıklama formlarında ve finansal raporlamada, metinlerin ortalama eğitim düzeyine sahip bir yatırımcı tarafından anlaşılabilir olması, piyasalardaki bilgi akışını etkilemekte ve etkin piyasa oluşumuna katkı sağlamaktadır (Lee, 2012). Özellikle menkul kıymetleri piyasada işlem gören şirketler tarafından kamuya açıklanan finansal raporların ve formların okunabilirliği; şirket performansı, finansal planlama etkinliği, yatırımcı kararları ve yatırımlarının verimliliği, kurumsal yönetim, karlılık durumu ve faaliyet etkinliği, likidite ve beklenen getiri ile bilgi asimetrisi gibi birçok faktörle yakından ilişkili olup, piyasanın ve yatırımcıların korunması açısından önemli bir yere sahiptir (Loughran ve McDonald, 2011; Lang ve Stice-Lawrence, 2015; Hasan ve Habib, 2020). Şirketlerin kamuyu bilgilendirirken açıkladığı formların anlaşılabilirliğinin finansal alana özgü terimlere ağırlık verilerek, karmaşık veya uzun cümleler kullanılarak düşürülmesi piyasada bilgi asimetrisi yaratmakta ve yatırımcıların finansal varlıklarını yanlış yönetmesi probleminde de sebep olmaktadır (You ve Zhang, 2009; Bai vd., 2018).

Finansal formların ve raporların düzenlenmesinde uzun ve karmaşık olmayan cümlelerle hazırlanmış, anlaşılır bir metnin kullanılması koşulu, ABD Menkul Kıymetler ve Borsa Komisyonunca (SEC) düzenlenip yürürlüğe konmuştur. Buna göre yayınlanan kılavuzda, finansal formların raporlanmasında kullanılacak esaslar belirlenmiş olup, yalın bir dil kullanılmasının gerekliliği belirtilmiştir (SEC, 1998). Söz konusu kılavuzda ayrıca, kamuya açıklanacak finansal rapor ve formların mümkün olan en fazla sayıda ve farklı profilde yatırımcı tarafından da anlaşılabilmesi için yalın dilin nasıl kullanılması gerektiğine ilişkin dikkat edilmesi gereken standartlar sıralanmaktadır. Bu bağlamda, açıklanan formlardaki metinlerin içeriğindeki sözcüklerin anlaşılmasının zorluğu, söz dizimsel uzunluğu ve karmaşıklığı gibi dilsel özellikleri, metnin okunabilirlik düzeyinin tespit edilmesinde kullanılmaktadır.

Bu çalışmada, ABD'de çeşitli normlarla da düzenlenen okunabilirlik düzeyi aracılığıyla yatırımcıların finansal kararlar alırken başvurdukları kamuyu bilgilendirici formların okunabilirliğinin çeşitli alternatif indekslerle tespit edilmesi amaçlanmıştır. Ayrıca çalışmada, şirket büyüklüklerine göre her bir form türünün okunabilirlik skorlarının farklılaşp farklılaşmadığı da incelenmiştir. Şirket değerleri göz önünde bulundurularak, özellikle bireysel yatırımcının ilgisini çeken teknoloji şirketleri (NY Times, 2019) çalışmanın örnekleme dahil edilmiştir. Bu ilgiyi ortaya koyan göstergelerden biri olan Finviz (2020) raporunda ele alınan ve son beş yıl için teknoloji şirketlerine ilişkin finansal performans göstergelerine bakıldığında, teknolojinin tüm alt sektörlerinde son 5 yıldaki satış büyüme oranları pozitif (ortalama: %12,89) iken, gelecek 5 yıla ilişkin hisse başı kazanç oranları da önceki 5 yıla göre (ortalama: %13,02) artış eğilimindedir. Ayrıca teknoloji şirketlerinin toplam piyasa değeri 2020 yılı itibarıyla 9,5 trilyon dolar olup, yazılım – altyapı, tüketici elektroniği, yarı iletkenler, yazılım – uygulama, bilgi teknolojileri, iletişim malzemeleri, yarı iletken ekipmanları ve donanım alt sektörlerinin sahip olduğu yüksek fiyat-kazanç oranı, gelecek dönemde yüksek nakit akışına işaret etmektedir (NYSE, 2020).

Yükselen piyasa değerine paralel olarak, yatırımcıların yüksek kazanç beklentisi ile yatırım yaptıkları teknoloji şirketlerinin SEC düzenlemelerine bağlı olarak yayınlamakla yükümlü oldukları formlar (8-K, 10-K, 10-Q, Form 3 ve Form 4) bulunmaktadır. Gerek bireysel gerekse kurumsal yatırımcılar, yatırım kararları öncesinde söz konusu formlara başvurmakta olup, yatırım

kararlarını formlar aracılığıyla açıklanan bilgi ve göstergelere göre almaktadır. Bu bağlam göz önünde bulundurularak çalışmanın araştırma sorusu, “NYSE’de işlem gören teknoloji şirketlerinin 2010-2019 döneminde yayınladıkları 8-K, 10-K, 10-Q, Form 3 ve Form 4 kamuya açıklama formları, SEC’in belirlediği standartlar doğrultusunda yatırımcılar tarafından anlaşılabilir bir düzeyde midir ve söz konusu formların okunabilirlik skorlarının yayınladıkları şirket ölçeğine göre farklılık göstermekte midir?” şeklinde ifade edilebilir. Araştırma sorusunun cevaplanması adına gerçekleştirilen uygulamada, 144 şirkete ait toplam 23.103 adet kamuya açıklama formunun okunabilirlik düzeyleri hakkında karşılaştırmalı bir değerlendirme yapabilmek adına, Loughran ve McDonald (2014) ile Zhou, Jeong ve Green (2017) çalışmalarında belirtilen metodoloji referans alınmıştır.

SEC’in finansal dokümanların oluşturulmasında kamu yararı gözeterek okunabilirliği belirli bir düzeye taşıyabilmek için yayınladığı “Sade İngilizce Kuralı” (The Plain English Rule) göz önünde bulundurulduğunda, yatırımcıya sunulan güncel finansal doküman, form ve raporların okunabilirlik düzeyinin incelenmesi önemlidir. Okunabilirlik kavramı Türkçe yazında da kullanılan bir kavram olmakla birlikte, söz konusu literatür tarandığında (bkz. Bölüm 3) bu kavramın daha çok eğitim bilimleri alanına odaklanan araştırmalarda kullanıldığı görülmektedir. Finansal dokümanların okunabilirliği hakkında ise, SEC örneğine benzer şekilde bir kuralın Türkiye’deki düzenleyici kurumlar tarafından ortaya konmuş olmamasına rağmen, finansal rapor ve formların okunabilirliği konusunun Türkiye’deki yatırımcıların korunması bağlamında da değerlendirilebilir olması, Türkiye’deki piyasalar için oluşturulabilecek Türkçe finansal dokümanlarda kullanılan dilin okunabilirliğinin artırılması ve özellikle kamuya açıklama formlarının yatırımcılar tarafından anlaşılabilir hale getirilmesi gerekmektedir.

Mevcut durumda Türkiye piyasalarında NYSE benzeri bir uygulama olmadığı bilinmekle beraber, alan literatüründe de Türkçe finansal dokümanların okunabilirliği ile ilgili herhangi bir çalışma veya öneriye rastlanmamıştır. Bir anlamda okunabilirlik düzeyinin artmasının da bir sonucu olarak değerlendirilebilecek yatırımcının korunması hususunda konuyu farklı açılardan ele alan iki araştırma bulunmaktadır. Bunlardan ilki, Çakın (2010) tarafından sermaye piyasalarında yatırımcıların korunması konusunda olup, kamuya açıklanan dokümanların okunabilirliği konusuna temas edilmemektedir. Üzeler (2014) tarafından gerçekleştirilen çalışmada ise, 6502 sayılı Tüketicinin Korunması Hakkındaki Kanunda tüketici ile sözleşmelerin en az on iki punto büyüklüğünde, anlaşılabilir bir dilde, açık, sade ve okunabilir bir şekilde oluşturulmasına ilişkin hükmünün sermaye piyasası ilişkilerinde uygulama alanı bulabileceğinden bahsedilmektedir. Her iki çalışmada da okunabilirlik kavramına ve bunun uygulamada nasıl sağlanacağı ilişkin net bir ifadeye yer verilmemiştir. Bu açıdan değerlendirildiğinde, çalışmamızın Türkçe literatürde metin madenciligi ve okunabilirliği tanıtıcı bir özelliği barındırması açısından araştırmacılara katkı sağlaması beklenmektedir. Ayrıca, dünya piyasalarında işlem gören ve yukarıda belirtildiği üzere en yüksek piyasa değerine sahip olan teknoloji sektöründe faaliyet gösteren işletmelere ilişkin kamuya açıklama formlarının okunabilirliği üzerine yabancı literatürde de yapılan bir çalışmaya rastlanmamış olması, çalışmanın literatüre katkısını artıran bir diğer neden olarak görülmektedir.

Çalışmanın ilk bölümünde okunabilirlik kavramı, okunabilirlik indeksleri ve bu indekslerin hesaplama adımları hakkında detaylı bilgiye yer verilmiştir. Ardından okunabilirlik indeksleri konusunda yapılmış çalışmalara ilişkin literatür taraması yapılmıştır. Üçüncü bölümde ise çalışma kapsamında kullanılan veri seti hakkında bilgi verilmiş, farklı analiz teknikleri detaylandırılmıştır. Uygulama bölümünde ise teknoloji alanında faaliyet gösteren ve NYSE’ye kote olmuş 144 şirkete ilişkin farklı kamuya açıklama formlarının okunabilirlikleri çeşitli indeksler kullanılarak analiz edilmiş, bu analizlerden elde edilen bulgulara yer verilmiştir. Çalışmanın son bölümünde ise, elde edilen bulgulara dayalı olarak değerlendirme ve önerilerde bulunulmuştur.

2. OKUNABİLİRLİK İNDEKSLERİ

Okunabilirlik kavramı, metinlerin okunaklı olması ve anlaşılabilirliği ile ilişkilendirilmekte olup, genel olarak bir metnin okuyucular tarafından hızlı, kolay ve anlaşılabilir şekilde okunabilmesi olarak tanımlanmaktadır (Klare, 1984). Bir önceki bölümde de değinildiği üzere, literatürde okunabilirlik konusunda yapılan teorik ve ampirik çalışmalarda, metnin biçimsel özelliklerinin metnin anlaşılabilirliği üzerindeki etkisi tartışılmış, birçok çalışmada da okuyucunun dikkatini metne çekebilmek için geliştirilebilecek tasarımlara ve tipografik unsurların kullanımına dikkat çekilmiştir (Di Marco, 2010;

Lupton, 2014). Metinlerin okuyucuya uygunluğu ve okurlar tarafından anlaşılabilirliği olan okunabilirlik kavramı, birçok araştırmada metnin dış yapı özelliklerini oluşturan harf büyüklüğü ve stili, vurgular, sayfa düzeni, çözünürlük, özel efektler ve renkler gibi okunaklılığı sağlayan unsurlardan ayrı olarak incelenmektedir (Bailin ve Grafstein, 2016). Metni oluşturan kelimelerin karmaşıklığı olarak da ifade edilen okunabilirlik, kullanılan dildeki yeni kelimelerin sayısı ve bunların kavranmasındaki kolaylık olarak da nitelendirilmektedir (Chandler, 2001).

Okunabilirlik üzerine yapılan matematiksel formül çalışmalarında, metnin yapısal zorluğunun belirlenmesinde tümce ve sözcük özelliklerinin kullanılarak niceliksel bir değerlendirmenin gerçekleştirilmesi amaçlanmaktadır. Buna göre, istatistiksel geçerliliği kanıtlanmış pek çok formül geliştirilmiş olup, üretilen bilgi ve belgelerin okunabilirlik düzeylerinin niceliksel bir değerlendirmeye kolay ve hızlı şekilde tespit edilmesi hedeflenmiştir (Bailin ve Grafstein, 2016).

Metinlerin okunabilirlik düzeylerinin tespiti için Flesch (1948) tarafından yapılan çalışmada, metin içerisindeki kelime ve cümle uzunluğunun metnin anlaşılabilirliği ile ilişkili olup olmadığı araştırılmıştır. Çalışmada, bir metinden seçilen 100 kelimelik bir kesit için kelime ve cümle uzunlukları hesaplanmaktadır. Kelime uzunluğunu (wl) belirlemek için toplam hece sayısı toplam kelime sayısına bölünmekte; cümle uzunluğunun (sl) tespiti için ise toplam kelime sayısı toplam cümle sayısına bölünmektedir. Buna göre, Flesch okuma kolaylığı (FRE) formülü aşağıdaki şekilde ifade edilmektedir (Flesch, 1948):

$$FRE=206,835-(0,846 \times wl)-(1,015 \times sl).$$

Hesaplama sonucunda elde edilen okuma kolaylığı skoru, eğitim düzeyi puanını belirlemek için Kincaid vd. (1975) tarafından yeniden uyarlanarak metinlerin tahmini sınıf seviyesindeki karşılığı elde edilmiştir. Uyarlanan Flesch – Kincaid İndeksi (FKI) şu şekilde formüle edilmektedir:

$$FKI=(0,39 \times sl)+(11,8 \times wl)-15.59.$$

Formül sonucunda elde edilen skor, Amerikan eğitim seviyesini yani, ilgili metni anlayabilmek için gerekli olan eğitim yılını vermektedir. Örneğin, işlem sonucunda elde edilen puan 13,2 ise metni okuyan kişinin metni anlayabilmesi için ortalama 13 yıl eğitim almış olması gerekmektedir.

Ortalama cümle uzunluğu ve metin içerisindeki 3 ve daha fazla heceli (multisyllabic- çok heceli veya kompleks) kelime oranı değişkenlerini temel alarak oluşturulan SMOG (Simple Measure of Gobbledygook) okunabilirlik indeksi McLaughlin (1969) tarafından ortaya atılmıştır. Buna göre diğer indekslerden farklı olarak metnin başından, ortasından ve sonundan alınan cümle örneklemelerinin, metnin anlamsal ve söz dizimsel etkileşiminden kaynaklanan olası problemleri elimine edebileceği savunulmaktadır. Çalışmada, bir metindeki toplam kelime sayısının cümle uzunluğu konusunda fikir verdiği ve bu durumun metin içerisindeki karmaşık kelime sayısı ile ilişkili olabileceği belirtilmektedir. Ayrıca, bir metindeki cümle uzunluğu arttıkça kelime başına düşen ortalama hece sayısının da artacağı belirtilmektedir (McLaughlin, 1969). SMOG skoru hesaplanırken en az 30 cümleden oluşan bir metin seçilerek, metinde bulunan çok heceli kelime sayısı tespit edilmektedir. Sonrasında, okuyucunun metni anlayabilmek için sahip olması gereken eğitim yılının bulunabilmesi için aşağıdaki denklem kullanılarak okunabilirlik skoru hesaplanmaktadır:

$$SMOG = 1,0430 \times \sqrt{\text{çok heceli kelime sayısı} \times \frac{30}{\text{cümle sayısı}}} + 3,1291.$$

Gunning (1952) tarafından kitap, dergi ve gazete yazarlarını desteklemek için geliştirilen indekste cümle karmaşıklığını ölçmek için ortalama cümle uzunluğu, kullanılan kelimelerin zorluğunu ölçmek için ise çok heceli kelimelerin sayısı değerlendirilmeye alınmıştır. Gunning-Fog (GF) indeksinde, kısa cümlelerin okunabilirliğinin uzun cümlelere göre daha kolay olduğu yaklaşımı benimsenmektedir. Hecelerin değerlendirilmesinde özel adlar, hece çizgisi olan ve İngilizce -s, -es, -er, -ed, -ing takılarına sahip olan kelimeler çok heceli olarak kabul edilmemektedir (Gunning, 1952). GF İndeksi aşağıdaki şekilde formüle edilmektedir:

$$GF = 0,4 + \left[\left(\frac{\text{kelime sayısı}}{\text{cümle sayısı}} \right) + \left(100 \times \left(\frac{\text{çok heceli kelime sayısı}}{\text{kelime sayısı}} \right) \right) \right].$$

Amerikan eğitim seviyesine göre skor üreten bir diğer okunabilirlik indeksi, Coleman ve Liau (1975) tarafından geliştirilmiştir. Söz konusu indeks diğerlerinden farklı olarak, kelime başına düşen hece sayısı yerine kelimelerdeki karakter sayısına dayalı hesaplanmaktadır. Metin içerisindeki karakterler bilgisayar programlarıyla çok heceli kelimelere göre daha kolay sayılabildiği için, okunabilirlik skoru daha hızlı elde edilebilmektedir (Coleman ve Liau, 1975). Coleman-Liau İndeksi (CLI) aşağıdaki formülle hesaplanmaktadır:

$$CLI=(0,0588 \times l)-(0,296 \times s)-15,8.$$

Formüle ifade edilen l değişkeni 100 kelimelik kesitteki ortalama karakter sayısını, s değişkeni ise 100 kelimelik kesitteki ortalama cümle sayısını belirtmektedir.

Smith ve Senter (1967) tarafından önerilen Otomatik Okunabilirlik İndeksi (ARI), hesaplamada metnin içerisindeki ortalama kelime uzunluğunun kaç karakterden oluştuğunu göz önünde bulundurmaktadır. Karakterlerin ve kelimelerin elektronik olarak sayımı, cümle ve paragrafların sayımından daha hızlı olduğundan, okunabilirlik skorunun gerçek zamanlı tespitinde söz konusu indeks sıklıkla kullanılmaktadır. ARI skoru aşağıdaki gibi hesaplanmaktadır:

$$ARI = 4,71 \times \left(\frac{\text{karakter sayısı}}{\text{kelime sayısı}} \right) + 0,5 \times \left(\frac{\text{kelime sayısı}}{\text{cümle sayısı}} \right) - 21,43.$$

Belirtilen indeks skorlarına karşılık gelen Amerika Birleşik Devletlerindeki eğitim ve okunabilirlik seviyeleri Tablo 2'deki gibi verilmiştir. Buna göre, skoru 8-9 olan bir metnin standart düzeyde okunabilirlik seviyesine sahip olduğu ve ortalama 9. sınıf ve üzerindeki eğitim seviyesine sahip kişilerin anlayabileceği nitelikte olduğu söylenebilir. 17 ve üzerinde skora sahip metinlerin okunabilirlik düzeyi çok zor olmakla birlikte, lisansüstü seviyede bir eğitim düzeyi gerektirmektedir.

Tablo 2

Eğitim ve Okunabilirlik Seviyelerine Göre İndeks Skorları

İndeks Skoru	Okunabilirlik Düzeyi	Eğitim Düzeyi
0-5	Çok kolay	İlkokul
6	Kolay	6.sınıf
7	Biraz kolay	7. sınıf
8-9	Standart düzeyde	8. ve 9. sınıf
10-12	Biraz zor	10-12. sınıf
13-16	Zor	Üniversite (lisans)
≥17	Çok zor	Yüksek lisans ve üstü

Kaynak: DuBay (2004)

3. LİTERATÜR ARAŞTIRMASI

Okunabilirlik kavramının çeşitli kaynaklardaki tanımına bakıldığında genel olarak herhangi bir metnin farklı düzeydeki okuyucu(lar) tarafından dile ilişkin bağlamlar çerçevesinde anlaşılabilirliği olarak açıklandığı görülmektedir. Çoğunlukla ilişkili olduğu “anlaşılabilirlik” ile karıştırılan okunabilirlik, metnin karakteri ile ilgili olarak anlaşılabilirlik, okunan metnin anlamlandırılacak okuyucunun seviyesi ile doğrudan bağlantılıdır (Güyer vd., 2009: 758). Klare (1963) dilbilimsel bakış açısıyla yaptığı tanımda okunabilirliği anlaşılabilirlik ile harmanlayarak “yazım tarzından dolayı metnin okunması ve anlaşılmasının kolaylığı” olarak ifade etmiştir. DuBay’e (2004) göre ise okunabilirlik bir metnin diğerlerinden daha kolay okunmasıdır ancak bu tanım yazı biçimi ve düzen ile ilişkili olan “okunaklılık” ile karıştırılmaktadır. Buna göre okunabilirlik, bir metnin okuyucu tarafından anlaşılma derecesi olarak ifade edilirken, anlaşılabilirlik ise okuyucunun metni okuyabilmesi için gerekli ve okunabilirlikle direkt ilişkili olan eğitim seviyesi olarak belirtilmektedir (DuBay, 2004).

Hisse senetleri borsada işlem gören şirketler hem finansal yatırımcılar hem de analistleri bilgilendirmek için işletmenin operasyonları ve finansal durumunu açıklayan pek çok doküman yayınlamakla yükümlüdür. Bu dokümanların kapsamı ve içeriği farklı ülkelerin menkul kıymetler borsaları için düzenleyici kurumlar tarafından detaylı olarak tanımlanmıştır. Bu konuda düzenleyici kurum ABD’de SEC¹, Türkiye’de ise Sermaye Piyasası Kuruludur² (SPK). Kamusal yarar gözetilerek

1 SEC’in borsada hisse senetleri işlem gören farklı nitelikteki şirketler için yayınlamayı zorunlu tuttuğu raporların listesi için: <https://www.sec.gov/forms>

2 SPK’nın finansal raporlamaya ilişkin esasları içeren mevzuatı için: <https://www.mevzuat.gov.tr/>

paylaşılması zorunlu kılınan bu metinler de tıpkı sağlık, eğitim veya askeri alanlarda paylaşılan ve öğretici nitelikteki dokümanlar gibi yoğun teknik bilgi ve terim içerir. Bu metinlerin okunabilirliği ise tartışmalıdır. Zira herhangi bir metnin okunabilirliği metnin yapısının yanı sıra okuyucunun dil yeteneği ve eğitim seviyesi ile de ilişkilidir (Al-Tamimi vd., 2014:373). Okuyucunun yeterlilik seviyesi genel olarak gelişim, öğrenim, deneyim ve yıllarca süren kapsamlı bir okuma deneyiminin ürünü olarak ulaşılan bir okuma yeteneği olarak tanımlanmaktadır (Thomas, 2013: 148-149). Ekonomik İş Birliği ve Kalkınma Örgütüne (OECD) göre temel bir yaşam becerisi olan finansal okuryazarlığı oluşturan faktörler finansal bilgi, finansal tutum ve finansal davranıştır (OECD, 2017). Bu çalışmada ele alınan metinlerin de finansal nitelikli oldukları düşünüldüğünde, doğru anlaşılması ve buradan hareketle doğru finansal kararlar alınabilmesi aynı zamanda finansal okuryazarlık seviyesinin de yeterliliğini gerektirmektedir (Lusardi ve Tufano, 2009).

İlerleyen bölümlerde de detaylandırılacağı üzere, bir metnin okunabilirlik seviyesinin belirlenmesinde farklı analiz ve yaklaşımlardan yararlanılmaktadır. Ele alınan metnin zorluk seviyesi ile ilgili bir değerlendirme yapan Fakhfakh'a (2015) göre metinde yer alan kelimelerin ve cümlelerin uzunluğu okuma gücünün derecesini yükseltmektedir (bkz. Tablo 3). Vajjala ve Meurers'a (2014) göre ise cümle bazında karşılaştırma ve değerlendirme yapabilen güvenilir bir okunabilirlik modeli bulunmamaktadır. Bu kapsamda modellerin geliştirilmesi okunabilirliği zor olan bir metnin nasıl sadeleştirilmesi gerektiği ile ilgili sistematik bir bakış açısı da sunabilecektir.

Tablo 3

Kelimelerin / Cümlelerin Uzunluğuna Göre Metinlerin Dilsel Zorluk Seviyesi

Dilsel Birimler	Karakter Sayısı	Zorluk Seviyesi
Kısa kelimeler	4 ve daha az harf	Düşük
Orta uzunlukta kelimeler	5-9 arası harf	Orta
Uzun kelimeler	10'dan fazla harf	Yüksek
Kısa cümleler	15'den az kelime	Düşük
Orta uzunlukta cümleler	15-20 arası kelime	Orta
Uzun cümleler	20'den fazla kelime	Yüksek

Kaynak: Fakhfakh (2015)

1998 yılında SEC tarafından yayınlanan dokümanda kullanılan “Sade İngilizce” ifadesinin bağlamı, başkan Arthur Levitt tarafından şöyle açıklanmıştır: “Yatırımcıların federal menkul kıymetler yasalarımızla sağlanan korumalardan tam olarak yararlanabilmesi için açıklama belgelerini okuması ve anlaması gerekmektedir. Birçok yatırımcı ne avukat ne muhasebeci ne de yatırım bankacısı olduğundan, belgeleri yatırımcıların anlayabileceği bir dilde yazmaya başlamalıyız.” (SEC, 1998:3). Söz konusu düzenleme, genel olarak, araştırmacı ve yatırımcılar ile paylaşılan finansal dokümanlarının hazırlanmasında okuyucunun anlamasını güçleştirecek legal ve teknik terminolojiden kaçınılması, kısa cümleler ve gündelik dil kullanımı, devrik cümlelere yer verilmemesi, karmaşık bilginin tablolar halinde sunulması gibi düzenlemeleri içermektedir (Loughran ve McDonald, 2010). Bu uygulamanın pek çok farklı dilde oluşturulacak finansal dokümanlar için de geçerli olacağı açıktır.

Finansal metinlerin okunabilirliği ile ilgili ABD’de yapılmış ilk çalışmalardan (Still, 1972; Dolphin ve Wagley, 1977) olan metin analizleri günümüzde de yoğun gelişim gösteren yöntemlerden biri olmuştur. Elektronik ortamda toplanan metinlerin hacminin giderek artmasına paralel olarak çeşitlenen analitik yöntemlerden biri olan metin madenciliği; eldeki dilsel verinin işlenmesi, analiz edilmesi, özetlenmesi ve anlaşılabilirliğinin değerlendirilmesinden oluşan yöntemler bütünüdür. Böylelikle herhangi bir konuda oluşturulmuş dağınık ve çok boyutlu metinler bütünüün değerlendirilmesinden hareketle karar vericilere kapsayıcı nitelikte bilgi sunulmaktadır. Söz konusu inceleme sağlık, finans gibi özel terimleri barındıran alanlarda, doğal dil işleme teknikleri gerektiren konularda ve metnin semantik olarak doğru düzenlenmesinin değerlendirileceği durumlarda yapıldığında yaklaşımın önemi ortaya çıkmaktadır (Aggarwal ve Zhai, 2012:3). Son yıllarda okunabilirlik analizleri metin madenciliği, doğal dil işleme algoritmaları ve makine öğrenmesi yöntemleri ile birleştirilerek uygulanmaktadır. Pek çok araştırmacı tarafından yapılan farklı çalışmalarda finansal raporlarının okunabilirlik düzeyinin şirketin karlılığının doğrudan tahmincisi olduğu tespit edilmiştir, bunun tersi de geçerlidir (Fisher vd., 2016:168). Konu ile ilgili yabancı kaynaklarda yapılan literatür taraması Tablo 4’de verilmiştir. Loughran ve McDonald (2016:1193) tarafından özellikle 2000 yılı öncesinde konu ile ilgili yapılmış çalışmalarda ele alınan örneklemin yetersizliğinin vurgulanması nedeniyle çalışma kapsamında tarama yapılırken 2005 yılı ve sonrasında gerçekleştirilmiş çalışmalar kapsama alınmıştır.

Tablo 4

Literatür Taraması

Kaynak	Çalışma İçeriği
Li (2008)	Yazar, yıllık finansal raporların okunabilirliğini (örn. 10-Ks) Fog İndeksi ve yıllık raporda yer alan kelime sayısına göre ölçmüştür.
Bloomfield (2008)	Li'nin (2008) bakış açısını takip eden yazar düşük performans gösteren firmaların raporlarında daha uzun kelime ve cümleler kullandığını, firma performansı ile raporların okunabilirlik seviyesi arasındaki ilişkileri göstermiştir.
Biddle, Hilary ve Verdi (2009)	Çalışmada raporlama kalitesi Fog İndeksine göre yüksek olan firmaların daha yüksek sermaye yatırım verimliliğine sahip oldukları tespit edilmiştir.
You ve Zhang (2009)	Çalışmada SEC'in EDGAR 10-K dosyalarının yayınlanma tarihlerine göre anlık ve gecikmeli piyasa tepkisi incelenmiştir. 10-K formlarının yayın tarihlerindeki günlerde alışılmadık işlem hacimleri ve hisse senedi fiyat hareketleri belgelenmiştir. Daha karmaşık 10-K raporları olan şirketler için yatırımcıların fiyat değişikliklerine daha tepkisz oldukları görülmüştür.
Loughran ve McDonald (2010)	1994-2007 yılları arasında 42.357 adet 10-K dokümanı için farklı indekslerle okunabilirlik analizi yapılmıştır. Okunabilirlik düzeyinin incelenen dönem boyunca artan bir trend izlediği gözlenmiştir.
Miller (2010)	Çalışmanın sonuçlarına göre küçük yatırımcılar yüksek Fog İndeksi ve kelime sayısına sahip, diğer bir deyişle okunabilirlik düzeyi düşük 10-K raporları yayınlayan firmalara anlamlı olarak düşük düzeyde yatırım yapmaktadır.
Lehavy, Li ve Merkley (2011)	Çalışmada yazarlar okunabilirlik ile analistlerin düzeyi arasındaki bağlantıya odaklanılmışlardır. Fog İndeksi ile yapılan değerlendirmeye göre daha okunabilir yıllık raporların daha düşük profilli analist dağılımına ve daha yüksek kazanç tahmin doğruluğuna sahip olduğu bulunmuştur.
Rennekamp (2012)	Yazar, okunabilirlik seviyesi farklı ancak içerdiği bilgi düzeyi ve uzunluğu eşit olan perakende sektörüne ait finansal dokümanların 234 katılımcı tarafından değerlendirilmesini isteyerek deneysel bir çalışma gerçekleştirmiştir. Buna göre SEC "Sade İngilizce" standardına göre okunabilirliği artırılan metinlerdeki pozitif ve negatif nitelikli bilgilere dair katılımcıların tepkileri daha az okunabilir olanlara nazaran daha yüksek çıkmıştır.
Lee (2012)	Çalışma, firmaların zorunlu olarak yayınladıkları çeyrek raporlarının (10-Q) okunabilirliği ile hisse senedi fiyatlarının bilgi etkinliği arasındaki ilişkiyi araştırmaktadır. Okunabilirlik Fog İndeksi ve metnin uzunluğuna dayalı olarak ölçülmüştür.
Lawrance (2013)	Çalışmanın sonuçlarına göre hem Fog İndeksi hem de yıllık rapordaki kelime sayısı perakende sektöründeki yatırımcıların hisse senedi varlıkları ile ilişkilidir. Çalışmanın örnekleminde, ABD'deki toplam 78.000 hane halkı için 1994-1996 dönemine ait gerçek portföy varlıkları bulunmaktadır. Bireysel yatırımcıların yatırım yaptığı firmaların yıllık raporlarının daha az kelime içerdiği ve daha iyi okunabilirliğe sahip olduğu tespit edilmiştir.
Lundholm, Rogo ve Zhang (2014)	Yazarlar hisseleri ABD'deki borsalarda işlem gören yerli ve yabancı firmaların finansal dokümanlarının okunabilirlik düzeylerini Fog İndeksi ile karşılaştırmıştır. Buna göre yabancı menşeli firmaların daha fazla yatırım çekebilmek için yerlilere nazaran daha okunabilir formlar yayınladığı görülmüştür.
Loughran ve McDonald (2014)	Fog İndeksinin finansal uygulamalarda eksik ve hatalı kullanımına karşın 10-K dokümanlarının dosya büyüklüğünün okunabilirlik düzeyi olarak basitçe değerlendirilebilecek bir değişken olduğu ve kullanım kolaylığının yanı sıra Fog İndeksinden daha iyi performans gösterdiği ortaya konmuştur.
De Franco v.d. (2015)	Analist raporlarının okunabilirliğinin de yatırımcı davranışıyla ilişkili olduğu görüşüne dayanan yazarlar, okunabilirlik seviyesi için 2002-2009 döneminde yayınlanan 365.000'den fazla yıllık raporun bir örneğini üç farklı okunabilirlik indeksinin (Fog, Flesch ve Flesch-Kincaid) bir araya getirilmesi ile analiz etmişlerdir. Yazarlar, daha okunabilir analist raporlarının izleyen üç gün içinde önemli ölçüde daha yüksek işlem hacmi ile ilişkilendirildiğini keşfetmiştir.
Fakhfakh (2015)	Yazar okunabilirliği finansal bilgi kalitesi ve finansal iletişim etkinliğini belirleyen bir faktör olarak nitelendirmekte olup, kelime uzunluğu, cümle ve paragraf büyüklüğü gibi dilsel normların okuma zorluğunu ölçen en önemli değişkenlerden olduğu sonucuna varmıştır.
Tan, Wang ve Zhou (2015)	Yazarlar okunabilirliğin (yüksek ve düşük) ve karşılaştırmalı performans tutarlılığının (tutarsız ve tutarsız) yatırımcıların kararlarını nasıl etkilediğini araştırmış ve karşılaştırmalı performans tutarsız olduğunda finansal dokümanın okunabilirlik düzeyinin düşük olduğunu, yatırımcıların da bu tip firmaları yatırım için tercih etmediğini göstermişlerdir.
Loughran ve McDonald (2016)	Çalışmada, finansal metinlerin okunabilirliğinin değerlendirilmesinde yoğun ilgi gören sayısal analizlerin yanı sıra metin analizinin de önemine değinilerek bu alanda bir uygulama çalışması gerçekleştirilmiştir.
Guay, Samuels ve Taylor (2016)	Fog İndeksi dahil altı farklı okunabilirlik ölçümüne dayanarak incelenen yıllık raporlara göre okunabilirlik düzeyi düşük raporları olan şirketlerin hisse satış ve nakit akışları için yönetime yönelik tahminler yayımlayarak negatif yönlü okunabilirlik etkisini azaltma eğiliminde olduğu tespit edilmiştir.
Bonsall IV vd. (2017)	Yazarlar makalede SEC'in "Sade İngilizce" kurallarını içine alan yeni bir okunabilirlik indeksini (Bog İndeks) önermişlerdir.
Dyer, Lang, ve Stice-Lawrence (2017)	Çalışmada 1996-2013 döneminde yayınlanan 10-K raporlarının okunabilirliğinin incelenmesi Latent Dirichlet Allocation (LDA) yaklaşımı ile gerçekleştirilmiştir. Buna göre metnin içerisinde yer alan gerçeğe uygun değer, iç kontroller ve risk faktörü açıklamaları metnin karakteristiğini açıklamada önemli değişkenler olarak bulunmuştur.
Lahtinen ve Shipe (2017)	Şirket büyüklükleri ve finansal rapor okunabilirliği üzerine yapılan çalışmalarda, şirketlerin sağladığı hizmet sayısı arttıkça kamuya açıklanan rapor ve formların okunabilirliğinin zorlaştığı tespit edilmiştir. Çalışmada, okunabilirliğin zorlaşmasındaki en önemli nedenlerden birinin bireysel yatırımcıların finansal raporlarda okuduklarını anlamak için finansal jargonu sahip olmadıkları, yatırım danışmanlarına ihtiyaç duydukları ve bu durumun şirketleri anlaşılması zor bir yazım diline yönelttiği belirtilmektedir.
Bushee, Gow ve Taylor (2018)	Firmaların yayınladığı finansal metinlerdeki dilin karmaşıklığının sağlanan bilginin de karmaşıklığını yansıtacağı varsayımı ile dilsel karmaşıklık bilgisi asimetrisiyle zat ilişkileri olan iki örtük bileşeni (bilgi ve gizleme) bir araya getirdiğinin gösterilmesi amaçlanmıştır.
Lim, Chalmers ve Hanlon (2018)	Çalışmada firmaların iş stratejisinin yıllık finansal rapor okunabilirliğinin belirleyicisi olduğu gösterilmiştir. Okunabilirlik düzeyi ölçütü olarak Fog İndeksi kullanılmıştır.
Chakrabarty vd. (2018)	Çalışma kurumsal yıllık finansal raporların (10-K) okunabilirliği üzerine hisse senedi opsiyonlarına dayalı tazminatın rolünü incelemek üzere gerçekleştirilmiştir.

Luo, Li ve Chen (2018)	Yazarlar 2001-2015 yılları arasında Çin A-hissesi olan 19.221 firma-yıl verisini kullanarak, yıllık rapor okunabilirliği ve kurumsal ajans maliyetleri arasındaki ilişkiyi araştırmıştır. Okunabilirliğin rapor uzunluğu ve / veya dosya boyutu ile temsil edildiği çalışmada yıllık raporlarının okunabilirliği yüksek olan firmaların daha düşük ajans maliyeti ile karşılaştıkları görülmüştür.
Hesarzadeh ve Bazrafshan (2018)	Çalışmanın amacı, kurumsal rapor okunabilirliğinin düzenleyici denetleme riskini azaltıp azaltmadığını araştırmaktır. Çalışmada okunabilirlik Fog İndeksi kullanılarak ölçülmüştür. Bulgular, kurumsal raporların okunabilirliğinin düzenleyici denetleme riskini etkileyen faktörler ile ilişkili olduğunu bu riski azalttığını ortaya koymaktadır.
Boubaker, Gounopoulos ve Rjiba (2019)	Yıllık finansal raporların dilsel karmaşıklığının firmaların stok likiditesi üzerindeki etkisini araştıran çalışmada hesaplamalı dilbilim teknikleri kullanılmış, okunabilirlik düzeyi düşük raporların daha düşük stok likiditesi ile ilişkili olduğu tespit edilmiştir. Buna göre çalışmada, okunması zor yıllık raporların yatırımcıların bu raporlarda yer alan bilgileri işleme ve analiz etme yeteneğini engellediğine ve böylece yatırım istekliliklerini azalttığına dair kanıtlar sunulmaktadır.
Luo ve Chen (2018)	Yazarlar okunabilir finansal raporların şirket içerisindeki yöneticilerin fırsatçı yatırım davranışlarını takip etmede kullanılabileceği ve bu durumun temsil maliyetlerinde bir düşüşe sebep olacağını belirtmektedir.
Jha (2019)	Yayımlanan finansal metinler üzerinde sosyal sermayenin etkisi incelenmiştir. Yüksek sosyal sermayeye sahip bölgelerde kurulmuş ABD firmalarının yıllık raporlarının okunabilirlik düzeyinin yüksek olduğu görülmüştür.
Kim, Wang ve Zhang (2019)	Çalışmada okunabilirlik düzeyi düşük 10-K raporlarına sahip firmaların hisse senedi fiyatlarında ani düşüş riski ile karşı karşıya olduğu gösterilmektedir.
Cassell, Cunningham ve Lisic (2019)	Çalışmada işleme akıcılığı teorisi uygulanarak, ele alınan şirketin bir SEC yorum mektubuna ilk tepkisinin okunabilirliği ve incelemeye kaynaklanan olumsuz sonuçlarla ilişkili olup olmadığı incelenmiştir. Buna göre okunabilirlik düzeyi düşük şirket yanıtlarının, daha uzun yanıt süreleri ve şirketin incelenmiş bir dosyayı yeniden düzenleme veya değiştirme olasılığının daha yüksek olması ile ilişkili olduğu görülmüştür.
deHaan vd. (2020)	S&P 500 indeksinde hemen hemen aynı yönetmeliklere, risklere ve brüt getirilere sahip ancak büyük ölçüde farklı fiyatlanabilen fonların açıklamalarındaki örtüklük durumu incelenmiştir. Yatırım fonları için özel olarak tasarlanmış tedbirler kullanarak, yüksek ücretleri gereksiz yere karmaşık açıklamalarla gizlemeye çalışan fonlar tespit edilmiş ve metinlerin okunabilirlik düzeyi ile fiyat dağılımının ilişkisine değinilmiştir.

Ülkemizde okunabilirlik üzerine yapılan çalışmalar incelendiğinde bunların temelde eğitim alanında yoğunlaştığı görülmektedir. Ateşman (1997) tarafından yapılan ölçüme dayalı ilk çalışma, Flesch'in (1948) yaklaşımının Türkçeye uyarlanmasıdır. Bunun yanı sıra Boztaş ve diğerleri (2014) sağlık sektöründeki metinlerin okunabilirliği üzerine bir çalışma gerçekleştirmiş, Adalı (2018) ise yüksek lisans tezinde Sosyal Güvenlik Mevzuatının okunabilirliğini ele alarak farklı alanlarda konunun önemini ve uygulanabilirliğini göstermiştir. Finansal alanda yapılan tek çalışma ise Keskin ve diğerlerine (2020) ait olup; bankacılık, imalat ve telekomünikasyon sektörlerine ait bağımsız denetçi raporlarının Otomatik Okunabilirlik İndeksi (ARI) ile değerlendirilmesi ve bunda cinsiyet rolünün etkisinin tespit edilmesi üzerinedir. Özcan ve Türü (2004) tarafından 1292 yatırımcı ile telefonla anket yöntemiyle gerçekleştirilen çalışmada ise katılımcılara halka açık şirketler ve hisse senetleri ile ilgili önemli bilgiler içeren dokümanlar ile mali tabloları inceleyip incelemedikleri, inceleyenlere de dokümanları anlamakta zorluk çekip çekmedikleri sorulmuştur. Sonuçlara göre yatırımcıların %85'i bu tip raporları ve dokümanları incelemekte ancak anlamakta zorluk çekmektedirler.

4. VERİ SETİ VE YÖNTEM

Çalışmada, 01.01.2010-01.10.2019 tarihleri arasında NYSE'de işlem gören ve teknoloji alanında faaliyet gösteren 144 şirkete ilişkin toplam 23.103 adet kamuya açıklama formu kullanılmıştır. İlgili formlar SEC EDGAR elektronik platformundan elde edilmiş olup; tamamlanmamış ve eksik olan form kayıtlarıyla birlikte, form eklerinde bulunan SMS ve e-mail yazışmaları veri setinden çıkartılmıştır. Çalışma kapsamında incelenen formların şirketler tarafından raporlanması ve kamuya açıklanması zorunlu olduğundan, kelimelerin tekil anlamları ve cümlelerde hangi sırada bir araya gelerek cümle anlamını oluşturduğuna ilişkin kurallar, diğer bir deyişle, söz dizimsel ve anlamsal olarak uygunluğu da SEC tarafından düzenlenmiştir. Buna göre formların etken ve kısa cümleli yapıda, sade ve anlaşılır bir dilde, okunaklı bir şekilde ve gerekli olmayan detayları içermeyen sade kelimelerle düzenlenmiş olması gerektiği çeşitli kurullarla zorunlu hale getirilmiştir (SEC, 1998). Çalışmada ayrıca, Lo vd. (2017) tarafından şirket karlılığını ölçen ve finansal form okunabilirliğini etkileyen önemli faktörlerden biri olarak gösterilen şirket büyüklüğü değişkeni kullanılarak, şirket büyüklüklerine göre form okunabilirliklerinin anlamlı derecede farklılaşıp farklılaşmadığı araştırılmıştır.

Veri seti, R programlama dili kullanılarak analiz edilmiş olup, XBRL, XML, pdf ve HTML türündeki formların işlenebilir hale getirilebilmesi için öncelikle metin dosyasına dönüştürme işlemi gerçekleştirilmiştir. Veri ön işleme sürecinde sırasıyla aşağıdaki adımlar izlenmiştir:

1. Formlarda bulunan grafiklerin veya XML, XBRL türündeki istenmeyen ifadelerin tespit edilerek veri setinden çıkartılması,
2. Unicode ve HTML türündeki metin ifadelerinin ASCII formatına dönüştürülmesi,
3. Formlardaki ifadelerde yer alan para cinsi gibi simge ve sembollerin yazı formatına dönüştürülmesi,
4. Formlardaki söz dizimsel yapının dil bilgisi kurallarına göre çözümlenebilmesi amacıyla ayrıştırma ağacının (parse tree) oluşturulması,
5. Ayrıştırma ağacının temizlenmesi: Metin belgeleri yarı yapılandırılmış veya yapılandırılmamış bir ifadeler yığınından oluşmaktadır. Bu safhada ayrıştırma ağacında bulunan sayfa numaraları, tablolar, istenmeyen XML ve HTML etiketleri gibi unsurların temizlenerek sadece okunabilir yazının kalması amaçlanmıştır.
6. Aynı paragrafa ait tüm cümlelerin aynı satırda gruplandırılması: Okunabilirlik indeksinin hesaplanabilmesi için paragraf sayımı önemli olduğundan, aynı paragrafa ait tüm cümleler aynı satırda gruplandırılarak paragraf yapısı büyük oranda korunmuştur.

Form içerisindeki karakterlerin sayımında, Gillick'in (2009) çalışmasında belirtildiği şekilde, beş veya daha fazla kelimenin oluşturduğu cümleler dikkate alınmıştır. Hece sayımı ise Carnegie Mellon Üniversitesi Telaffuz Sözlüğü (CMUdict) kullanılarak gerçekleştirilmiş olup, ilgili sözlükte bulunmayan telaffuzların varlığı durumunda R programlama dili *NLP* paketi (Hornik ve Hornik, 2018) kullanılarak standart doğal dil işleme yardımıyla hece sayısı hesaplanmıştır. Kelimelerin sayımında kısa çizgi, tırnak işareti, parantez ve denden gibi noktalama işaretleri ihmal edilmiştir. Form içerisindeki cümlelerin tespiti için yukarıda belirtilen çalışmadan hareketle, beş ve daha fazla kelimenin oluşturduğu dizilimler kullanılmıştır. Cümlelerin başlangıç ve bitişinin tespitinde ise, %0,35 hata oranı ile doğru tahmin gerçekleştiren Cümle Sınır Tespiti (SBD) algoritmasından yararlanılmıştır (Gillick, 2009). Bir satır başından diğer satır başına kadar olan kelimeler bütünü olarak tanımlanan paragraf sayımı için, en az on kelimededen oluşan metinler dikkate alınmıştır. HTML uzantılı dosyalarda, temizleme betiği (script) dosya içerisinde <p>,
, <i>, veya <tr> gibi etiketlere rastlandığında satır başı, başka bir deyişle yeni paragraf oluşturmaktadır. Biçimlendirilmiş metin dosyalarında ise boş satırlar yeni paragraf başlangıcını göstermektedir. XML yapısındaki dokümanlar yapılandırılmış içeriğe sahip olduğundan belge yapısında herhangi bir hata yoksa içeriğin bilgi girişi doğru ve eksiksiz kabul edilmekte olup, bu yapı metin madenciliği uygulamalarında da sıklıkla tercih edilmektedir (Chakrabarty v.d., 2018). Veri setini oluşturan finansal formlara ilişkin tanımlar Tablo 5'te verilmiştir.

Tablo 5

İncelenen Form Türleri

Form Türü	Tanım
3	Menkul kıymet ihraç eden şirketlerin yöneticileri, çalışanları ve herhangi bir menkul kıymet sınıfının %10'unundan fazlasına sahip olan hissedarlar, gerçekleştirdikleri işlemleri belirlenen süreler içerisinde SEC'e raporlamaktadır. Menkul kıymetlerin mülkiyet beyannamesi niteliğinde doldurulan Form 3 söz konusu yönetici, çalışan ve hak sahiplerinin varlıklarını açıklamak için kullanılmaktadır. Ayrıca ilgili form içeriden öğrenenlerin ticaretini düzenlemek için önemli bir adım olarak nitelendirilmektedir. Form 3 doldurulup raporlandıktan sonra kamu denetimine açık hale gelmektedir. Ayrıca söz konusu kişilerin menkul kıymet borsalarında takibi ve şüpheli davranışlarının tespiti Form 3'de sağlanan bilgiler aracılığıyla gerçekleştirilmektedir. Formda içeriden öğrenenin (insider) kimliği, şirket içerisindeki pozisyonu, sahibi olduğu menkul kıymetin türü, miktarı ve ihraççı şirkete ilişkin beyan bulunmaktadır.
4	İçeriden öğrenenlerin menkul kıymetlerinde meydana gelen değişiklikler Form 4'de raporlanmakta olup, işlemin gerçekleştirildiği günden itibaren iki iş günü içerisinde SEC'e bildirilmelidir. Menkul kıymet hak sahipliğine ilişkin ilk beyan Form 3 ile verilmekte olup, hak sahipliğine ilişkin değişiklikler Form 4'te bildirilmektedir. Formda gerçekleştirilen işlemin tarihi, türü ve miktarı beyan edilmektedir.
8-K	Şirket bünyesinde gerçekleşen ve hissedarları ilgilendiren önemli olayların ve değişikliklerin duyurulmasını sağlayan 8-K, SEC'e bildirilen en önemli formlardan biridir. Mali yıl içerisindeki iflas, şirket devri, yönetici istifası, anlaşma feshi, maddi varlık edinimi veya satışı, tesis değişikliği, hissedar haklarında yapılan değişiklikler, kredi temerrüt takasları, sertifika / ana sözleşme veya tüzük değişiklikleri ve şirket genel kurulu oylama sonuçları gibi olaylar söz konusu formda dört gün içerisinde raporlanmakta ve kamuya duyurulmaktadır.
10-K	10-K formu, halka açık bir şirket tarafından şirketin kendi finansal performansı hakkında SEC'e yıllık olarak sunduğu en kapsamlı rapordur. Form, kamuyu aydınlatma kapsamında şirketlerin raporladıkları en ayrıntılı doküman olmakla birlikte, şirketlerin olağan genel kurullarından önce hissedarlarla paylaşılmaktadır. Söz konusu form şirket geçmişi, organizasyon yapısı, idari tazminatlar, öz sermaye yapısı, bağlı ortaklıklar ve denetlenmiş finansal tablolar gibi bilgileri içermektedir.

10-Q

Menkul kıymet ihraç eden şirketlerin her üç ayda bir SEC'e sunmasının gerekli olduğu 10-Q formu denetlenmemiş olup, 10-K ile benzer nitelikte bilgiler içermektedir. Formda, genellikle şirketlerin finansal pozisyonlarıyla ilgili daha az ayrıntıda bilgi sunulmaktadır. İlgili form yılın ilk üç çeyreği için düzenlenmekte ve son çeyrek için ayrıntılı 10-K formu beyan edilmektedir. Şirketler halka açıklıklarının büyüklüğüne bağlı olarak, çeyrek sonundan itibaren 40, 45 veya 10 gün içerisinde ilgili formu doldurmak zorundadır. Örneğin, şirketin halka açık menkul kıymet değerinin en az 700 milyon dolar olması durumunda, şirketin ilgili çeyrek mali yılını takiben en geç 40 gün sonra 10-Q formunu düzenleyip SEC'e bildirmesi zorunludur.

Kaynak: Yazarlar tarafından oluşturulmuştur.

Aşağıdaki uygulama bölümünde teknoloji alanında faaliyet gösteren ve NYSE'ye kote olmuş 144 şirkete ilişkin kamuya açıklama formlarının okunabilirlik analizleri ile bu analizlerden elde edilen bulgulara yer verilecektir.

5. UYGULAMA

Formların raporlandığı yıllar itibariyle dağılımlarına bakıldığında çoğunluğu şirketin piyasa değerini etkileyebilecek nitelikteki önemli olayların veya değişikliklerin raporlandığı 8-K formlarının oluşturduğu ve son 15 yılın en büyük değer artışının yaşandığı 2013 yılını takip eden 2014 yılında ise, içeriden öğrenenlerin raporladıkları Form 3 ve 4'lerin sayısında ciddi bir artış kaydedildiği görülmektedir (bkz. Tablo 6). Ayrıca aynı sene içerisinde, sisteme en yüksek sayıda form girişi gerçekleştirildiği görülmektedir. Diğer taraftan 10-K ve 10-Q formlarının dağılımı yaklaşık olarak aynı oranda seyretmekte olup, yıllar itibariyle önemli bir değişiklik tespit edilmemiştir.

Tablo 6

Yıllara Göre Form Sayıları

Yıllar	Form Türü					Toplam
	10-K	10-Q	3	4	8-K	
2010	114 (%6,6)	337 (%19,5)	19 (%1,1)	106 (%6,1)	1155 (%66,7)	1731 (%100)
2011	113 (%6,4)	342 (%19,4)	24 (%1,4)	81 (%4,6)	1200 (%68,2)	1760 (%100)
2012	118 (%6,5)	355 (%19,5)	21 (%1,2)	91 (%5,0)	1232 (%67,8)	1817 (%100)
2013	117 (%6,2)	360 (%18,9)	25 (%1,3)	106 (5,6%)	1294 (%68,0)	1902 (%100)
2014	127 (%2,7)	393 (%8,4)	153 (%3,3)	2564 (%54,8)	1446 (%30,9)	4683 (%100)
2015	133 (%4,0)	393 (%11,7)	155 (%4,6)	1253 (37,3%)	1421 (%42,4)	3355 (%100)
2016	115 (%5,4)	386 (%18,3)	113 (%5,4)	177 (%8,4)	1321 (%62,5)	2112 (%100)
2017	129 (%6,6)	375 (%19,2)	18 (%0,9)	99 (%5,1)	1329 (%68,2)	1950 (%100)
2018	137 (%6,1)	408 (%18,2)	29 (%1,3)	103 (%4,6)	1561 (%69,7)	2238 (%100)
2019*	105 (%6,8)	278 (%17,9)	29 (%1,9)	69 (%4,4)	1074 (%69,1)	1555 (%100)
Toplam	1208 (%5,2)	3627 (%15,7)	586 (%2,5)	4649 (%20,1)	13033 (%56,4)	23103 (%100)

*01.10.2019 itibariyle.

Tablo 7'de her bir form türüne ilişkin okunabilirlik indeksini oluşturan değişkenlerin tanımsal istatistikî ölçütleri verilmiştir. Ortalama karakter, hece, kelime, cümle, paragraf ve çok heceli kelime sayısı açısından 10-K formları diğer form türlerine göre yüksek bulunmuştur. Metnin fiziki niteliği ve okunabilirliğinin değerlendirilmesinde etkili faktörlerden olan kelime ve cümle uzunluklarının 10-K ve 10-Q formlarında çok yüksek olduğu tespit edilmiştir. Buna karşın, cümle başına düşen ortalama kelime sayısı ve paragraf başına düşen ortalama kelime sayısının içeriden öğrenenlerin kullandığı Form 3 ve Form 4'de yüksek olduğu görülmektedir. Normallik varsayımı sağlanamadığından dolayı, form türüne göre ortalama skor farklılıklarının araştırılmasında parametrik olmayan Kruskal-Wallis testi kullanılmıştır.

Tablo 7

Okunabilirlik Değişkenleri Tanımlayıcı İstatistikler ve Kruskal-Wallis Testi Sonuçları

Değişken	Form Türü	Ortalama Skor	S.S.	Q1	Medyan	Q3	H*
Karakter sayısı	a 10-K (b, c, d, e)	320,376	134,391	243,979	300,712	375,593	11,341**
	b 10-Q (a, c, d, e)	146,371	92,265	84,103	126,973	188,184	
	c Form 3 (a, b, d, e)	3,734	2,180	3	3,181	3,947	
	d Form 4 (a, b, c)	5,938	4,646	4	4,330	6,756	
	e 8-K (a, b, c, d)	13,115	54,687	2,415	3,391	6,782	
Hece sayısı	a 10-K (b, c, d, e)	89,699	37,357	68,280	84,539	105,312	11,329**
	b 10-Q (a, c, d, e)	40,621	25,510	23,096	35,098	52,726	
	c Form 3 (a, b, d, e)	1,019	598	3	881	1,103	
	d Form 4 (a, b, c)	1,593	1,273	4	1,129	1,802	
	e 8-K (a, b, c, d)	3,584	15,227	651	913	1,836	
Kelime sayısı	a 10-K (b, c, d, e)	47,639	20,063	36,341	44,766	55,660	11,456**
	b 10-Q (a, c, d, e)	21,828	13,818	12,542	18,833	28,113	
	c Form 3 (a, b, d, e)	505	238	3	482	589	
	d Form 4 (a, b, c)	511	344	4	438	552	
	e 8-K (a, b, c)	2,004	8,540	362	507	1,010	
Cümle sayısı	a 10-K (b, c, d, e)	1,893	690	1,491	1,820	2,207	17,760**
	b 10-Q (a, c, d, e)	830	431	509	755	1,079	
	c Form 3 (a, b, d, e)	8	5	3	8	10	
	d Form 4 (a, b, c)	7	8	4	6	9	
	e 8-K (a, b, c)	61	171	21	26	40	
Paragraf sayısı	a 10-K (b, c, d, e)	656	274	505	627	770	17,512**
	b 10-Q (a, c, d, e)	292	196	174	258	377	
	c Form 3 (a, b, d, e)	5	10	3	1	7	
	d Form 4 (a, b, c)	2	5	4	1	1	
	e 8-K (a, b, c, d)	28	91	11	13	19	

Çok heceli- kompleks kelime sayısı	a	10-K (b, c, d, e)	12392	5089	9498	11712	14507	11,238**
	b	10-Q (a, c, d, e)	5451	3359	3132	4769	7063	
	c	Form 3 (a, b, d, e)	132	80	92	115	148	
	d	Form 4 (a, b, c)	210	175	111	147	239	
	e	8-K (a, b, c, d)	469	1969	89	126	253	
Cümle başına düşen ortalama kelime sayısı	a	10-K (c, d, e)	25	3	23	25	26	12,711**
	b	10-Q (c, d, e)	26	8	23	25	27	
	c	Form 3 (a, d, e)	76	60	3	64	87	
	d	Form 4 (a, b, c)	123	119	4	85	149	
	e	8-K (a, b, c, d)	23	9	17	20	25	
Paragraf başına düşen ortalama kelime sayısı	a	10-K (c, d, e)	71	18	64	69	74	13,544**
	b	10-Q (c, d, e)	73	18	63	70	78	
	c	Form 3 (a, b, d, e)	288	252	3	278	468	
	d	Form 4 (a, b, c)	300	249	4	229	414	
	e	8-K (a, b, c, d)	46	88	26	33	48	

Not: Tabloda parantez içinde ve *italik* olarak verilen harfler, farklı harflerle kategorize edilmiş form türleri için çoklu karşılaştırma testi sonucunda birbirinden istatistiksel olarak anlamlı derecede ($p < 0,01$) farklılaşan grupları göstermektedir.

** $p < 0,01$

Tablo 7’de gösterildiği üzere, form türlerinde ortalama karakter sayısı, hece sayısı, paragraf ve çok heceli kelime sayısı ile birlikte, cümle ve paragraf başına düşen ortalama kelime sayısı %99 güven düzeyinde istatistiksel olarak anlamlı derecede farklılaşmaktadır. Kruskal-Wallis testine göre birbirinden istatistiksel olarak anlamlı formları tespit etmek için post-hoc yaklaşımlarından Dunn’ın (1964) sıra toplamlarına dayalı çoklu karşılaştırma testi gerçekleştirilmiştir. Buna göre, ortalama kelime ve cümle sayısında, Form 3 ve Form 4 birbirinden anlamlı derecede farklı olmayıp, diğer form türleri birbirlerinden ve Form 3 ile Form 4’den istatistiksel olarak anlamlı derecede farklılaşmaktadır. Paragraf başına düşen ortalama kelime sayısı 10-K ve 10-Q formlarında benzer olup, diğer form türlerinde ortalama kelime sayısı istatistiksel olarak anlamlı derecede farklılaşmaktadır. Genel olarak değerlendirildiğinde, her bir değişkene ilişkin ortalama skorlar 10-K ve 10-Q formlarında yüksek bulunup anlamlı olarak farklılaşsa da Form 3 ve Form 4’ün, diğer form türlerine göre birbirlerine daha yakın skorlara sahip olduğu söylenebilir.

Tablo 8

Farklı Okunabilirlik İndekslerine Ait Skor ve Tanımlayıcı İstatistikler

İndeks Türü	Form Türü	Ortalama Skor	S.S.	Q1	Medyan	Q3	H*	
Dosya boyutu (MB)	a	10-K (b, c, d, e)	12,71	8,07	7,46	12,19	16,18	17,156**
	b	10-Q (a, c, d, e)	7,01	4,93	3,66	6,18	9,26	
	c	Form 3 (a, b)	0,02	0,11	0,011	0,017	0,020	
	d	Form 4 (a, b)	0,01	0,01	0,012	0,015	0,021	
	e	8-K (a, b)	0,53	2,20	0,03	0,11	0,35	
Flesch-Kincaid	a	10-K	16,35	0,96	15,79	16,24	16,75	13,694**
	b	10-Q	16,45	3,10	15,46	16,23	17,03	
	c	Form 3 (d)	37,97	26,36	24,79	32,85	39,16	
	d	Form 4 (c)	68,65	52,63	36,28	51,99	85,46	
	e	8-K	14,50	3,68	12,14	13,38	15,60	
Gunning Fog	a	10-K	20,40	0,97	19,79	20,31	20,81	13,171**
	b	10-Q	20,39	3,13	19,45	20,18	20,93	
	c	Form 3 (d)	40,95	25,13	28,76	35,39	43,90	
	d	Form 4 (c)	65,32	50,54	33,81	46,42	78,88	
	e	8-K	18,85	3,73	16,43	17,88	20,07	
Coleman-Liau	a	10-K	22,54	0,62	22,18	22,47	22,83	4,403**
	b	10-Q	22,45	1,25	21,92	22,26	22,65	
	c	Form 3 (d)	27,42	11,46	20,55	21,36	30,88	
	d	Form 4 (c)	51,54	20,16	35,75	56,55	65,13	
	e	8-K	22,07	1,38	21,15	21,83	22,73	
ARI	a	10-K	22,70	1,28	21,95	22,49	23,15	13,124**
	b	10-Q	23,14	4,20	21,86	22,65	23,79	
	c	Form 3 (d)	51,69	34,75	33,62	44,99	53,57	
	d	Form 4 (c)	94,44	68,88	52,13	74,57	117,47	
	e	8-K	21,39	4,80	18,36	19,89	22,61	
Smog	a	10-K	17,66	,68	17,24	17,61	17,97	13,240**
	b	10-Q	17,61	1,23	17,00	17,52	18,06	
	c	Form 3 (d)	27,76	9,82	22,50	25,78	28,55	
	d	Form 4 (c)	41,09	20,76	26,33	34,68	51,27	
	e	8-K	16,33	2,39	14,64	15,73	17,39	

Not: Tabloda parantez içinde ve italik olarak verilen harfler, farklı harflerle kategorize edilmiş form türleri için çoklu karşılaştırma testi sonucunda birbirinden istatistiksel olarak anlamlı derecede ($p < 0,01$) farklılaşan grupları göstermektedir.

** $p < 0,01$

Farklı form türlerinin okunabilirlik indekslerine ilişkin ortalama skorlar ve Kruskal Wallis testine dayalı farklar ve post-hoc sonuçları Tablo 8'de gösterilmiştir. Buna göre, Flesch – Kincaid İndeksine göre 10-K ve 10-Q formlarının skorlarının birbirine benzer olup, en az lisans seviyesinde eğitime sahip kişilerce anlaşılabilir nitelikte, “zor” okunabilirlik düzeyinde olduğu tespit edilmiştir. Söz konusu formlar için Smog İndeks skoru 17, diğer indeks skorları da 20'nin üzerinde olup, okunabilirlik

düzeyi “çok zor” olarak nitelendirilmektedir. 8-K formu için elde edilen skorlar 10-K ve 10-Q formlarıyla benzerlik göstermekte olup, söz konusu formlara ilişkin skorlar arasında %99 güven düzeyinde istatistiksel olarak anlamlı derecede farklılık bulunmamaktadır. Buna karşın, Form 3 ve 4’ün diğer formlardan daha da zor anlaşılabilir olduğu ve söz konusu formlara ilişkin okunabilirlik skorlarının birbirlerinden ve diğer form türlerinden istatistiksel olarak anlamlı derece farklılaştığı sonucuna ulaşılmıştır. Diğer taraftan, Loughran ve McDonald (2014) tarafından gerçekleştirilen finansal tablo açıklamalarının okunabilirliğine ilişkin çalışmada, formların dosya büyüklüğünün okunabilirlik indeksleriyle ilişkili olduğu ve belge ayrıştırma gerektirmeden hesaplanabilen, basit fakat önemli bir değişken olarak nitelendirilebileceği belirtilmektedir. Dosya boyutu açısından bakıldığında, işletmelerin detaylı bilgilerini kamuya açıkladığı 10-K formlarının boyutunun diğer form türlerine göre anlamlı derecede yüksek olduğu, Form 3 ve Form 4’e ait dosya boyutlarının istatistiksel olarak anlamlı derecede farklılaşmadığı görülmektedir.

Tablo 9

Okunabilirlik İndeksleri Korelasyon Tablosu

İndeks Türü (N: 23.103)	Flesch-Kincaid	Coleman-Liau	Gunning Fog	ARI	Smog	Dosya Boyutu (MB)
Flesch-Kincaid	1	0,811*	0,996*	0,999*	0,987*	-0,057*
Coleman-Liau		1	0,757*	0,828*	0,851*	-0,051*
Gunning Fog			1	0,993*	0,976*	-0,058*
ARI				1	0,988*	-0,065*
Smog					1	-0,047*
Dosya Boyutu (MB)						1

* $p < 0,01$

Tablo 9’da okunabilirlik indeksleri arasındaki doğrusal ilişki Pearson Korelasyon Katsayıları ile incelendiğinde, her bir indekste üretilen skorların birbirleriyle istatistiksel olarak anlamlı pozitif yönlü ve güçlü bir ilişkiye sahip olduğu görülmektedir. Korelasyonun yüksek olması farklı indekslerle ölçüm sonucunda benzer okunabilirlik sonuçlarının alındığını ve dolayısıyla okunabilirlik ölçümlerinin tutarlı olduğunu göstermektedir. Bununla birlikte tüm indeks skorları ile dosya boyutu arasındaki ilişkinin istatistiksel olarak anlamlı ve ters yönlü olduğu görülmektedir. Buna göre dosya boyutu (MB) arttıkça tüm indeks türlerinde okunabilirlik skoru azalmaktadır.

Tablo 10

Şirket Büyüklüğüne Göre Raporlanan Form Sayıları

Form Türü	Şirket Büyüklüğü				Toplam (n=144)
	Mikro (n=38)	Küçük (n=29)	Orta (n=34)	Büyük (n=43)	
10-K	328	232	274	374	1208
10-Q	990	692	823	1122	3627
Form 3	89	147	130	220	586
Form 4	435	1186	1097	1931	4649
8-K	3562	2292	2721	4458	13033
Toplam	5404	4549	5045	8105	23103

Çalışma kapsamında şirket büyüklüğünün, şirketlerin kamuya açıkladıkları finansal formların okunabilirlik düzeyi ile ilişkili olup olmadığının tespiti amacıyla teknoloji alanında faaliyet gösteren ve NYSE’de işlem gören toplam 144 şirket piyasa değerine göre sınıflandırılmıştır. Bu doğrultuda, şirket büyüklüklerine ilişkin büyüklük sınıflarının eşik değerleri belirlenirken persentillere dayalı piyasa değerleri kullanılmıştır. Buna göre, 0 – 300 milyon dolar piyasa değerine kadar olan şirketler “Mikro”, 301 milyon dolar – 2 milyar dolara kadar olan şirketler “Küçük”, 2 – 10 milyar dolar “Orta”, 10 milyar dolar ve üstü şirketler ise “Büyük” olarak etiketlenmiştir. Şirket büyüklüğüne göre incelenen 23.103 formun dağılımı Tablo 10’da verilmiştir.

Tablo 11’de gösterildiği üzere, büyük şirketlerde içeriden öğrenenler statüsüne sahip yönetici ve çalışan sayısının daha fazla olduğu veya söz konusu statüye sahip kişilerin sahip oldukları menkul kıymetlerdeki değişikliklerin daha sık gerçekleştiği / raporlandığı varsayımından hareketle, raporlanan Form 3 ve Form 4 sayısı büyük şirketlerde daha yüksek olarak izlenmektedir.

Tablo 11

Şirket Büyüklüğüne Göre Farklı Okunabilirlik İndekslerine Ait Skor ve Tanımlayıcı İstatistikler

İndeks Türü	Form Türü	Şirket Büyüklüğü							
		(a) Mikro		(b) Küçük		(c) Orta		(d) Büyük	
		Ortalama Skor	S.S.	Ortalama Skor	S.S.	Ortalama Skor	S.S.	Ortalama Skor	S.S.
Flesch-Kincaid	10-K	15,97 (b, c, d)	0,98	16,35 (a, c)	0,90	16,60 (a, b)	0,91	16,50 (a, b)	0,90
	10-Q	16,29 (c, d)	3,95	16,40 (c, d)	4,63	16,58 (a, b)	1,50	16,71 (a, b)	1,52
	Form 3	54,85 (b, d)	49,67	39,38 (a, d)	21,70	33,33 (a)	11,86	32,95 (a, b)	17,67
	Form 4	76,34 (b)	50,04	57,72 (a, c, d)	49,74	78,89 (b)	67,69	67,80 (b)	42,89
	8-K	14,50	3,77	14,46 (d)	3,88	14,54	3,80	14,50 (b)	3,43
Coleman	10-K	22,24 (b, c, d)	0,42	22,59 (a, c)	0,56	22,71 (a, b)	0,65	22,69 (a, b)	0,7
	10-Q	22,21 (c, d)	0,81	22,29 (c, d)	1,16	22,69 (a, b)	1,47	22,60 (a, b)	1,39
	Form 3	37,01 (b, c, d)	15,62	27,20 (a, d)	10,59	24,85 (a)	8,36	25,20 (a, b)	9,52
	Form 4	57,93 (b, c)	17,03	39,25 (a, c, d)	21,88	53,81 (a, d)	19,54	56,36 (b, c)	16,58
	8-K	22,09	1,32	22,14 (d)	1,62	22,07	1,35	22,20 (b)	1,29
Gunning Fog	10-K	20,17 (b, c, d)	1,02	20,36 (a, c)	0,93	20,61 (a, b)	0,95	20,46 (a, b)	0,94
	10-Q	20,30 (c, d)	3,99	20,30 (c, d)	4,72	20,45 (a, b)	1,44	20,50 (a, b)	1,51
	Form 3	55,38 (d)	49,27	42,45 (d)	20,00	36,99	10,88	36,45 (a, b)	16,11
	Form 4	71,54 (b)	49,35	57,42 (a, c, d)	46,67	75,20 (b, d)	65,77	63,15 (b, c)	41,26
	8-K	18,89	3,79	18,77 (d)	3,95	18,89	3,88	18,85 (b)	3,47
ARI	10-K	22,27 (b, c, d)	1,27	22,69 (a, c)	1,23	22,95 (a, b)	1,30	22,89 (a, b)	1,23
	10-Q	22,86 (c, d)	5,36	23,15 (c, d)	6,33	23,21 (a, b)	2,02	23,34 (a, b)	2,00
	Form 3	75,31 (c, d)	64,27	53,31 (d)	29,10	45,35 (a)	16,06	44,80 (a, b)	23,44
	Form 4	105,53 (b)	64,53	78,34 (a, c, d)	65,89	107,96 (b)	88,07	94,14 (b)	55,92
	8-K	21,37 (d)	4,97	21,43 (d)	5,11	21,39	4,95	21,37 (a, b)	4,40
Smog	10-K	17,50 (b, c, d)	0,72	17,64 (a, c)	0,62	17,82 (a, b)	0,65	17,71 (a, b)	0,66
	10-Q	17,52 (c, d)	1,41	17,50 (c, d)	1,55	17,68 (a, b)	0,95	17,71 (a, b)	0,99
	Form 3	33,20 (d)	16,96	28,66 (d)	8,77	26,12	5,18	25,93 (a, b)	7,59
	Form 4	43,79 (b)	20,32	36,10 (a, c, d)	20,28	45,01 (b)	24,61	41,32 (b)	17,97
	8-K	16,35	2,42	16,26 (d)	2,51	16,37	2,47	16,32 (b)	2,25

Not: Tabloda parantez içinde ve italik olarak verilen harfler, farklı harflerle kategorize edilmiş farklı büyüklüklerdeki şirket türleri için Dunn'ın (1964) sıra toplamlarına dayalı çoklu karşılaştırma testi sonucunda birbirinden istatistiksel olarak anlamlı derecede ($p < 0,01$) farklılaşan grupları göstermektedir.

** $p < 0,01$

Şirket büyüklüklerine göre her bir form türü farklı okunabilirlik indeksi skorları bağlamında değerlendirildiğinde, 10-K

formu okunabilirlik skorlarının küçük ve büyük şirketler ile orta ve büyük şirketlerde anlamlı derecede farklılaşmadığı, diğer tüm karşılaştırmalar için ise istatistiksel olarak anlamlı bir fark olduğu sonucuna ulaşılmıştır. Bu durumda örneğin, mikro şirketler ile küçük şirketlerin 10-K formu okunabilirlik skorları istatistikî olarak anlamlı derecede farklılaşmaktadır. Aynı şekilde, mikro ve orta ölçekteki, mikro ve büyük ölçekteki, küçük ve orta ölçekteki işletmelerin raporladıkları 10-K formlarının okunabilirlik düzeyi %99 güven düzeyinde farklılaşmaktadır. Buna göre okunabilirliği en zor olan ve yüksek lisans üstü seviye gerektiren 10-K formları orta ve büyük ölçekli işletmeler tarafından üretilmektedir. Orta ve büyük, mikro ve küçük şirketlerde 10-Q formları için okunabilirlik skorları anlamlı derecede farklılaşmamakta olup, diğer şirket ölçeği karşılaştırmalarında anlamlı fark görünmektedir. Bu durumda, orta ve büyük ölçekli şirketlerde 10-Q formlarının okunabilirlik düzeyinin diğer şirketlere göre daha yüksek olduğu ve söz konusu formların çok zor anlaşılabilir olarak nitelendirilebileceği söylenebilir. 8-K formu açısından sadece küçük ve büyük işletmeler ile mikro ve büyük işletmelerin tüm indekslerdeki okunabilirlik skorları istatistiksel olarak anlamlı derecede farklılaşmaktadır. Tüm indeks türlerinde küçük ve büyük ölçekteki işletmelerin 8-K formu okunabilirlik skorları arasında anlamlı bir fark tespit edilmiştir.

İçeriden öğrenenlerin raporladıkları Form 3'lerin okunabilirlik düzeylerine bakıldığında, mikro şirketlerce raporlanan formların diğer ölçekteki şirketlerin raporladıkları formlardan anlamlı derecede yüksek olduğu göze çarpmaktadır. Ayrıca, küçük ölçekteki şirketlerin raporladıkları Form 3'lerin de "çok zor" seviyede okunabilir olduğu ve okunabilirlik skorlarının diğer ölçekteki şirketlerden farklılaştığı tespit edilmiştir. Buna göre, mikro ve küçük ölçekteki şirketlerdeki yönetici, çalışan ve hak sahiplerinin gerçekleştirdikleri işlemleri raporladıkları formun okunabilirlik düzeyi, orta ve büyük ölçekteki şirketlerde raporlanan formların okunabilirlik düzeyinden anlamlı derecede yüksek bulunmuştur. Bütün indeks türlerinde orta ve büyük şirketlerdeki okunabilirlik skorları arasında anlamlı bir fark görülmemektedir.

Aynı şekilde, Form 4'e ilişkin okunabilirlik skorları değerlendirildiğinde tüm ölçeklerdeki şirketlerde raporlanan formların "çok zor" düzeyde olduğu dikkat çekmektedir. İçeriden öğrenenlerin kendi varlıklarında meydana gelen değişiklikleri raporladıkları Form 4'ün okunabilirlik düzeyi küçük ölçekteki işletmelerde "çok zor" seviyede olmasına rağmen, diğer ölçekteki şirketlerde raporlanan form skorlarının küçük ölçekteki şirkette raporlanan form skorlarından anlamlı derecede yüksek olduğu sonucuna ulaşılmıştır.

6. SONUÇ VE TARTIŞMA

Okunabilirlik indeksleri eğitim bilimleri başta olmak üzere, anket çalışmaları, raporlar ve genel olarak bir metnin anlaşılabilirliğini ölçmek amacıyla gerçekleştirilen geniş bir uygulama alanına sahiptir. Çalışmada da gösterildiği üzere, okunabilirlik skorunun elde edilmesinde birçok farklı indeks kullanılmakta olup, anlamsal zorluk ve söz dizimsel karmaşıklık için her indeksin hesaplamasında genel olarak ortalama kelime ve cümle uzunluğuna dayalı skorlar üretilmektedir.

Bu çalışmada ele alınan konu bağlamında araştırma örneğine 01.01.2010 - 01.10.2019 tarihleri arasında NYSE'de işlem gören 144 teknoloji şirketi alınmış olup, toplamda 23.103 adet finansal kamuya açıklama formu incelenmiştir. Yapılan analiz sonucunda, şirketlerin olağan genel kurullarından önce hissedarlarıyla paylaştıkları ve şirketin finansal durumunun ayrıntılı olarak raporlandığı 10-K ve 10-Q formlarındaki karmaşık (kompleks) kelime sayısının yüksek olduğu tespit edilmiştir. Bu durum, teknoloji alanında faaliyet gösteren şirketlerin hissedarları ve yatırımcılarının şirket tarafından raporlanıp kamuya açıklanan formları anlayabilmesi için en az lisansüstü düzeyinde bir eğitim seviyesi veya uzmanlığa sahip olmaları gerektiğine işaret etmektedir. Finans alan yazını içerisinde değerlendirildiğinde bu sonuç, Bonsall vd. (2017) ve Lo vd. (2017) tarafından NYSE'de işlem gören şirketlerin yıllık faaliyet raporlarının okunabilirlik seviyesinin tespiti üzerine gerçekleştirilen çalışmanın sonuçlarıyla uyumlu bulunmuştur. Buna göre, ilgili formların çok zor okunabilirlik seviyesinde olması piyasalarda bilgi asimetrisi doğurabileceği gibi, menkul kıymet fiyat tahmini, kredi ve risk derecelendirme, borçlanma maliyeti, likidite ve işlem düzeyi, işletme karlılığı ve yatırım etkinliğini de olumsuz yönde etkileyecektir (Hasan, 2018). 10-K ve 10-Q formları şirket büyüklüklerine göre değerlendirildiğinde ise, orta ve büyük ölçekli teknoloji işletmelerinde üretilen formların genel olarak diğer ölçekteki işletmelerin kamuya açıkladığı formlardan daha zor anlaşılabilir bir düzeyde olduğu sonucuna ulaşılmıştır. Bu durum, kamuya açıklanan formların "Sade İngilizce" kuralına uygunluğu başta olmak üzere, Lo vd. (2017) ve Hesarzadeh v.d. (2018) çalışmalarında şirket karlılığı, finansal yönetim etkinliği, kurumsal iletişim ve denetim gibi

konularla ilişkilendirilen şirket büyüklüğüyle birlikte, finansal form okunabilirliğinin teknoloji alanında faaliyet gösteren işletmeler için ayrıntılı olarak değerlendirilmesi gerektiğini göstermektedir. Nitekim, Li (2008)'de belirtildiği üzere, şirket büyüklüğü iş operasyonlarındaki ve organizasyondaki karmaşıklığı arttırdığı için daha zor okunabilirlik seviyesinde finansal raporlar üretilmesine sebep olmaktadır. Kim v.d. (2018) çalışmasında ortaya konan, şirket finansal formlarının okunabilirlik düzeyiyle menkul kıymet fiyatlarındaki düşüşün pozitif ilişkisi de göz önünde bulundurulduğunda, teknoloji şirketlerinin finansal piyasalardaki değer artışının yanı sıra; finansal raporların okunabilirlik düzeyindeki artışın menkul kıymet fiyatları açısından bir risk oluşturduğu düşünülebilir. Ayrıca, hissedarları ilgilendiren önemli şirket olaylarının raporlandığı 8-K formlarının da okunabilirlik seviyesinin çok zor olduğu sonucuna ulaşılması, yukarıda belirtilen hususu güçlendirmektedir.

Şirket içerisinde menkul kıymet fiyatlarını etkileme gücüne sahip, ayrıcalıklı veriye erişimi bulunan yönetici ve çalışanların finansal piyasalardaki işlemlerinin kamuya açıklandığı Form 3 ve Form 4'ün okunabilirlik düzeyine bakıldığında ise yatırımcıların, yönetici ve çalışanların menkul kıymetlerinde meydana gelen değişiklikleri okuyup anlamada güçlük çektiği ve yatırımcıların yine en az lisansüstü eğitim düzeyine sahip olması gerektiği göze çarpmaktadır. Bu durum Loughran ve McDonald (2013) çalışmasında da belirtildiği şekilde, teknoloji şirketlerine yatırım yapmak isteyen yatırımcıların şirketin varlıklarına ve kurumsal yönetimine ilişkin risk karakteristiklerini doğru bir şekilde değerlendirememesi ve bununla birlikte gelen menkul kıymet fiyatlarındaki oynaklık problemini ortaya çıkaracaktır. Form 3 ve Form 4 okunabilirlik düzeyinin mikro boyuttaki işletmelerde çok daha yüksek olması ise, söz konusu şirketlerdeki yönetici ve çalışanların finansal piyasalarda gerçekleştirdikleri işlemlerin yatırımcılar tarafından anlaşılmasının çok zor olduğuna işaret etmektedir. Neill vd. (2008) tarafından belirtildiği şekilde, içeriden öğrenenlerin anormal getirilerinin küçük ölçekli işletmelerde daha yüksek olduğu göz önüne alındığında, söz konusu işletmelerdeki içeriden öğrenenlerin Form 3 ve Form 4 gibi işlem raporlarının okunabilirlik düzeyinin finansal getirilerle değerlendirilmesi hususu önem kazanmaktadır.

Genel olarak çalışmada incelenen finansal formların okunabilirlik düzeyinin “çok zor” olması, diğer bir ifadeyle, söz konusu formların yatırımcılar tarafından anlaşılabilmesi, finansal bilgilerin belirsizliğini ve karmaşıklığını arttırmakta olup, piyasalarda asimetrik bilgi sorununu doğuran önemli etkenlerden biridir. Anlaşılması zor ve bilgi eksikliği yaratan finansal formlar yatırımcıların piyasalardaki davranışlarını bozma eğilimi yaratacağından söz konusu formlar yatırımcılar tarafından güvenilir referanslar olarak kabul edilmeyecektir. Bu durum ayrıca, piyasalarda şeffaflığa engel olmakla birlikte, yatırım verimliliği, kurumsal yönetim, şirket etkinliği gibi piyasayı ve yatırımcıları ilgilendiren faktörleri de olumsuz yönde etkileyecektir. Piyasada ortaya çıkabilecek bilgi asimetrisi problemlerine yönelik olarak, şirketlerce yayınlanan rapor ve formların şekil ve içeriği konusunda getirilen standartların geliştirilmesi gerekmekte olup, şirketlerin faaliyet gösterdikleri teknolojik alana ilişkin dilin sadeleştirilerek anlaşılır bir seviyeye getirilmesi, teknik terimlerin form ve raporlarda, gerekirse, açıklama veya dipnot şeklinde verilmesi önerilmektedir.

Yukarıda da belirtildiği üzere mevcut literatürde, şirketlerin kamuyu aydınlatma amacıyla resmî kurumlara bildirdikleri metin dosyasının boyutu, okunabilirliğin bir ölçüsü olarak kabul edilmektedir. Söz konusu ölçüt veri ayrıştırma ve metin çıkarma gibi işlemleri gerektirmemektedir. Bununla birlikte, dosya boyutunun diğer indekslerle korelasyonuna bakıldığında negatif yönde anlamlı ve düşük bir ilişkinin olduğu tespit edilmiştir. Loughran ve McDonald (2014) tarafından belirtildiği üzere, şirketlerin raporladıkları dosya boyutunun kurumsal bilgi ortamı ve menkul kıymet fiyatlarındaki volatilitenin negatif yönde etkilediği göz önünde bulundurulduğunda ise özellikle 10-K ve 10-Q formlarının yüksek boyutta olduğu görülmektedir. Finans bakış açısıyla değerlendirildiğinde bu durum, okunabilirlik indeksleriyle şirketlerin finansal performans göstergeleri arasındaki ilişkinin incelenmesinin faydalı olabileceğini düşündürmektedir.

Diğer taraftan, okunabilirlik indeksleriyle cümle uzunluklarının doğru bir şekilde ölçülmesi amaçlandığı için, form içerisinde bulunan bağlaçlar, noktalama işaretleri, kısaltmalar, bölüm başlıkları, matematiksel semboller ve sektöre ilişkin terminolojik kelimeler, cümle sınırlarının belirlenmesini zorlaştırmaktadır. Bu sebeple, belirli bir işlem yüküne sahip okunabilirlik skoru hesaplama sürecini optimize etmek amacıyla, analiz edilecek finansal dokümanların tanımlanması ve içeriğin anlaşılabilmesinde hibrit doğal dil işleme tekniklerinden yararlanılmasının yazım yanlışları ve imla hatalarının düzeltilerek uzun cümlelerin okunabilirliğinin artırılmasına da katkı sağlayacağı düşünülmektedir. Ayrıca söz konusu tekniklerin kullanımı, doküman içerisinde bulunan liste, tablo veya farklı türdeki yapılandırılmamış veri alanlarından hiçbir kısıta bağlı kalmadan anlamlı

ifadelerin elde edilmesinde faydalı olacaktır. Günümüzde işletmelerin kamuya açıkladığı veya işletme faaliyetlerine yönelik ürettiği doküman miktarı düşünüldüğünde, dokümanlar içerisindeki anlamlı bilgilerin çıkarılmasına yönelik olarak özellik seçme, boyut indirgeme gibi işletme maliyetlerini azaltıcı makine öğrenmesi ve derin öğrenme yaklaşımlarının geliştirilmesi önerilmektedir.

Çalışmada, hali hazırda Türkiye örneğinde formların doldurulmasında dikkat edilmesi gerekli olan dilbilimsel standartlara ilişkin referans alınabilecek bir düzenleyici kural veya doküman bulunmaması ve metin madenciliği için finans alanında oluşturulmuş teknik terimler sözlüğü olmaması nedeniyle, uygulama NYSE örnekleme üzerinden gerçekleştirilmiştir. Hem farklı endekslerle ölçümlenen okunabilirlik düzeyleri karşılaştırılmış, hem de şirket ölçeği düzeyinde bir karşılaştırma yapılmıştır. Çalışmada ayrıca, ilerde planlanabilecek Türkiye örneklemleri için şirket ölçeklerine göre okunabilirlik skorlarının değerlendirilmesinin önemine de dikkat çekilmektedir. Çalışma kapsamında ele alınan ve incelenen finansal formların okunabilirlik düzeyi ile ilgili ABD’de SEC tarafından yapılan düzenlemelerin benzerlerinin Borsa İstanbul’da işlem gören şirketlerin kamuya yayınlamakla yükümlü olduğu finansal formlar için de uygulanmasının yerli yatırımcıları korumak ve doğru finansal kararlar alabilmelerini sağlayabilmek adına yasal bir zemine oturtulması önerilmektedir. Okunabilirlik düzeyi ile ilgili Türkçe finansal metinler üzerine çalışmalar gerçekleştirilebilmesi için terimler sözlüğünün oluşturulması da önemlidir.

Hakem Değerlendirmesi: Dış bağımsız.

Çıkar Çatışması: Yazarlar çıkar çatışması bildirmemiştir.

Finansal Destek: Yazarlar bu çalışma için finansal destek almadığını beyan etmiştir.

Peer-review: Externally peer-reviewed.

Conflict of Interest: The authors have no conflict of interest to declare.

Grant Support: The authors declared that this study has received no financial support.

Kaynaklar/References

- Adalı, A. (2018). *Sosyal Güvenlik Mevzuatının Okunabilirlik Düzeyi ve Buna İlişkin Bir Uygulama*. (Yayımlanmamış Y.L. Tezi), Pamukkale Üniversitesi Sosyal Bilimler Enstitüsü, Denizli.
- Aggarwal, C. C., Zhai, C. (2012). An Introduction to Text Mining. C.C. Aggarwal, C. Zhai, içinde *Mining Text Data* (S. 1-10), Boston: Springer.
- Al Tamimi, A. K., Jaradat, M., Al-Jarrah, N., Ghanem, S. (2014). AARI: Automatic Arabic Readability Index, *The International Arab Journal of Information Technology*, 11(4), 370-378.
- Ateşman, E. (1997). Türkçede Okunabilirliğin Ölçülmesi, *Dil Dergisi*, 58, 171-174.
- Bailin, A., Grafstein, A. (2016). *Readability: Text And Context*, Boston: Springer.
- Biddle, G., Hilary, G., Verdi, R. (2009). How Does Financial Reporting Quality Relate to Investment Efficiency?, *Journal of Accounting And Economics*, 48, 112-31.
- Bloomfield, R. (2008). Discussion of Annual Report Readability, Current Earnings, and Earnings Persistence, *Journal of Accounting and Economics*, 45(2-3), 248-252.
- Bonsall Iv, S. B., Leone, A. J., Miller, B. P., Rennekamp, K. (2017). A Plain English Measure of Financial Reporting Readability. *Journal of Accounting and Economics*, 63(2-3), 329-357.
- Boubaker, S., Gounopoulos, D., Rjiba, H. (2019). Annual Report Readability and Stock Liquidity, *Financial Markets, Institutions and Instruments*, 28(2), 159-186.
- Bushee, B.J., Gow, I.D., Taylor, D.J. (2018). Linguistic Complexity İn Firm Disclosures: Obfuscation or Information?, *Journal of Accounting Research*, 56(1), 85-121.
- Cassell, C. A., Cunningham, L. M., Lisic, L. L. (2019). The Readability of Company Responses to SEC Comment Letters and SEC 10-K Filing Review Outcomes, *Review of Accounting Studies*, 24(4), 1252-1276.
- Kim, C., Wang, K., Zhang, L. (2019). Readability of 10-K Reports and Stock Price Crash Risk. *Contemp Account Res*, 36, 1184-1216.
- Chakrabarty, B., Seetharaman, A., Swanson, Z., Wang, X. (2018). Management Risk Incentives and the Readability of Corporate Disclosures, *Financial Management*, 47(3), 583-616.
- Chakraborty, G., Pagolu, M., Garla, S. (2013). *Text Mining and Analysis: Practical Methods, Examples, and Case Studies Using SAS*, North Carolina: Sas Institute Inc.
- Chandler, S.B. (2001). *Legibility And Comprehension of Onscreen Type*, (Yayımlanmamış Doktora Tezi), Virginia Polytechnic Institute And State University, Blacksburg, Virginia, Usa.
- Coleman, M., Liau, T. L. (1975). A Computer Readability Formula Designed for Machine Scoring, *Journal of Applied Psychology*, 60(2), 283-284.
- Çakın, T. (2010). *Yatırımcıların Korunması Kapsamında Aracılık Faaliyetlerinde İş Yapış Kurallarına İlişkin Yurtdışı Düzenlemeler ve Türkiye İçin*

- Öneriler, 18 Eylül 2020 Tarihinde SPK Hukuk İşleri Dairesi: <https://www.spk.gov.tr/siteapps/yayin/yayingoster/1156> Adresinden Alındı.
- De Franco, G., Hope, O. K., Vyas, D., Zhou, Y. (2015). Analyst Report Readability, *Contemporary Accounting Research*, 32(1), 76-104.
- Dehaan, E., Song, Y., Xie, C., Zhu, C., (2020). *Disclosure Obfuscation in Mutual Funds*, 18 Eylül 2020 Tarihinde [ssrn: https://ssrn.com/abstract=3540215](https://ssrn.com/abstract=3540215) Adresinden Alındı.
- Di Marco, J. (2010). *Digital Design for Print and Web: An Introduction to Theory, Principles and Techniques*, New Jersey: John Wiley & Sons, Inc.
- Dolphin, R., Wagley, R. A. (1977). Reading the Annual Report, *Financial Executive*, 45(6), 20–22.
- Drake, M. S., Roulstone, D. T., Thornock, J. R. (2016). The Usefulness of Historical Accounting Reports, *Journal of Accounting And Economics*, 61(2-3), 448-464.
- Dubay, W. H. (2004). *The Principles of Readability*. Costa Mesa: Impact Information.
- Dunn, O. J. (1964). Multiple Comparisons Using Rank Sums, *Technometrics*, 6(3), 241-252.
- Dyer, T., Lang, M., Stice-Lawrence, L. (2017). The Evolution of 10-K Textual Disclosure: Evidence From Latent Dirichlet Allocation, *Journal of Accounting and Economics*, 64(2-3), 221-245.
- Fakhfakh, M. (2015). The Readability of International Illustration of Auditor's Report: An Advanced Reflection on the Compromise Between Normative Principles and Linguistic Requirements, *Journal of Economics, Finance and Administrative Science*, 20(38), 21-29.
- Finviz (2020). *Financial Visualizations*, 21 Temmuz 2020 Tarihinde Finviz: <https://finviz.com/groups.Ashx?G=industryvesg=technologyvev=210veo=Marketcap> Adresinden Alındı.
- Fisher, I. E., Garnsey, M. R., Hughes, M. E. (2016). Natural Language Processing in Accounting, Auditing and Finance: A Synthesis of the Literature with A Roadmap for Future Research, *Intelligent Systems in Accounting, Finance and Management*, 23(3), 157-214.
- Flesch, R. (1948). A New Readability Yardstick, *Journal of Applied Psychology*, 32, 221-233
- Gillick, D. (2009). Sentence Boundary Detection and the Problem with the Us, *Proceedings of Human Language Technologies: The 2009 Annual Conference of the North American Chapter of The Association for Computational Linguistics, Companion Volume: Short Papers* İçinde, (s. 241-244).
- Guay, W., Samuels, D., Taylor, D. (2016). Guiding through the Fog: Financial Statement Complexity and Voluntary Disclosure, *Journal of Accounting and Economics*, 62(2-3), 234-269.
- Gunning, R. (1952). *The Technique of Clear Writing*, New York: Mcgraw-Hill.
- Güyer T., Temur T. Solmaz E. (2009). Bilgisayar Destekli Metin Okunabilirliği Analizi, *Türk Eğitim Bilimleri Dergisi*, 7(4), 751-766.
- Hasan, M.M. (2020) Readability of Narrative Disclosures in 10-K Reports: Does Managerial Ability Matter?, *European Accounting Review*, 29:1, 147-168
- Hesarzadeh, R., Bazrafshan, A. (2018). Corporate Reporting Readability and Regulatory Review Risk, *Baltic Journal of Management*, 13(4), 488-507.
- Hornik, K., Hornik, M. K. (2018). *Package 'Nlp'*, 21 Eylül 2020 Tarihinde Cran: [Http://cran.salud.gov.sv/web/packages/Nlp/Nlp.Pdf](http://cran.salud.gov.sv/web/packages/Nlp/Nlp.Pdf) Adresinden Alındı.
- Jha, A. (2019). Financial Reports and Social Capital, *Journal of Business Ethics*, 155(2), 567-596.
- Keskin, D. A., Tutar, S., Öktem, B., Akçay, B. (2020). ARI Okunabilirlik Endeksine göre Cinsiyetin Kilit Denetim Konularının Okunabilirliği Üzerine Etkisi: Sektörel Bir İnceleme, *Marmara Üniversitesi Sosyal Bilimler Enstitüsü Öneri Dergisi*, 15(53), 209-228.
- Kim, C., Wang, K., Zhang, L. (2019). Readability of 10-K Reports and Stock Price Crash Risk, *Contemporary Accounting Research*, 36(2), 1184-1216.
- Kincaid, J. P., Fishburne, R. P., Rogers, R. L., Chissom B. S. (1975). Derivation of New Readability Formulas (Automated Readability Index, Fog Count, and Flesch Reading Ease Formula) for Navy Enlisted Personnel, *Research Branch Report*, 8-75.
- Klare, G. R. (1963). *The Measurement of Readability*, Ames: Iowa State University Press.
- Klare, G. R. (1984). *Readability. Handbook of Reading Research*, New York: Longman.
- Lawrence, A. (2013). Individual Investors and Financial Disclosure, *Journal of Accounting and Economics*, 56(1), 130-147.
- Lee, Y. J. (2012). The Effect of Quarterly Report Readability on Information Efficiency of Stock Prices, *Contemporary Accounting Research*, 29(4), 1137-1170.
- Lehavy, R., Li, F., Merkley, K. (2011). The Effect of Annual Report Readability on Analyst Following and The Properties of Their Earnings Forecasts, *The Accounting Review*, 86(3), 1087–1115.
- Li, F. (2008). Annual Report Readability, Current Earnings, and Earnings Persistence, *Journal of Accounting and Economics*, 45(2-3), 221-247.
- Lim, E. K., Chalmers, K., Hanlon, D. (2018), The Influence of Business Strategy on Annual Report Readability, *Journal of Accounting and Public Policy*, 37(1), 65-81.
- Lo, K., Ramos, F., & Rogo, R. (2017). Earnings management and annual report readability. *Journal of Accounting Research*, 54(4), 1187–1230.
- Loughran, T., Mcdonald, B. (2010). *Measuring Readability in Financial Text*, 21 Temmuz 2020 tarihinde: http://securitieseditor.com/Wp/Wp-Content/Uploads/2014/05/Plain_English_V5.Pdf adresinden alındı.
- Loughran, T., Mcdonald, B. (2013). Ipo First-Day Returns, Offer Price Revisions, Volatility, And Form S-1 Language. *Journal Of Financial Economics*, 109: 307–26.
- Loughran, T., Mcdonald, B. (2014). Measuring Readability in Financial Disclosures, *The Journal of Finance*, 69(4), 1643-1671.
- Loughran, T., Mcdonald, B. (2016). Textual Analysis in Accounting and Finance: A Survey, *Journal of Accounting Research*, 54(4).1187-1230.
- Lundholm, R. J., Rogo, R., Zhang, J. L. (2014). Restoring the Tower of Babel: How Foreign Firms Communicate with US Investors, *The Accounting Review*, 89(4), 1453-1485.
- Luo, J. H., Li, X., Chen, H. (2018). Annual Report Readability and Corporate Agency Costs, *China Journal of Accounting Research*, 11(3), 187-212.
- Lupton, E. (2014). *Type On Screen: A Critical Guide for Designers, Writers, Developers, and Students*. New York: Princeton Architectural Press.

- Lusardi, A., Tufano, P. (2009). *Debt Literacy, Financial Experiences, and over in indebtedness*, 21 Temmuz 2020 Tarihinde National Bureau of Economic Research: <https://www.nber.org/papers/w14808> adresinden alındı.
- McLaughlin, G. H. (1969). Smog Grading-A New Readability Formula, *Journal of Reading*, 22, 639-646.
- Miller, B. P. (2010). The Effects of Reporting Complexity on Small and Large Investor Trading, *The Accounting Review*, 85(6), 2107-2143.
- Neill, B. J., Sadeghi, M., & Watt, E. (2008). Are insider trades profitable? Evidence from directors' trade on the Australian stock exchange. *Corporate Ownership & Control*, 5(3-1), 176-187.
- NY Times (2019, 11 Aralık). *Big tech is under attack, and investors couldn't care less*, 21 Temmuz 2020 Tarihinde Ny Times: <https://www.nytimes.com/2019/12/10/Business/Tech-Stocks-Apple-Amazon-Microsoft-Google.html> Adresinden Alındı.
- NYSE (2020). *Why Tech Companies Choose NYSE?*, 21 Temmuz 2020 Tarihinde NYSE: <https://www.nyse.com/Article/Technology-ipo/Listings/Why-Tech-Companies-Choose-Nyse> Adresinden Alındı.
- OECD (2017). *G20/OECD Infe Report on Adult Financial Literacy in G20 Countries*, 21 Temmuz 2020 Tarihinde OECD: <http://www.oecd.org/Daf/Fin/Financial-Education/G20-Oecd-Infe-Report-Adult-Financial-Literacy-In-G20-Countries.Pdf> adresinden Alındı.
- Özcan, O., Türü, T. E. (2004). Sermaye Piyasasında Yatırımcı Eğitiminin Önemi. *İktisat İşletme ve Finans*, 19(225), 58-72.
- Rennekamp, K. (2012). Processing Fluency and Investors' Reactions to Disclosure Readability, *Journal of Accounting Research*, 50(5), 1319-1354.
- Sec (1998), A Plain English Handbook: How To Create Clear Sec Disclosure Documents, 21 Temmuz 2020 Tarihinde U.S. Securities and Exchange Commission: <http://www.sec.gov/pdf/handbook.pdf> adresinden alındı.
- Smith, E. A., Senter, R. J. (1967). *Automated Readability Index*, 21 Temmuz 2020 tarihinde Cincinnati University: <https://apps.dtic.mil/dtic/tr/fulltext/U2/667273> adresinden alındı.
- Still, M. D. (1972). The Readability of Chairman's Statements, *Accounting and Business Research*, 3(9), 36-39.
- Tan, H. T., Wang, E. Y., Zhou, B. (2015). How Does Readability Influence Investors' Judgments? Consistency of Benchmark Performance Matters, *The Accounting Review*, 90(1), 371-393.
- Tekşan, K. (2016). Ömer Seyfettin'in Hikâyelerinin Okunabilirlik Düzeyleri, *Sosyal ve Liberal Bilimlerde Yeni Yönelimler*, 1, 25-36.
- Thomas, M. (2013). Looking ahead with Hope: Reviving The Reading Maturity Construct as Social Science for Adolescent and Adult Readers, *A Journal of Literacy and Language Arts*, 52(2), 4.
- Üzeler, R.S. (2014). *Yatırım Hizmeti Alımında Yatırımcının Korunması ve Tüketici Hukukunun Sermaye Piyasası Mevzuatında Uygulanabilirliği Sorunu*, 21 Temmuz 2020 Tarihinde SPK Hukuk İşleri Dairesi: <https://www.spk.gov.tr/siteapps/yayin/yayingoster/1117> adresinden alındı.
- Vajjala, S., Meurers, D. (2014). Readability Assessment for Text Simplification: From Analysing Documents to Identifying Sentential Simplifications, *International Journal Of Applied Linguistics*, 165(2), 194-222.
- You, H., Zhang, X. J. (2009). Financial Reporting Complexity and Investor Underreaction to 10-K Information, *Review of Accounting Studies*, 14(4), 559-586
- Zhou, S., Jeong, H., & Green, P. A. (2017). How consistent are the best-known readability equations in estimating the readability of design standards?. *IEEE Transactions on Professional Communication*, 60(1), 97-111.