



METİNSEL VERİ MADENCİLİĞİ İLE TÜRKİYE'DEKİ ÜNİVERSİTELERİN SÜRDÜRÜLEBİLİRLİK RAPORLARININ DEĞERLENDİRİLMESİ

EVALUATION OF SUSTAINABILITY REPORTS OF UNIVERSITIES IN TURKEY USING TEXTUAL DATA MINING

DOI:10.20854/bujse.1842519

Ülker Başar^{1*}, Merve Doğruel²

Öz

Bu çalışmanın amacı, Türkiye'deki üniversitelerin 2024 yılı sürdürülebilirlik raporlarındaki stratejik eğilimleri ve öncelik alanlarını, metin madenciliği teknikleri kullanarak nicel verilerle ortaya koymaktır. Çalışma, yapılandırılmamış metinsel verilerin analizi yoluyla yükseköğretimdeki sürdürülebilirlik söylemini objektif bir şekilde ölçmeyi hedeflemektedir. Araştırma kapsamında, 65 üniversiteye ait Türkçe sürdürülebilirlik raporları incelenmiş, 2024 yılında 26 üniversite tarafından yayımlanan sürdürülebilirlik raporları veri seti olarak belirlenmiştir. PDF formatındaki dokümanların, metin ön işleme ve tematik kodlama aşamaları MAXQDA ile yapılmış; terim frekansı ve TF-IDF hesaplamaları ise Python (pandas, scikit-learn) ile doğrulanmıştır. Çalışma, durak kelime (stop-word) temizliği ve normalizasyon işlemlerinin ardından Terim Frekans (TF) ve Kod Eş-Oluşum (Co-occurrence) analizlerine tabi tutulmuştur. Elde edilen veriler, Çevresel, Sosyal ve Yönetişim (ESG) kriterlerine dayalı bir sözlük (dictionary) yaklaşımıyla sınıflandırılmıştır. Toplam 19.446 adet anlamsal terimin analizi sonucunda, Türk üniversitelerinin sürdürülebilirlik yaklaşımının literatürdeki genel beklentinin aksine Sosyal (%45,71) > Çevresel (%39,39) > Yönetişim (%14,90) hiyerarşisinde şekillendiği tespit edilmiştir. "Enerji" (n=1.597) ve "Atık" (n=875) terimleri operasyonel odaklanmayı gösterirken; "Eğitim" ve "Sürdürülebilirlik" kavramları arasındaki güçlü ilişki (n=151), üniversitelerin ana misyonlarını sürdürülebilirlik hedefleriyle bütünleştirdiğini kanıtlamaktadır. Buna karşın, yönetim boyutundaki zayıflık, sürecin henüz stratejik planlama seviyesine tam entegre edilemediğini göstermektedir. Elde edilen bulgular, üniversitelerin sürdürülebilirliği sadece "Yeşil Kampüs" olarak değil, eğitim ve toplumsal katkı odaklı hibrit bir model olarak benimsediğini ortaya koymuştur.

Abstract

The purpose of this study is to quantitatively reveal the strategic tendencies and priority areas in the 2024 sustainability reports of Turkish universities using text mining techniques. The study aims to objectively measure the sustainability discourse in higher education by analyzing unstructured textual data. Within the scope of the research, Turkish sustainability reports from 65 universities were examined, and the sustainability reports published by 26 universities in 2024 were selected as the dataset. Text preprocessing and thematic coding for PDF documents were carried out using MAXQDA; term frequency and TF-IDF calculations were verified using Python (pandas, scikit-learn). Following stop-word cleaning and normalization, the study was subjected to Term Frequency (TF) and Code Co-occurrence analyses. The resulting data were classified using a dictionary approach based on Environmental, Social, and Governance (ESG) criteria. Analysis of 19,446 semantic terms revealed that the sustainability approach of Turkish universities is structured in the hierarchy Social (45.71%) > Environmental (39.39%) > Governance (14.90%), contrary to the general expectation in the literature. The terms "Energy" (n=1,597) and "Waste" (n=875) indicate an operational focus, while the strong relationship between "Education" and "Sustainability" (n=151) shows that universities are integrating their core missions with sustainability goals. In contrast, the weakness in the governance dimension indicates that the process has not yet been fully integrated into the strategic planning level. The findings reveal that universities adopt sustainability not only as a "Green Campus" but as a hybrid model focused on education and societal contribution.

Anahtar Kelime: metin madenciliği, sürdürülebilirlik raporları, ui greenmetric, yapılandırılmamış veri analizi, yükseköğretim.

Keywords: text mining, sustainability reports, ui greenmetric, unstructured data analysis, higher education.

^{1,*} Sorumlu Yazar: İstanbul Esenyurt Üniversitesi, İşletme ve Yönetim Bilimleri Fakültesi, Yönetim Bilişim Sistemleri Bölümü, ulkerbasar@esenyurt.edu.tr, <https://orcid.org/0009-0000-6720-4161>

² İstanbul Esenyurt Üniversitesi, İşletme ve Yönetim Bilimleri Fakültesi, Yönetim Bilişim Sistemleri Bölümü, mervedogruel@esenyurt.edu.tr, <https://orcid.org/0000-0003-2299-7182>

1. GİRİŞ

Birleşmiş Milletler’in 2030 Gündemi kapsamında belirlenen Sürdürülebilir Kalkınma Amaçları (SKA), karmaşık küresel sorunların çözümünde çok paydaşlı ve bütüncül bir yaklaşımı zorunlu kılmaktadır (World Sustainable Development Goals Platform, 2025). Bilgi üretimi ve inovasyon kapasiteleriyle bu sürecin merkezinde yer alan yükseköğretim kurumları; yalnızca eğitim faaliyetleriyle değil, aynı zamanda araştırma, toplumsal katılım ve kurumsal liderlik misyonlarıyla da 17 hedefin tamamının hayata geçirilmesinde kritik bir rol üstlenmektedir. Geleceğin liderlerini yetiştirerek sürdürülebilirlik ilkelerini müfredata entegre eden ve iklim eyleminden yoksulluğun azaltılmasına kadar pek çok alanda kanıta dayalı çözümler üreten üniversiteler, sürdürülebilirliğe yönelik toplumsal dönüşümün ana aktörleri konumundadır. Bu bağlamda, kampüs operasyonlarını optimize eden ve sürdürülebilirlik ilkelerini yönetim süreçlerine entegre eden üniversiteler; hükümetler, endüstri ve sivil toplum ile iş birliği yaparak yenilikçi uygulamaları ölçeklendiren birer "Yaşayan Laboratuvar" (Living Lab) potansiyeli taşıyan yapılar olarak görülmektedir (Stuckrath, Rosales-Carreón et. al., 2025).

Şeffaflık, hesap verebilirlik ve kurumsal performansın merkezî bir unsuru haline gelen sürdürülebilirlik raporlaması, kuruluşların Çevresel, Sosyal ve Yönetişim (ESG) taahhütlerini paydaşlarına sunmalarını sağlayan temel bir araçtır. Yatırımcılar, çalışanlar ve toplum nezdinde güven inşa eden bu raporlar; kuruluşların sürdürülebilirlik hedeflerini belirlemelerine, ilerlemelerini izlemelerine ve ESG kriterlerini stratejik planlama süreçlerine entegre etmelerine olanak tanımaktadır. Literatürdeki pek çok çalışma, yüksek kaliteli raporlama standartları ile kurumsal değer ve piyasa performansı arasında pozitif bir korelasyon olduğunu ortaya koymaktadır (Zennaro, Corazza, et. al., 2024). Günümüzde GRI, SASB ve AB’nin Kurumsal Sürdürülebilirlik Raporlama Direktifi (CSRD) gibi küresel çerçevelerin yaygınlaşması, raporların karşılaştırılabilirliğini ve güvenilirliğini artırmış; bu eğilimin bir yansıması olarak yükseköğretim kurumları da ESG performanslarını yıllık raporlar aracılığıyla belgelemeye başlamıştır (Global Reporting Initiative, 2025). Üniversiteler için bu raporlar, yalnızca yasal ve etik bir sorumluluğun yerine getirilmesi değil; aynı zamanda UI GreenMetric ve THE Impact Rankings gibi veri odaklı uluslararası sıralama sistemlerinde rekabet avantajı sağlayan stratejik birer veri kaynağı niteliğindedir (UI GreenMetric, 2025), (Times Higher Education, 2025).



Şekil 1. Sürdürülebilir Kalkınma için Küresel Amaçlar (World Sustainable Development Goals Platform, 2025).

Türkiye'de de yükseköğretim kurumları, küresel eğilimlere paralel olarak sürdürülebilirlik raporlamasına artan bir ilgi göstermektedir. Ancak, bu raporlamaların niceliği, niteliği ve standardizasyonu üzerine sistematik çok fazla inceleme yapılmamıştır. Türkiye'de üniversitelerin raporlama eğilimindeki artış, analiz edilmesi gereken verinin artışından dolayı zorluklar yaşanmaktadır. Standart bir veri tabanı şemasından yoksun, PDF formatında ve serbest metin düzeninde sunulan bu raporlar, "yapılandırılmamış veri" (unstructured data) niteliği taşımaktadır (Adnan & Akbar, 2019). Bu veri türü; optik karakter tanıma (OCR) kaynaklı hatalar, tutarsız biçimlendirmeler, alana özgü jargon ve dilsel belirsizlikler gibi gürültü unsurlarını barındırması nedeniyle işlenmesi oldukça karmaşık bir yapıya sahiptir (Adnan & Akbar, 2019).

Geleneksel kural tabanlı yöntemler veya insan gücüne dayalı manuel içerik analizleri, büyük boyuttaki ve karmaşık veri yığınlarını işlemeye yetersiz kalmaktadır. Mevcut yöntemlerin yaşadığı ölçeklenebilirlik (scalability) sorunları ve araştırmacı yanlılığı (bias) riski, verilerin güvenilir bir şekilde analiz edilmesini zorlaştırmakta; bu durum ise stratejik örüntülerin nesnel bir şekilde ortaya çıkarılmasını engelleyen ciddi bir "veri darboğazı" oluşturmaktadır.

Bilgisayar bilimlerindeki gelişmeler, özellikle Metin Madenciliği ve Doğal Dil İşleme (NLP) teknikleri, yapılandırılmamış veri yığınlarından "bilgi keşfini" (knowledge discovery) mümkün kılarak karar alma süreçlerini dönüştürmektedir. Bu algoritmalar, metinsel verileri matematiksel vektör uzaylarına taşıyarak; kelime frekanslarını, gizli anlamsal ilişkileri ve tematik örüntüleri sayısal olarak modelleme yeteneğine sahiptir. Son yıllarda Derin Öğrenme ve Büyük Dil Modelleri (BERT, GPT vb.) ile kaydedilen ilerlemeler, metin madenciliğinde otomasyonu, doğruluğu ve bağlam farkındalığını artırarak manuel çabayı minimize etmiştir.

Ancak, bu evrensel modellerin uygulanmasında dile ve alana özgü zorluklar devam etmektedir. Özellikle Türkçe gibi sondan eklemeli (agglutinative) dillerde, kelime köklerinin ve morfolojik yapıların doğru analiz edilmesi, raporlardaki anlamsal bütünlüğün korunması açısından kritik bir mühendislik problemi olarak öne çıkmaktadır. Tüm bu teknik zorluklara rağmen NLP yöntemleri; nitel verilerin nicel metriklere dönüştürülmesini sağlayarak, kuruluşların ve araştırmacıların yapılandırılmamış veri yığınlarından eyleme geçirilebilir, ölçülebilir ve karşılaştırılabilir iç görüler elde etmesine olanak tanımaktadır.

Literatür incelendiğinde Türkiye'deki yükseköğretim kurumlarının sürdürülebilirlik raporlama pratiklerinin henüz başlangıç aşamasında olduğu görülmektedir. Yükseköğretim Kurulu (YÖK) 2025 yılı verilerine göre Türkiye'de 208 üniversite bulunmasına rağmen, bunların yalnızca 65'inin erişilebilir bir sürdürülebilirlik raporu mevcuttur. Mevcut araştırmalar, üniversitelerin stratejik planlarının Birleşmiş Milletler Sürdürülebilir Kalkınma Amaçları (SKA) ile entegrasyonunun zayıf olduğunu; raporlamaların ise genellikle uluslararası standartlara uygun, denetlenmiş belgelerden ziyade web sitesi duyuruları ile sınırlı kaldığını ortaya koymaktadır.

Bu alandaki en temel eksiklik ise metodolojik yaklaşımlarda göze çarpmaktadır. Türkiye özelinde yapılan az sayıdaki çalışma, çoğunlukla durum tespiti yapan betimsel analizlere dayanmaktadır. Türkçe raporların zengin morfolojik yapısının NLP teknikleriyle çözümlendiği ve rapor içerik derinliğinin UI GreenMetric gibi uluslararası performans skorlarıyla algoritmik olarak ilişkilendirildiği kapsamlı bir çalışmaya rastlanmamıştır. Dolayısıyla literatürde, bu raporları metin madenciliği teknikleriyle analiz edecek mühendislik tabanlı araştırmalara ihtiyaç duyulmaktadır.

Bu çalışmanın amacı, Türkiye’deki üniversitelerin sürdürülebilirlik raporlarını metin madenciliği teknikleri kullanarak sistematik bir şekilde analiz etmek; bu raporlarda öne çıkan tematik eğilimleri ve odak alanlarını ortaya koymaktır. Çalışma, 2024 yılında yayımlanan Türkçe sürdürülebilirlik raporlarını kapsamakta ve UI GreenMetric sıralama verileri ile tanımlayıcı bir bağlam oluşturmaktadır. Çalışmanın özgün katkısı üç boyuttadır:

1. Türkiye bağlamında yükseköğretim sürdürülebilirlik söylemlerine dair kapsamlı metin madenciliği analizi sunması,
2. Hibrit bir nicel-nitel metodolojik yaklaşım (MAXQDA ve otomatik ön işleme) önermesi,
3. Bulguları uluslararası performans göstergeleri (GreenMetric) ile ilişkilendirerek söylem ile performans arasındaki genel bağlantıya dair ön bulgular sunmasıdır.

Bu kapsamda çalışmanın temel hedefi, Türkiye’deki üniversiteler tarafından yayımlanan Türkçe sürdürülebilirlik raporlarının içerik özelliklerini sistematik biçimde incelemek ve bu raporlarda ortaya çıkan söylemsel örüntüleri uluslararası sürdürülebilirlik performans göstergeleri ile ilişkilendirmektir. Bu doğrultuda çalışma şu araştırma sorularına yanıt aramaktadır:

- Türkçe sürdürülebilirlik raporlarında en sık vurgulanan temalar ve kavramsal odaklar nelerdir?
- Rapordaki tematik yapı, Çevresel–Sosyal–Yönetişim (ESG) kategorileri açısından nasıl bir dağılım sergilemektedir?
- Raporlama yapan üniversitelerin kurumsal türü (Devlet/Vakıf) ile rapor içerik yoğunluğu arasında farklılık var mıdır?

Makale, giriş bölümünden sonra ilgili literatürün gözden geçirildiği bölümle devam etmektedir. Ardından, veri setinin oluşturulması, metin ön işleme adımları ve analitik çerçevenin detaylandırıldığı Materyal ve Yöntem bölümü yer almaktadır. Bulgular bölümünde tanımlayıcı istatistikler, terim frekansları, tematik dağılımlar ve ilişki analizleri sunulmaktadır. Tartışma bölümünde, bulguların metodolojik, teorik ve pratik çıkarımları değerlendirilmekte, literatürle karşılaştırılmakta ve sınırlılıklar belirtilmektedir. Makale, gelecek araştırmalar için öneriler ve sonuç bölümleri ile tamamlanmaktadır.

2. İLGİLİ ÇALIŞMALAR

Bu bölümde, ilgili literatür taranarak; yükseköğretimde sürdürülebilirlik raporlamasının önemi, kurumsal rapor analizlerinde kullanılan metin madenciliği uygulamaları, Türkçe metin işlemenin getirdiği morfolojik zorluklar ve nitel analiz yazılımlarının (MAXQDA, NVivo) sağladığı metodolojik katkılar incelenmiştir. Bölümün sonunda, mevcut araştırmalardaki boşluklar değerlendirilerek bu çalışmanın literatüre sağlayacağı özgün katkı ortaya konulmuştur.

2.1. Yükseköğretimde Sürdürülebilirlik Raporlaması

Yükseköğretim kurumlarında sürdürülebilirlik raporlaması, özel sektöre kıyasla henüz gelişim aşamasındadır ve benimsenme düzeyi ile kapsam açısından kurumlar arasında önemli farklılıklar bulunmaktadır (Mo & Wang, 2023). Mevcut raporlar incelendiğinde üniversitelerin genellikle çevresel ve ekonomik boyutlara odaklandığı, buna karşın sosyal boyutları ve

toplumsal katkı programlarını ihmal etme eğiliminde olduğu görülmektedir. Avrupa üniversiteleri bu alanda en aktif kurumlar arasında yer alsa da, küresel ölçekte raporlama kültürü sınırlı kalmakta; süreç genellikle gönüllülük esasına dayalı ve parçalı bir yapı sergilemektedir (Filho vd., 2022; Sepasi vd., 2019). Küresel Raporlama Girişimi (GRI) ve STARS gibi standartlaştırılmış çerçevelerin kullanımı artsa da, bu standartların yükseköğretim dinamiklerine uyum süreci hâlen devam etmektedir (Caeiro vd., 2020).

Sürdürülebilirlik raporları, yükseköğretim kurumlarının sürdürülebilirlik ilkelerini içselleştirmelerinde, paydaş iletişimini güçlendirmelerinde ve bu ilkeleri stratejik planlarına entegre etmelerinde kritik bir rol oynamaktadır (Lakhno, 2024). Ancak, stratejik planlar ile yayımlanan raporlar arasındaki uyumsuzluk, yüzeysel içerikler veya "yeşil aklama" (greenwashing) riski ve dış denetim (güvence) eksikliği gibi zorluklar sürecin güvenilirliğini zedelemektedir (Di Tullio & Rea, 2025). Bu noktada raporların kalitesi belirleyici olmaktadır; şeffaf, kapsamlı ve düzenli güncellenen raporlar, kurumlara rekabet avantajı sağlamakta ve paydaş güvenini artırmaktadır (Yanez vd., 2019).

Raporlamada şeffaflığı ve denetim kalitesini artırmak amacıyla GRI, STARS ve dijital raporlama standardı olan XBRL gibi araçlar yaygınlaşmaktadır; ancak bu sistemlerin etkin kullanımı, teknik uzmanlık ve kaynak gerektirmektedir (Basheer vd., 2024). Yükseköğretim kurumlarının çok boyutlu etkilerini ölçmek için çeşitli kıyaslama araçları geliştirilse de, sürdürülebilirliğin dışsal ve geleneksel olmayan yönlerini ölçümlemede eksiklikler devam etmektedir (Mohammed vd., 2025; Filho vd., 2022).

Sonuç olarak, yükseköğretimde sürdürülebilirlik raporlaması yaygınlaşma eğiliminde olsa da, içerik derinliği ve tutarlılık açısından henüz istenilen seviyeye ulaşmamıştır. Kurumsal stratejilerle uyumlu, şeffaflığı ve paydaş katılımını önceleyen, dijital araçlarla desteklenmiş özelleştirilmiş raporlama çerçevelerinin geliştirilmesi, bu alanda anlamlı bir ilerleme kaydedilmesi için hayati önem taşımaktadır.

2.2. Kurumsal Rapor Analizlerinde Metin Madenciliği ve NLP

Metin madenciliği ve Doğal Dil İşleme (NLP) teknolojileri, büyük hacimli ve yapılandırılmamış metinsel verilerden anlamlı iç görüler elde edilmesini sağlayarak kuruluşların kurumsal rapor analiz süreçlerini kökten değiştirmektedir. Bu teknolojiler; finansal tabloların, sürdürülebilirlik bildirimlerinin ve risk raporlarının otomatik analizinde, karar alma mekanizmalarının desteklenmesinde, uyumluluk denetimlerinde ve şeffaflığın artırılmasında giderek daha yaygın bir kullanım alanı bulmaktadır.

Analiz yöntemlerinin evrimine bakıldığında; ilk aşamada temaları ve duygu durumlarını belirlemek için anahtar kelime aramaları, kelime sayımları ve sözlük tabanlı yaklaşımlar kullanılmış, ancak bu yöntemler metnin bağlamını yakalamada sınırlı kalmıştır (Lewis & Young, 2019). Günümüzde ise Naive Bayes, Destek Vektör Makineleri (SVM) ve Sinir Ağları gibi makine öğrenmesi algoritmaları ile Latent Dirichlet Allocation (LDA) gibi konu modelleme tekniklerinin yerini; duygu analizi, konu çıkarımı ve özetlemede çok daha yüksek doğruluk sunan BERT ve GPT gibi derin öğrenme tabanlı dönüştürücü (transformer) modelleri almıştır (Kraus & Feuerriegel, 2017).

Modellerin karmaşıklığı arttıkça, kural tabanlı ve istatistiksel yöntemlerin Açıklanabilir NLP (XNLP) yaklaşımlarıyla entegre edilmesi, sonuçların yorumlanabilirliğini ve otomatik analizlere duyulan güveni artırmaktadır (Ong vd., 2024). Ancak, raporlama formatlarındaki tutarsızlıklar ve verilerin yapılandırılmamış doğası, model performansını ve raporlar arası

karşılaştırılabilirliği olumsuz etkileyebilmektedir (Lewis & Young, 2019). Ayrıca, özellikle yüksek riskli finansal kararlarda şeffaflığın sağlanması ve algoritmik önyargıların (bias) minimize edilmesi hayati önem taşımaktadır (Trachova & Lysak, 2025). Bu teknolojilerin kurumsal dünyada benimsenme hızı, gelişen yasal düzenlemeler ve açıklanabilir, denetlenebilir yapay zeka sistemlerine duyulan ihtiyaç ekseninde şekillenmeye devam etmektedir (Lewis & Young, 2019).

Sonuç olarak metin madenciliği ve NLP, iç gözü çıkarımını otomatikleştirerek kurumsal rapor analizinde devrim yaratırken; veri kalitesi, yorumlanabilirlik ve etik kullanım konusundaki zorluklar, araştırmaların standardizasyon ve finansal karar alma süreçleriyle entegrasyona odaklanmasını zorunlu kılmaktadır.

2.3. Türkçe Metin İşleme ve Dilin Morfolojik Zorlukları

Türkçe, kelimelerin köklerine getirilen çok sayıda yapım ve çekim ekiyle türetildiği "sondan eklemeli" (agglutinative) yapısı nedeniyle, Doğal Dil İşleme (NLP) çalışmalarında İngilizce gibi dillerden ayrılan özgün yapısal zorluklar barındırmaktadır. Bu morfolojik zenginlik; tokenizasyon (belirteçlere ayırma), morfolojik analiz, metin normalizasyonu ve duygu analizi gibi temel NLP görevlerinin karmaşıklığını artırmaktadır.

Teorik olarak sonsuz sayıda kelime biçiminin üretilebilmesi, özellikle kök bulma (lemmatization), gövdeleme (stemming) ve morfolojik belirsizlik giderme (disambiguation) süreçlerini zorlaştırmaktadır (Ofłazer, 2014). Buna ek olarak, ünlü ve ünsüz uyumu gibi fonolojik kurallar, algoritmik analiz süreçlerine ek bir katman daha eklemektedir (Tohma & Kutlu, 2020). Kelime dağarcığının bu denli genişlemesi, veri setlerinde "veri seyrekliği" (data sparsity) ve "yüksek boyutluluk" sorunlarına yol açarak, standart NLP modellerinin genelleme yeteneğini sınırlamaktadır (Ofłazer, 2014). Ayrıca Türkçe'de morfoloji, sözdizimi (syntax) ile sıkı bir etkileşim hâlinindedir; bu durum cümle yapısının çözülmesini ve ayrıştırılmasını (parsing) doğrudan etkilemektedir (Sak, 2011).

Bu zorlukları aşmak için literatürde sonlu durum dönüştürücüleri (finite-state transducers) ve ortalama algılayıcı (averaged perceptron) tabanlı belirsizlik giderme araçları geliştirilmiştir (Demir & Topcu, 2022). Özellikle sosyal medya verisi gibi gürültülü içeriklerin işlenmesinde, grafik tabanlı ve sinirsel ağ (neural) tabanlı normalizasyon yaklaşımları Türkçe NLP araçlarının başarısını önemli ölçüde artırmıştır (Demir & Topcu, 2022).

Güncel çalışmalarda ise kelime dağarcığı boyutu ile model performansı arasındaki dengeyi optimize etmek için Byte Pair Encoding (BPE) ve WordPiece gibi "alt kelime" (sub-word) modellerinin, karakter düzeyindeki yöntemlerden daha etkili olduğu görülmüştür (Toraman vd., 2022). Hibrit karakter-kelime modelleri ve BERT gibi dönüştürücü (transformer) tabanlı mimariler, geleneksel yöntemlere kıyasla daha yüksek performans göstererek yoğun ön işleme ihtiyacını azaltmaktadır (Toraman vd., 2022).

Sonuç olarak; Türkçenin zengin morfolojisi, veri seyrekliği ve sözdizimsel karmaşıklığı NLP çalışmaları için önemli bir engel teşkil etse de, hibrit ve derin öğrenme tabanlı modern mimariler bu alandaki performansı artırmaktadır. Yine de hem modern hem de tarihsel metinlerin işlenmesinde mevcut boşlukların giderilmesi için özelleştirilmiş çözümlere ihtiyaç duyulmaktadır.

2.4. Nitel Analiz Yazılımları (MAXQDA, NVivo) ve Metodolojik Katkıları

Modern nitel araştırmaların merkezinde yer alan MAXQDA ve NVivo gibi Bilgisayar Destekli Nitel Veri Analizi (CAQDAS) yazılımları; veri yönetimi, kodlama, görselleştirme ve iş birliği araçları sunarak araştırma süreçlerini dönüştürmektedir. Bu platformlar, karmaşık nitel veri setlerinin sistematik yönetimini kolaylaştırmakta; kodlama, not alma (memoing) ve görselleştirme süreçlerini dijitalleştirerek karma yöntem (mixed-methods) araştırmalarını güçlü bir şekilde desteklemektedir (Kalpokas & Radivojevic, 2021). Araştırma sürecinde dijital denetim izleri (audit trails) oluşturarak ve analitik kararların detaylı dokümantasyonunu sağlayarak, çalışmanın şeffaflığını ve metodolojik titizliğini (rigor) artırmaktadırlar (Woods vd., 2016).

Yazılımların metodolojik odakları incelendiğinde; MAXQDA'nın genellikle yorumlayıcı yaklaşımlar ve tematik içerik analizi için tercih edildiği, NVivo'nun ise Gömülü Teori (Grounded Theory) ve netnografi çalışmalarında daha yaygın kullanıldığı görülmektedir (Saillard, 2011). Disiplinler arası uyarlabilirlikleri sayesinde bu araçlar sağlık, eğitim ve iletişim çalışmalarında giderek artan bir kullanım oranına sahiptir (Woods vd., 2016).

Son yıllarda Yapay Zekâ (AI) algoritmalarının bu platformlara entegrasyonu analiz hızını artırsa da, araştırmacının süreçteki etkinliği ve epistemolojik uyum konusunda yeni tartışmaları beraberinde getirmektedir (Paulus & Marone, 2024). Bu bağlamda, özellikle kültürel açıdan hassas veya karmaşık araştırma bağlamlarında metodolojik şeffaflık ve düşünömsellik (reflexivity) önemini korumaktadır (Star vd., 2025). Literatür, yazılımın sağladığı hesaplama gücü ile geleneksel manuel kodlamanın yaratıcılığını birleştiren hibrit yaklaşımların hem titizliği hem de analitik derinliği optimize ettiğini göstermektedir (Maher vd., 2018).

Sonuç olarak MAXQDA ve NVivo, veri yönetimini ve şeffaflığı geliştirerek nitel araştırmayı dönüştürmüş olsa da; etkinlikleri ancak araştırma tasarımıyla bilinçli bir entegrasyon sağlandığında ve sınırlılıkları eleştirel bir gözle değerlendirildiğinde maksimize edilebilmektedir.

Ulusal ve uluslararası literatür tarandığında; kurumsal raporların metin madenciliği ile incelendiği çalışmaların arttığı, ancak bu çalışmaların büyük çoğunluğunun İngilizce metinler üzerine odaklandığı görülmektedir. Türkiye bağlamında ise aşağıdaki eksiklikler tespit edilmiştir:

1. Üniversite sürdürülebilirlik raporlarına odaklanan çalışmaların çoğu manuel içerik analizi ile sınırlıdır; yapılandırılmamış veri analizi teknikleri kullanılması daha iyi sonuçlar elde edecektir.
2. Türkçe raporların morfolojik özelliklerini dikkate alan NLP tabanlı sistematik bir analiz eksikliği bulunmaktadır.
3. En önemlisi, rapor içeriklerinin metin madenciliği ile elde edilen nicel metriklerinin, UI GreenMetric gibi somut performans çıktıları ile karşılaştırmalı olarak değerlendirildiği hibrit bir çalışma olarak katkı sağlaması hedeflenmektedir.

Bu çalışma, Türkçe metin madenciliği tekniklerini bir performans tahmin aracı olarak kullanarak literatürdeki bu metodolojik ve bağlamsal boşluğu doldurmayı hedeflemektedir.

3. MATERYAL VE YÖNTEM

3.1. Veri Seti ve Örneklem

Araştırmanın evrenini, Yükseköğretim Kurulu (YÖK) 2025 yılı verilerine göre Türkiye'de faaliyet gösteren 129'u devlet, 79'u vakıf (4'ü vakıf MYO dahil) olmak üzere toplam 208 yükseköğretim kurumu oluşturmaktadır [2]. Veri toplama sürecinde, üniversitelerin resmi web siteleri taranarak 2022-2024 yılları arasında yayımlanan sürdürülebilirlik raporları incelenmiştir. Yapılan tarama sonucunda; evreni oluşturan 208 üniversitenin 65 tanesinin (%31,2) belirtilen yıllar aralığında en az bir kez sürdürülebilirlik raporu yayımladığı tespit edilmiştir. Rapor yayımlayan bu üniversitelerin 47'si devlet üniversitesi, 18'i ise vakıf üniversitesidir. Raporlama eğiliminin yıllara göre değişimi incelendiğinde; 2022 yılında 21, 2023 yılında 37 ve 2024 yılında 37 olmak üzere son iki yılda raporlama sayısında bir istikrar sağlandığı görülmüştür.

Araştırmanın ana odak noktası olan 2024 yılı verileri detaylandırıldığında, yayımlanan 36 raporun 26'sinin Türkçe, 10'unun ise İngilizce olduğu belirlenmiştir. Metin madenciliği analizlerinde dil birliğini sağlamak amacıyla 26 adet Türkçe rapor araştırmanın nihai veri seti (corpus) olarak kabul edilmiştir. Bu analiz setinde yer alan üniversitelerin 19'u devlet, 7'i ise vakıf üniversitesidir.

Ayrıca, raporlama davranışı ile kurumsal sürdürülebilirlik performansı arasındaki ilişkiyi ölçmek amacıyla UI GreenMetric 2025 verileri kullanılmıştır. Raporu bulunan 65 üniversite ile GreenMetric sıralamasında yer alan 142 Türk üniversitesi karşılaştırıldığında, 61 üniversitenin her iki veri setinde de yer aldığı (kesişim kümesi) tespit edilmiştir [1].

GreenMetric sıralama verileri, 27 üniversiteden 24'ü için mevcuttur (3 üniversite GreenMetric'a kayıtlı değildir). Türkiye sıralaması ortalaması 60,54'tür (min=4, max=117), bu da örneklemdeki üniversitelerin Türkiye'deki 117 üniversite arasında orta sıralarda yer aldığını göstermektedir. Dünya sıralaması ortalaması ise 611,83'tür (min=85, max=1183), bu da küresel ölçekte orta-alt segmentte bir konuma işaret etmektedir. İki sıralama arasındaki fark, Türk üniversitelerinin ulusal ve uluslararası sürdürülebilirlik performansı arasında önemli bir boşluk olduğunu düşündürmektedir.

Tablo 1. Veri Setinin Tanımlayıcı İstatistikleri.

Parametre	Değer	Açıklama
Toplam İncelenen Üniversite Sayısı	208	YÖK 2025 Güncel listesi [2].
Sürdürülebilirlik Raporu Olan Üniversite Sayısı	65	208 Üniversite içerisinde 2022-2024 yılları arasında rapor yayımlayan üniversiteler.
2024 Rapor Sayısı	36	2024 yılında üniversiteler tarafından yayımlanan toplam sürdürülebilirlik rapor sayısı.
2024 Raporlarının Ortalama Sayfa Sayısı	73,185	2024 yılına ait raporların ortalama uzunluğu.
2024 Türkçe Rapor Sayısı	26	2024 yılı içinde

		yayımlanan Türkçe raporlar.
2024 İngilizce Rapor Sayısı	10	2024 yılı İngilizce raporlar.
2024 Türkçe Raporu Olan Devlet Üniversitesi	19	Türkçe rapor yayımlayan devlet üniversiteleri.
2024 Türkçe Raporu Olan Vakıf Üniversitesi	7	Türkçe rapor yayımlayan vakıf üniversiteleri.
2023 Rapor Sayısı	37	Yıllık rapor sayısı.
2022 Rapor Sayısı	21	Yıllık rapor sayısı.
Devlet Üniversitesi Sayısı	129	YÖK listesi [2].
Raporu Olan Devlet Üniversitesi Sayısı	47	2022-2024 rapor yayımlayan devlet üniversiteleri.
Vakıf Üniversitesi Sayısı	79	75 Vakıf Üniversitesi, 4 tane vakıf meslek yüksekokulu [2].
Raporu Olan Vakıf Üniversite Sayısı	18	2022-2024 rapor yayımlayan vakıf kurumları.
GreenMetric’te Yer Alan Üniversite Sayısı	142	GM 2025 veri seti [1].
Hem Raporu Hem GreenMetric’te Yer Alan Üniversite Sayısı	61	Kesişim Kümesi.
GreenMetric Türkiye Sıralama Ortalaması	60.54	GM Veri Seti [1].
GreenMetric Dünya Sıralama Ortalaması	611.83	GM Veri Seti [1].
2024 Rapor Ortalama Sayfa Sayısı	68.5	2024 Türkçe yayımlanan sürdürülebilirlik raporları sayfa sayısı ortalaması.

3.2. Metin Çıkarımı ve Ön İşleme Pipeline

Bu çalışmada analiz edilen sürdürülebilirlik raporları PDF formatında olup, yapılandırılmamış metin niteliği taşımaktadır. Bu nedenle raporların incelenebilmesi için MAXQDA içerisinde üç aşamalı bir metin çıkarımı ve ön işleme süreci uygulanmıştır.

Birinci aşamada, raporlar MAXQDA’nın “PDF Import” modülü kullanılarak doğrudan çalışma alanına aktarılmıştır. MAXQDA, PDF içerisindeki metni otomatik olarak ayrıştırarak düzenli bir metin yapısına dönüştürmekte; tablo, şekil, başlık formatları ve sayfa numarası gibi analiz sürecine gürültü katabilecek unsurları büyük ölçüde hariç bırakmaktadır. Bu sayede raporların tamamı, kodlamaya uygun bütünlüklü metin belgeleri hâline gelmiştir.

İkinci aşamada, metin temizleme ve normalizasyon işlemleri gerçekleştirilmiştir. MAXQDA’nın “Document Browser – Find & Replace” aracı kullanılarak satır sonu kırılmaları, fazla boşluklar, tekrarlayan karakterler ve bozuk Türkçe karakter çözümleri düzeltilmiştir. Ayrıca her raporun ilk sayfalarında yer alan zorunlu yasal ifadeler, kapak metinleri ve gereksiz formatlanmış içerikler analiz dışı tutulmuştur. Bu adım, hem tematik kodlamanın doğruluğunu artırmış hem de dilsel gürültüyü en aza indirmiştir.

Üçüncü aşamada, raporlar içerik analizi için kodlanabilir hâle getirilmiş ve tematik yapı MAXQDA'nın Kod Sistemi aracılığıyla oluşturulmuştur. ESG çerçevesi temel alınarak önceden belirlenmiş çevresel, sosyal ve yönetim kodları dedüktif olarak eklenmiş; bunun yanında raporların kendi iç yapısından türeyen yeni kavramlar indüktif kodlama ile sisteme dahil edilmiştir. Her rapor ayrı bir “case” olarak tanımlanmış ve kod yoğunlukları Code Matrix Browser, Code Relations Browser ve Code Frequency araçlarıyla analiz edilmiştir. Bu süreç, Python tabanlı otomatik NLP yöntemlerinde karşılığı bulunan kelime frekansı, ortak görülme (co-occurrence) ve tematik kümeleme analizlerinin MAXQDA ortamındaki eşdeğerini oluşturmaktadır.

Bu üç aşamalı pipeline sayesinde tüm raporlar dilsel gürültüden arındırılmış, içerik açısından standardize edilmiş ve sistematik tematik analiz için hazır hâle getirilmiştir.

3.3. Analitik Çerçeve

Veri setindeki yapılandırılmamış metinlerin analizi için hibrit bir yaklaşım benimsenmiştir. Bu yaklaşım, istatistiksel metin madenciliği teknikleri (TF-IDF, Frekans Analizi) ile nitel içerik analizi yöntemlerini (Tematik Kodlama) birleştiren çok katmanlı bir analitik çerçeveye dayanmaktadır. Analiz sürecinde kullanılan temel matematiksel modeller ve kodlama stratejileri aşağıda detaylandırılmıştır.

3.4. Vektör Uzay Modeli ve Terim Frekans Matrisi (TF)

Raporların matematiksel olarak işlenebilmesi için öncelikle metinler “Kelime Çantası” (Bag of Words) modeli ile temsil edilmiştir. Her bir d dökümanı (üniversite raporu), V boyutundaki bir kelime dağarcığı üzerinde tanımlanan bir vektör olarak kabul edilmiştir. t teriminin d dökümanındaki ham frekansı $tf(t, d)$ hesaplanarak, $D \times V$ boyutunda bir Terim-Frekans (TF) matrisi oluşturulmuştur. Bu matris, üniversitelerin sürdürülebilirlik söylemlerindeki yoğunluğu sayısal olarak karşılaştırmaya olanak tanımıştır.

3.4.1. TF-IDF Ağırlıklandırma Yöntemi

Raporlarda sık geçen ancak ayırt edici özelliği düşük olan genel kelimelerin (örneğin; "üniversite", "yıl", "rapor" vb.) etkisini azaltmak ve sürdürülebilirlik bağlamında özgün terimleri öne çıkarmak amacıyla TF-IDF (Term Frequency-Inverse Document Frequency) ağırlıklandırma algoritması uygulanmıştır. Bir t teriminin d dökümanındaki ağırlığı ($W_{t,d}$) Denklem 1’deki formül ile hesaplanmıştır:

$$W_{t,d} = tf(t, d) \times \log\left(\frac{N}{df(t)}\right) \quad \text{Denklem (1)}$$

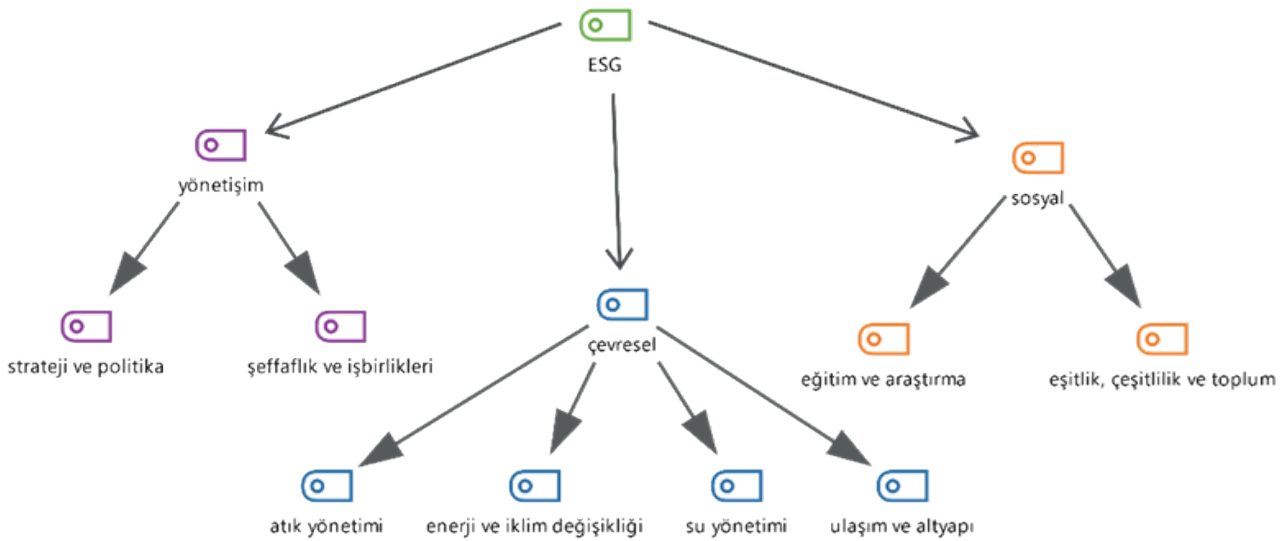
Burada N toplam rapor sayısını, $df(t)$ ise t teriminin geçtiği rapor sayısını ifade etmektedir. Bu işlem, MAXQDA’nın sözcük matrisi ve Python scikit-learn kütüphanesi referans alarak doğrulanmıştır.

3.4.2. ESG Tabanlı Hibrit Kodlama Stratejisi

Metin madenciliği bulgularını anlamlandırmak için Çevresel, Sosyal ve Yönetişim (ESG) kriterleri temel alınarak hiyerarşik bir kod sistemi oluşturulmuştur. Kodlama süreci iki aşamada gerçekleştirilmiştir:

1. Dedüktif (Tümdengelim) Kodlama: Literatürdeki standart ESG metrikleri ve UI GreenMetric kategorileri sisteme ön tanımlı kodlar olarak eklenmiştir (Örn: Enerji Verimliliği, Atık Yönetimi, Toplumsal Eşitlik).
2. İndüktif (Tümevarım) Kodlama: Metin taraması sırasında raporlara özgü yeni kavramlar (Örn: "Sıfır Atık Belgesi", "Kampüs Bostanı") tespit edilerek kod şemasına dinamik olarak dahil edilmiştir.

Literatürdeki standart ESG metrikleri ve UI GreenMetric kategorileri sisteme ön tanımlı kodlar olarak eklenmiş; rapor taramaları sırasında tespit edilen yeni kavramlar ise indüktif olarak sisteme dahil edilmiştir. Araştırma kapsamında oluşturulan ve analizde temel alınan hiyerarşik kodlama modeli Şekil 2'de sunulmuştur.



Şekil 2. MAXQDA Ortamında Oluşturulan ESG Tabanlı Hiyerarşik Kodlama Şeması ve Alt Kod Kırılımları.

3.4.3. İlişki Ağları ve Eş-Oluşum (Co-occurrence) Analizi

Sürdürülebilirlik kavramlarının birbirleriyle olan anlamsal bağlarını ortaya koymak amacıyla kod eş-oluşum analizi yapılmıştır. İki farklı kodun (C_i ve C_j) aynı segment içinde (paragraf veya cümle) birlikte görülme sıklığı hesaplanarak $C \times C$ boyutunda bir simetri matrisi elde edilmiştir. Bu matris, üniversitelerin stratejik planlarında hangi konuları (örneğin; "Eğitim" ve "İklim Eylemi") bütünleşik olarak ele aldığını haritalamak için kullanılmıştır.

3.5. Görselleştirme

Elde edilen yüksek boyutlu verilerin ve karmaşık ilişkilerin yorumlanabilirliğini artırmak amacıyla çoklu veri görselleştirme teknikleri kullanılmıştır. Üniversitelerin odaklandığı baskın temaları görselleştirmek için terim frekanslarına dayalı kelime bulutları oluşturulmuştur.

Kod eş-oluşum matrisleri, değişkenler arasındaki korelasyon gücünü renk yoğunluğu ile ifade eden ısı haritalarına dönüştürülmüştür. Bu yöntem, özellikle Çevresel ve Sosyal boyutlar arasındaki kesişim kümelerini tespit etmekte kullanılmıştır.

Üniversite türü (Devlet/Vakıf) ile tematik kodların dağılımı arasındaki ilişki, MAXQDA Code Matrix Browser kullanılarak görselleştirilmiş; böylece kurumsal türün rapor içeriği üzerindeki etkisi grafiksel olarak sunulmuştur.

4. BULGULAR

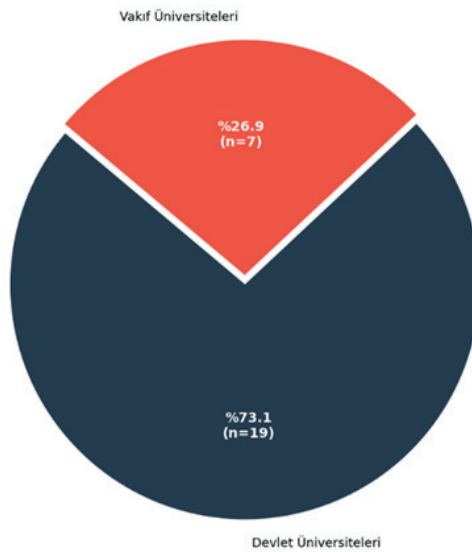
4.1. Tanımlayıcı İstatistikler

Metin madenciliği analizi öncesinde, veri setinin yapısal özellikleri incelenmiştir. Analize dahil edilen 2024 yılı Türkçe sürdürülebilirlik raporlarının ortalama uzunluğunun 68,5 sayfa olduğu hesaplanmıştır. Bu sayfa hacmi, raporların sadece özet bilgi içermediğini, derinlemesine veri ve metin yoğunluğuna sahip olduğunu göstermektedir. Raporlarda ortalama kelime sayısı ise 8902,5 olarak hesaplanmıştır. Raporlardaki benzersiz kelime sayıları ortalaması 3087,76 ve 90841,34 karakter sayısı ortalaması olarak belirlenmiştir.

Örneklem grubunun uluslararası sürdürülebilirlik performansı incelendiğinde; Türkçe rapor yayımlayan üniversitelerin UI GreenMetric 2025 Dünya Sıralaması ortalamasının 611,83, Türkiye Sıralaması ortalamasının ise 60,54 olduğu belirlenmiştir [1]. Bu veriler, araştırmaya konu olan üniversitelerin Türkiye genelinde sürdürülebilirlik farkındalığı yüksek ve performansı ölçülebilir bir grubu temsil ettiğini kanıtlamaktadır.

Analiz kapsamındaki raporların kurumsal dağılımı incelendiğinde, kamu üniversitelerinin raporlama sürecinde daha baskın bir rol oynadığı görülmektedir. 2024 yılı veri setinde yer alan üniversitelerin 19’u devlet üniversitesi, 7’si ise vakıf üniversitesi statüsündedir. Devlet üniversitelerinin bu sayısal üstünlüğü (%73,1), kamu kaynaklarını kullanan kurumların hesap verebilirlik ilkelerine uyum sağlama ve YÖK’ün stratejik hedefleriyle uyumlanma konusundaki hassasiyetlerini yansıtmaktadır. Vakıf üniversiteleri (%26,9) ise sürdürülebilirlik raporlamasını genellikle uluslararası öğrenci temini ve küresel sıralamalarda (UI GreenMetric, THE Impact Rankings) rekabet avantajı elde etmek amacıyla stratejik bir iletişim aracı olarak değerlendirmektedir.

Kurumsal yapıdaki bu dağılım, raporlama motivasyonlarının farklılaştığını, ancak her iki tür kurumun da küresel sürdürülebilirlik hedeflerine yöneliminin arttığını göstermektedir. Raporlama yapan üniversitelerin kurumsal türlerine göre dağılımı Şekil 3’de görselleştirilmiştir.



Şekil 3. 2024 Yılında Türkçe Sürdürülebilirlik Raporu Yayımlayan Üniversitelerin Kurum Türüne Göre Dağılımı.

4.2. Terim Frekans Analizi

Veri ön işleme ve durak kelime (stop-word) temizliği adımlarının ardından, üniversitelerin sürdürülebilirlik raporlarında en sık kullanılan kök kelimeler analiz edilmiştir. Bu analiz, kurumların sürdürülebilirlik olgusunu hangi kavramlarla ilişkilendirdiğini ortaya koymaktadır.



Şekil 4. Türk Üniversitelerinin Sürdürülebilirlik Raporlarında En Yüksek Frekansa Sahip Terimlerin Kelime Bulutu (Word Cloud) Gösterimi.

Şekil 4 incelendiğinde, raporların merkezinde beklendiği üzere 'sürdürülebilirlik' ve 'sürdürülebilir' kavramları yer almaktadır. Ancak kelime bulutundaki diğer baskın terimler (büyük puntolu ifadeler), üniversitelerin stratejik öncelikleri hakkında önemli ipuçları vermektedir:

1. Akademik Misyonun Önceliği: 'Eğitim', 'Araştırma', 'Akademik' ve 'Öğrenci' kelimelerinin yüksek frekansı, Türk üniversitelerinin sürdürülebilirliği sadece bir kampüs yönetimi işi olarak değil, aynı zamanda eğitim-öğretim müfredatının ve Ar-Ge faaliyetlerinin bir parçası olarak konumlandığını göstermektedir.
2. Operasyonel Odaklar: 'Enerji', 'Atık' ve 'Kampus' kelimelerinin belirginliği, üniversitelerin çevresel performansı ölçülebilir ve somut eylemler (enerji verimliliği, atık yönetimi) üzerinden raporlamaya eğilimli olduğunu kanıtlamaktadır.
3. Sosyal Boyut: 'Sosyal', 'Toplumsal' ve 'İnsan' kelimelerinin varlığı, raporların sadece çevre odaklı kalmadığını, ESG'nin sosyal boyutunun da metinlerde yer bulduğunu işaret etmektedir.

Tablo 2. 2024 Yılı Sürdürülebilirlik Raporlarında En Sık Kullanılan Terimler (TF Listesi).

Kelimeler	Frekans	Kelime	Frekans
Sürdürülebilirlik	1884	üretim	372
Sürdürülebilir	1823	ekonomik	369
Enerji	1597	dijital	348
Eğitim	1299	iklim	325
Sosyal	1114	Kadın	315
atık	875	Elektrik	303

sağlık	648	teknoloji	301
kalkınma	648	sanayi	300
çevre	616	cinsiyet	298
toplumsal	587	kurumsal	292
çevresel	559	yenilenebilir	287
uluslararası	540	kültür	284
proje	511	bilimsel	282
yaşam	459	kariyer	276
öğretim	429	ulusal	257
küresel	412	karbon	252
etkinlik	404	personel	240
doğa	377	beslenme	239
yeşil	374	finansal	238

Tablo 2'deki veriler incelendiğinde, üniversitelerin stratejik önceliklerine dair şu çarpıcı sonuçlar ortaya çıkmaktadır:

Listede "Sürdürülebilirlik" kavramlarından hemen sonra gelen en baskın terim "Enerji" (n=1.597) olmuştur. Bu durum, "İklim" (n=325) ve "Karbon" (n=252) terimlerinin oldukça önündedir. Bu bulgu, Türk üniversitelerinin sürdürülebilirliği öncelikle enerji verimliliği, elektrik tüketimi ve yenilenebilir enerji gibi ölçülebilir ve maliyet odaklı operasyonel parametreler üzerinden yönettiğini göstermektedir.

"Sosyal" (n=1.114) kelimesinin yüksek frekansına ek olarak, "Kadın" (n=315) ve "Cinsiyet" (n=298) kelimelerinin listede belirgin bir yer tutması, üniversitelerin Birleşmiş Milletler SKA-5 (Toplumsal Cinsiyet Eşitliği) hedefine yönelik duyarlılığının raporlara yansıdığını kanıtlamaktadır.

"Sağlık" (n=648) ve "Beslenme" (n=239) terimlerinin varlığı, kampüslerdeki gıda güvenliği ve sağlıklı yaşam olanaklarının sürdürülebilirlik kapsamında raporlandığını göstermektedir. Bu durum, özellikle tıp fakültesi olan üniversitelerin etkisi ve kampüs yemekhane hizmetlerinin raporlamadaki ağırlığı ile açıklanabilir.

"Dijital" (n=348) ve "Teknoloji" (n=301) kavramları, sürdürülebilirlik ile dijital dönüşümün (Twin Transition) üniversiteler nezdinde birlikte ele alındığına işaret etmektedir.

4.3. Tematik Kodlama Bulguları

Metin madenciliği analizinde kullanılan ESG sözlüğü (dictionary), alana özgü literatür taranarak oluşturulmuştur. Çevresel kategorisi için UI GreenMetric göstergeleri (Enerji, Atık, Su) temel alınmış; Sosyal kategorisi Birleşmiş Milletler Sürdürülebilir Kalkınma Amaçları (özellikle SKA 3, 4 ve 5) kapsamındaki anahtar kelimelerle (Sağlık, Eğitim, Cinsiyet Eşitliği) zenginleştirilmiştir. Yönetişim kategorisi ise GRI Standartları'nın 'Genel Açıklamalar' bölümünde yer alan kurumsal yönetim terminolojisine (Strateji, Etik, Şeffaflık) dayandırılmıştır.

Metin madenciliği analizinin ikinci aşamasında, toplam 19.446 adet sürdürülebilirlik odaklı terim; Çevresel (Environmental), Sosyal (Social) ve Yönetişim (Governance) boyutlarında

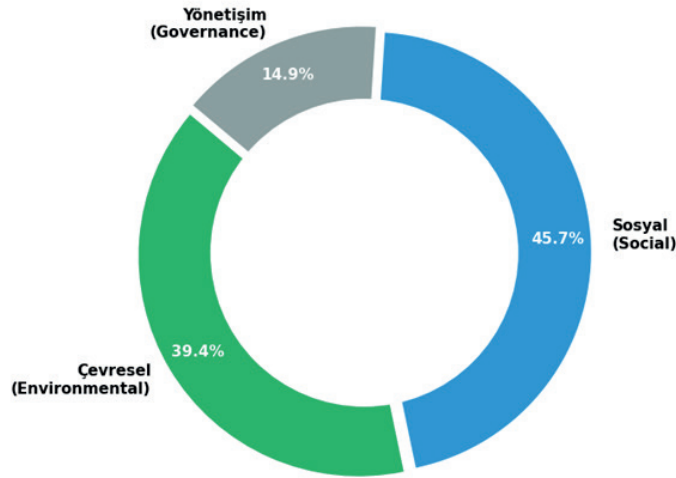
sınıflandırılmıştır. Analiz sonuçları, üniversitelerin sürdürülebilirlik raporlarında Sosyal > Çevresel > Yönetişim şeklinde bir hiyerarşik yapı olduğunu ortaya koymuştur.

Elde edilen tematik dağılım oranları Şekil 3'te görselleştirilmiş olup detaylı analiz aşağıda sunulmuştur:

Analiz sonucunda en yüksek yoğunluğa sahip kategori 8.889 terim ile Sosyal boyut olmuştur. Bu durum, üniversitelerin temel varlık nedeninin "Eğitim" ve "Öğretim" olmasıyla doğrudan ilişkilidir. Raporlarda sıkça geçen "Eğitim", "Öğrenci" ve "Sağlık" kavramları, sürdürülebilirliğin sadece ekolojik değil, aynı zamanda beşeri bir kalkınma modeli olarak ele alındığını göstermektedir. Ayrıca "Kadın" ve "Eşitlik" terimlerinin yoğunluğu, toplumsal cinsiyet eşitliği (SKA 5) hedeflerinin raporlara güçlü bir şekilde yansıdığını kanıtlamaktadır.

Çevresel boyut, 7.659 terim ile sosyal boyutu yakından takip etmektedir. Üniversitelerin kampüs operasyonlarına yönelik "Enerji", "Atık" ve "Yeşil Alan" projeleri bu kategoriyi domine etmektedir. Çevresel boyutun sosyal boyutun gerisinde kalması bir eksiklik değil; aksine üniversitelerin sürdürülebilirliği "yeşil kampüs" kavramının ötesine taşıyarak toplumsal fayda eksenine oturtması olarak yorumlanabilir.

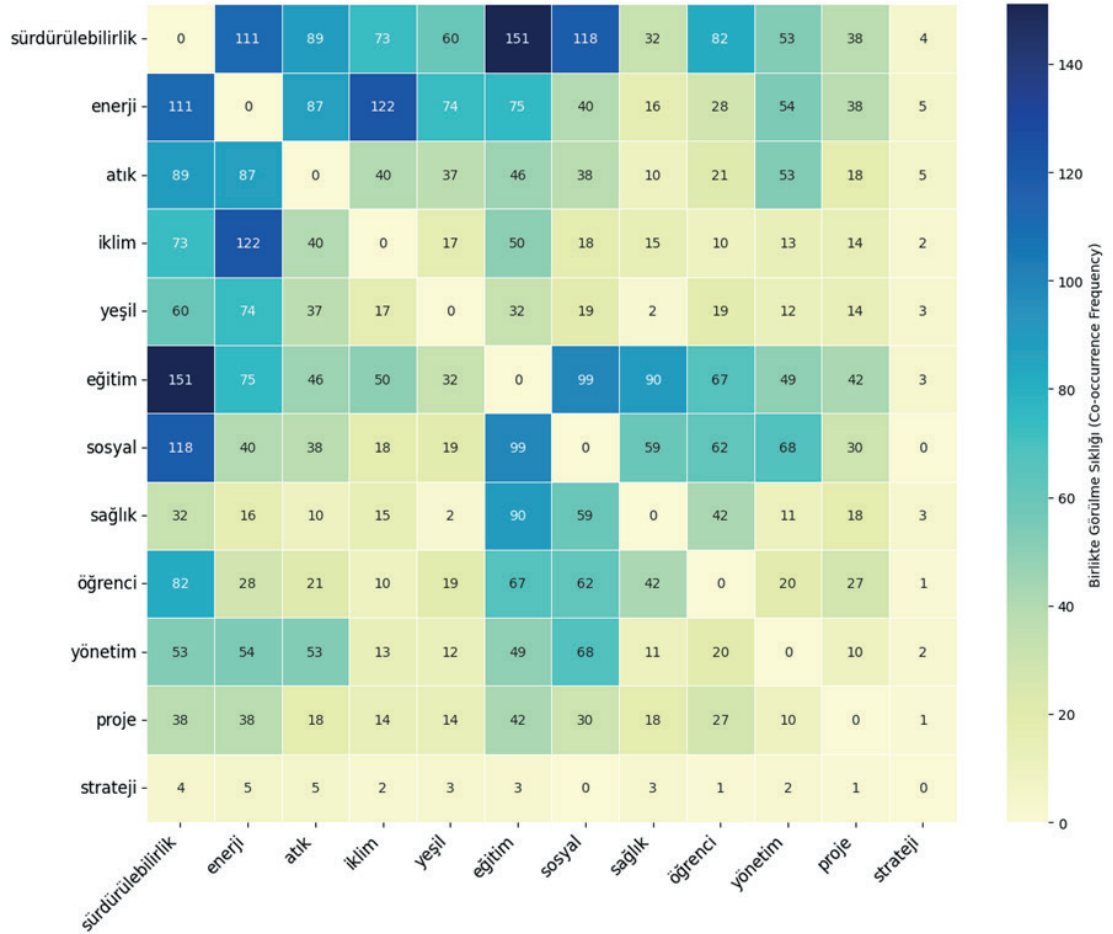
Yönetişim boyutu 2.898 terim ile oransal olarak en düşük paya sahiptir. Bu durum, sürdürülebilirlik süreçlerinin henüz stratejik yönetim, risk analizi ve finansal raporlama mekanizmalarına tam entegre edilmediğini; daha çok proje ve faaliyet bazlı yürütüldüğünü göstermektedir.



Şekil 5. Sürdürülebilirlik Raporlarında Tespit Edilen Terimlerin ESG Kategorilerine Göre Dağılımı.

4.4. Kod İlişki Analizi

Raporlardaki kavramların birbirinden bağımsız sıklıklarının ötesinde, birbirleriyle olan anlamsal bağlarını ve bağlamsal ilişkilerini ortaya koymak amacıyla Kod Eş-Oluşum (Co-occurrence) analizi yapılmıştır. Bu analizde, anahtar kavramların aynı cümle içerisinde birlikte geçme sıklıkları hesaplanarak Şekil 6'daki Isı Haritası (Heatmap) oluşturulmuştur.



Şekil 6. Sürdürülebilirlik Raporlarında Anahtar Kavramların Birlikte Görülme Sıklığını Gösteren Isı Haritası (Co-occurrence Heatmap).

Isı haritasındaki renk yoğunluğu ve kesişim değerleri incelendiğinde, üniversitelerin sürdürülebilirlik anlatısında şu stratejik bağlantılar öne çıkmaktadır:

Matristeki en güçlü ilişki, "Sürdürülebilirlik" ve "Eğitim" kavramları arasında tespit edilmiştir (151 eş-oluşum). Bu bulgu, Türk üniversitelerinin sürdürülebilirliği sadece kampüs operasyonlarına (atık, enerji vb.) indirgemediğini; aksine ders müfredatları, sertifika programları ve farkındalık eğitimleri yoluyla sürdürülebilirliği akademik misyonlarının merkezine yerleştirdiğini kanıtlamaktadır.

"Enerji" kavramı en çok "iklim" (122) ve "Sürdürülebilirlik" (111) kavramlarıyla birlikte anılmaktadır. Bu durum, üniversitelerin enerji yönetimi konusunu (verimlilik, yenilenebilir kaynaklar) sadece mali bir tasarruf kalemi olarak değil, küresel iklim değişikliği ile mücadele stratejisinin bir parçası olarak çerçevelediğini göstermektedir.

Önceki bölümlerdeki bulguları destekler nitelikte, "Sosyal" ve "Sürdürülebilirlik" kavramları arasında çok güçlü bir bağ (118) bulunmaktadır. Ayrıca "Eğitim - Sosyal" (99) ve "Eğitim - Sağlık" (90) ilişkilerinin yüksekliği, üniversitelerin toplumsal katkı hedeflerini eğitim ve sağlık hizmetleri üzerinden gerçekleştirdiğini doğrulamaktadır.

Matrisin en dikkat çekici negatif bulgusu ise "Strateji" satırındaki düşük değerlerdir (Genellikle <5). Üniversiteler sürdürülebilirlik hakkında çok sayıda "Proje" (42 kez eğitimle, 38 kez enerjiyle ilişkili) üretse de, bu faaliyetlerin metin içerisinde açıkça bir "Strateji" kavramıyla ilişkilendirilme oranı düşüktür. Bu durum, eylemlerin stratejik dokümantasyondan ziyade operasyonel düzeyde raporlandığına işaret etmektedir.

5. TARTIŞMA

Bu çalışma, Türkiye'deki üniversitelerin 2024 yılı sürdürülebilirlik raporlarını metin madenciliği teknikleriyle analiz ederek, yükseköğretimdeki sürdürülebilirlik söylemini nicel verilerle ortaya koymuştur. Literatürde yer alan ve genellikle manuel içerik analizi yöntemlerine dayanan önceki çalışmaların aksine, bu araştırma raporları yapılandırılmamış veri analizi perspektifiyle ele alarak metodolojik bir farklılık yaratmıştır. Analiz kapsamında işlenen 19.446 adet terim ve bu terimlerin anlamsal ilişkileri, üniversite raporlarının subjektif yorumlardan arındırılmış, matematiksel ve tekrarlanabilir bir profilini sunmaktadır.

Analiz bulgularının en dikkat çekici yönü, Türk üniversitelerinin sürdürülebilirlik anlayışında "Kampüs Operasyonları" ve "Eğitim Misyonu" arasında dengeli bir hibrit model geliştirmesidir. Literatürdeki genel eğilim ve kurumsal raporlama pratikleri genellikle Çevresel boyutun ezici üstünlüğüne işaret ederken, bu çalışmada Sosyal boyutun (%45,71) Çevresel boyutu (%39,39) geride bırakarak en baskın kategori olması özgün bir bulgudur. Bu durum, Türkiye'deki üniversitelerin kampüsleri sadece yönetilecek birer fiziki alan olarak değil, aynı zamanda barınma, beslenme ve sosyalleşme imkânı sunan birer yaşam merkezi olarak konumlandırılmasıyla açıklanabilir. Özellikle "Eğitim", "Sağlık" ve "Toplumsal Eşitlik" kavramlarının yüksek frekansı, üniversitelerin sürdürülebilirliği sadece ekolojik bir dönüşüm olarak değil, beşeri bir kalkınma hamlesi olarak benimsediğini göstermektedir.

Buna karşın Çevresel boyutun da ihmal edilmediği, aksine "Enerji" (n=1.597) ve "Atık" (n=875) terimlerinin raporların en sık kullanılan teknik kavramları olduğu görülmüştür. Bu bulgu, üniversitelerin artan enerji maliyetlerini yönetme kaygısı ve UI GreenMetric gibi uluslararası sıralamalarda rekabet etme motivasyonu ile paralellik göstermektedir. Kod eş-oluşum matrisinde "Enerji" ve "İklim" kavramları arasındaki güçlü ilişki, üniversitelerin operasyonel verimlilik çalışmalarını küresel iklim kriziyle mücadele stratejilerinin bir parçası olarak çerçevelediğini kanıtlamaktadır. Ancak, Yönetişim boyutunun (%14,90) diğer iki boyuta kıyasla oldukça zayıf kalması, sürdürülebilirliğin henüz üniversitelerin tepe yönetim stratejilerine ve finansal planlama süreçlerine tam olarak entegre edilemediğini düşündürmektedir. Raporlarda "Proje" kelimesinin sık geçmesine rağmen "Strateji", "Risk" ve "Performans" gibi yönetsel kavramların azlığı, sürdürülebilirlik çalışmalarının kurumsal bir politika olmaktan ziyade, birimler bazında yürütülen faaliyetler bütünü olarak kaldığına işaret etmektedir.

Uluslararası literatürle karşılaştırıldığında, Türk üniversitelerinin "Eğitim-Sürdürülebilirlik" entegrasyonu konusunda (Isı haritasındaki 151 kesişim değeri ile) küresel örneklerden daha bütünleşik bir yaklaşım sergilediği söylenebilir. Ancak raporlama standartları açısından, sayfa sayılarındaki yüksek varyasyon (Min: 15, Max: 150 sayfa) ve içerik formatındaki tutarsızlıklar, kurumsallaşma sürecinin henüz tamamlanmadığını göstermektedir. Bu bağlamda, Yükseköğretim Kurulu (YÖK) tarafından belirlenecek asgari bir "Sürdürülebilirlik Raporlama Şablonu", verilerin karşılaştırılabilirliğini ve şeffaflığını artıracaktır.

Sonuç olarak, bu çalışma Türkiye'deki üniversitelerin sürdürülebilirlik yolculuğunda önemli bir mesafe katettiğini, ancak sürecin "Proje" odaklılıktan "Strateji" odaklılığa evrilmesi gerektiğini ortaya koymaktadır. Çalışmanın sadece Türkçe raporlarla sınırlı olması ve PDF formatındaki görsel verilerin analiz dışı kalması bir sınırlılık olarak değerlendirilse de, elde edilen makine öğrenmesi tabanlı bulgular, politika yapıcılar ve üniversite yöneticileri için kanıta dayalı bir yol haritası sunmaktadır.

6. SONUÇ VE ÖNERİLER

Bu çalışma, Türkiye'deki üniversitelerin 2024 yılı sürdürülebilirlik raporlarını metin madenciliği ve doğal dil işleme (NLP) teknikleriyle analiz ederek, yükseköğretimdeki sürdürülebilirlik söyleminin kavramsal röntgenini çekmiştir. 26 üniversiteye ait raporların analizi sonucunda elde edilen bulgular, Türk üniversitelerinin sürdürülebilirlik yaklaşımının Sosyal (%45,71) > Çevresel (%39,39) > Yönetişim (%14,90) hiyerarşisinde şekillendiğini somut verilerle ortaya koymuştur. Literatürdeki genel beklentinin aksine, üniversitelerin sürdürülebilirliği öncelikli olarak eğitim, sağlık ve toplumsal eşitlik ekseninde tanımlaması; kampüslerin sadece fiziki birer alan değil, toplumsal birer yaşam merkezi olarak görüldüğünü kanıtlamaktadır. Çevresel boyutta ise "Enerji" (n=1.597) ve "Atık" (n=875) kavramlarının dominasyonu, "Yeşil Kampüs" hedeflerinin somut ve maliyet odaklı teknik projeler üzerinden yönetildiğini göstermektedir. Buna karşın, raporlarda "Proje" kelimesinin sıkça geçmesine rağmen "Strateji", "Risk" ve "Performans" gibi yönetim kavramlarının düşük frekansta kalması, sürdürülebilirliğin henüz kurumsal tepe yönetim mekanizmalarına tam entegre edilemediğini işaret etmektedir. Ayrıca kod eş-oluşum analizi, üniversitelerin "Eğitim" ve "Sürdürülebilirlik" kavramlarını güçlü bir şekilde ilişkilendirerek (n=151), ana misyonlarını sürdürülebilirlik hedefleriyle başarılı bir şekilde bütünleştirdiğini doğrulamaktadır.

Araştırmanın akademik ve kurumsal katkıları iki temel boyutta değerlendirilmektedir. Metodolojik açıdan çalışma, sürdürülebilirlik raporlarını subjektif manuel okumaların ötesine taşıyarak, Python tabanlı algoritmalarla analiz eden öncü araştırmalardan biridir. Nitel verilerin (metinlerin) nicel göstergelere dönüştürülmesi, raporlama kalitesinin ölçülmesinde objektif bir denetim aracı sunmaktadır. Kurumsal açıdan ise elde edilen ESG dağılım haritası, üniversite yönetimlerine mevcut durumlarını görme fırsatı tanımaktadır. Sosyal boyutun gücüne karşın yönetim boyutundaki zayıflığın veri ile kanıtlanması, gelecek stratejik planlarda bu alana odaklanılması gerektiğini göstermektedir.

Mevcut çalışmanın sınırları ve elde edilen bulgular ışığında, gelecek araştırmalar için çeşitli öneriler sunulmaktadır. Öncelikle, bu çalışmada uygulanan frekans bazlı analizin ötesine geçilerek, raporlardaki ifadelerin "duygu durumu" (sentiment analysis) incelenmeli ve üniversitelerin hedeflerine ulaşma konusundaki iyimserlik düzeyleri ölçülmelidir. Ayrıca, çalışmanın kapsamı genişletilerek Türk üniversitelerinin Türkçe ve İngilizce raporları arasındaki söylem farkları analiz edilebilir; böylece ulusal ve uluslararası paydaşlara yönelik iletişim stratejilerinin nasıl farklılaştığı ortaya konabilir. Son olarak, yapılacak boylamsal (longitudinal) çalışmalarla, sürdürülebilirlik söyleminin yıllar içindeki evrimi takip edilmelidir.

Politika yapıcılara ve uygulayıcılara yönelik olarak ise, sürdürülebilirlik raporlamasının standardizasyonu için Yükseköğretim Kurulu (YÖK) öncülüğünde ulusal bir rehberin hazırlanması önerilmektedir. Raporların sayfa sayıları ve içerik formatlarındaki tutarsızlık, GRI standartlarıyla uyumlu asgari bir veri setinin belirlenmesini zorunlu kılmaktadır. Ayrıca üniversitelerin, sürdürülebilirlik faaliyetlerini birim bazlı "Proje" yönetiminden çıkarıp, Rektörlük seviyesinde koordine edilen "Stratejik" bir yönetim modeline geçmeleri gerekmektedir. Raporlamada ise nitel betimlemelerden ziyade, ölçülebilir Anahtar Performans Göstergelerine (KPI) dayalı veri sunumu teşvik edilmelidir. Çalışmanın bulguları Türkiye'de üniversitelerin sürdürülebilirlik süreçlerinde "proje odaklı" bir yaklaşımı benimsediğini ortaya koymaktadır. Gelecekte bu sürecin kurumsal bir stratejiye dönüştürülmesi için YÖK tarafından standart bir sürdürülebilirlik raporlama çerçevesi geliştirilmesi önerilmektedir.

Bu çalışma yalnızca Türkçe olarak yayımlanmış 26 sürdürülebilirlik raporunu kapsamaktadır. PDF formatından kaynaklanan OCR hataları, raporların farklı formatlarda hazırlanmış olması ve görsel içeriklerin analiz dışı kalması metodolojik sınırlılıklar oluşturmaktadır. Ayrıca İngilizce raporlar dışlandığı için kurumlar arası karşılaştırmalarda dil temelli bir örneklem yanlılığı oluşma ihtimali vardır. Bu sınırlılıklar gelecekte daha geniş kapsamlı veri setleriyle giderilebilir.

7. TEŞEKKÜR

Bu çalışma, İstanbul Esenyurt Üniversitesi Bilimsel Araştırma Projeleri Birimi (BAP) tarafından proje numarası BAP-2025/19-A ile desteklenmiştir.

8. KAYNAKÇA

- Adnan, K. ve Akbar, R. (2019). An analytical study of information extraction from unstructured and multidimensional big data. *Journal of Big Data*, 6, Article 91. <https://doi.org/10.1186/s40537-019-0254-8>
- Basheer, N., Ahmed, V., Bahroun, Z., & Anane, C. (2024). Exploring sustainability assessment practices in higher education: A comprehensive review through content and bibliometric analyses. *Sustainability*, 16(13), Article 5799. <https://doi.org/10.3390/su16135799>
- Caeiro, S., Hamón, L., Martins, R., & Aldaz, C. (2020). Sustainability assessment and benchmarking in higher education institutions—A critical reflection. *Sustainability*, 12(2), Article 543. <https://doi.org/10.3390/su12020543>
- Demir, S., & Topcu, B. (2022). Graph-based Turkish text normalization and its impact on noisy text processing. *Engineering Science and Technology, an International Journal*, 35, Article 101192. <https://doi.org/10.1016/j.jestch.2022.101192>
- Di Tullio, P., & Rea, M. (2025). Institutionalisation of sustainability in universities: Insights from strategic planning and sustainability reporting practices in Italian universities. *Meditari Accountancy Research*. Advance online publication. <https://doi.org/10.1108/medar-05-2024-2467>
- Filho, W. L., Coronado-Marín, A., Sálvia, A. L., Silva, F. M., Wolf, F., LeVasseur, T., Kirrane, M., Doni, F., Paço, A., Blicharska, M., Schmitz, M., Grahl, A., & Moggi, S. (2022). International trends and practices on sustainability reporting in higher education institutions. *Sustainability*, 14(19), Article 12238. <https://doi.org/10.3390/su141912238>
- Kalpokas, N., & Radivojevic, I. (2021). Bridging the gap between methodology and qualitative data analysis software: A practical guide for educators and qualitative researchers. *Sociological Research Online*, 27(2), 313–341. <https://doi.org/10.1177/13607804211003579>

- Kraus, M., & Feuerriegel, S. (2017). Decision support from financial disclosures with deep neural networks and transfer learning. *Decision Support Systems*, 104, 38–48. <https://doi.org/10.1016/j.dss.2017.10.001>
- Lakhno, M. (2024). Green or green-washed? Examining sustainability reporting in higher education. *Higher Education Quarterly*. Advance online publication. <https://doi.org/10.1111/hequ.12513>
- Lewis, C., & Young, S. (2019). Fad or future? Automated analysis of financial text and its implications for corporate reporting. *Accounting and Business Research*, 49(5), 587–615. <https://doi.org/10.1080/00014788.2019.1611730>
- Maher, C., Hadfield, M., Hutchings, M., & De Eyto, A. (2018). Ensuring rigor in qualitative data analysis. *International Journal of Qualitative Methods*, 17, Article 1609406918786362. <https://doi.org/10.1177/1609406918786362>
- Mo, F., & Wang, D. (2023). Emerging ESG reporting of higher education institutions in China. *Heliyon*, 9(11), Article e22527. <https://doi.org/10.1016/j.heliyon.2023.e22527>
- Mohammed, Y., Fadel, M., Abdalhamed, M., & Khudhair, A. (2025). Advancing SDGs: Towards effective sustainability reporting for green universities via XBRL and its impact on green auditing quality. *Journal of Lifestyle and SDGs Review*, 5(3), Article e05920. <https://doi.org/10.47172/2965-730x.sdgsreview.v5.n03.pe05920>
- Oflazer, K. (2014). Turkish and its challenges for language processing. *Language Resources and Evaluation*, 48(4), 639–653. <https://doi.org/10.1007/s10579-014-9267-2>
- Ong, K., Mao, R., Satapathy, R., Filho, R., Cambria, E., Sulaeman, J., & Mengaldo, G. (2024). Explainable natural language processing for corporate sustainability analysis. *Information Fusion*, 103, Article 102726. <https://doi.org/10.1016/j.inffus.2024.102726>
- Paulus, T., & Marone, V. (2024). “In minutes instead of weeks”: Discursive constructions of generative AI and qualitative data analysis. *Qualitative Inquiry*, 31(4), 395–402. <https://doi.org/10.1177/10778004241250065>
- Saillard, E. (2011). Systematic versus interpretive analysis with two CAQDAS packages: NVivo and MAXQDA. *Forum: Qualitative Social Research*, 12(1), Article 21. <https://doi.org/10.17169/fqs-12.1.1518>
- Sak, H., Güngör, T., & Saraçlar, M. (2011). Resources for Turkish morphological processing. *Language Resources and Evaluation*, 45(2), 249–261. <https://doi.org/10.1007/s10579-010-9128-6>
- Sepasi, S., Braendle, U., & Rahdari, A. (2019). Comprehensive sustainability reporting in higher education institutions. *Social Responsibility Journal*, 15(2), 155–170. <https://doi.org/10.1108/SRJ-01-2018-0009>

- Star, J., Ringaert, L., & Larcombe, L. (2025). Qualitative methods case study: Using MAXQDA in Indigenous HIV journey mapping research. *International Journal of Qualitative Methods*. Advance online publication. <https://doi.org/10.1177/16094069251356667>
- Stuckrath, C., Rosales-Carreón, J., & Worrell, E. (2025). Conceptualisation of Campus Living Labs for the sustainability transition: An integrative literature review. *Environmental Development*. <https://doi.org/10.1016/j.envdev.2025.101143>.
- Tohma, K., & Kutlu, Y. (2021). Challenges encountered in Turkish natural language processing studies. *Natural and Engineering Sciences*, 6(1), 1–14. <https://doi.org/10.28978/nesciences.833188>
- Toraman, C., Yılmaz, E., Şahinuç, F., & Özçelik, O. (2022). Impact of tokenization on language models: An analysis for Turkish. *ACM Transactions on Asian and Low-Resource Language Information Processing*, 22(1), 1–21. <https://doi.org/10.1145/3578707>
- Woods, M., Paulus, T., Atkins, D., & Macklin, R. (2016). Advancing qualitative research using qualitative data analysis software (QDAS)? Reviewing potential versus practice in published studies using ATLAS.ti and NVivo, 1994–2013. *Social Science Computer Review*, 34(5), 597–617. <https://doi.org/10.1177/0894439315596311>
- Yáñez, S., Uruburu, Á., Moreno, A., & Lumbreras, J. (2019). The sustainability report as an essential tool for the holistic and strategic vision of higher education institutions. *Journal of Cleaner Production*, 207, 57–65. <https://doi.org/10.1016/j.jclepro.2018.09.171>
- Zennaro, G., Corazza, G., & Zanin, F. (2024). The effects of integrated reporting quality: a meta-analytic review. *Meditari Accountancy Research*. <https://doi.org/10.1108/medar-09-2023-2175>.

Kurumsal / İnternet Kaynakları

- Global Reporting Initiative. (2025). *Global Sustainability Standards Board (GSSB)*. <https://www.globalreporting.org/standards/global-sustainability-standards-board/>
- Times Higher Education. (2025). *Impact rankings*. <https://www.timeshighereducation.com/impactrankings>
- UI GreenMetric. (2025). *Overall rankings*. <https://greenmetric.ui.ac.id/rankings/overall-rankings-2025>
- World Sustainable Development Goals Platform. (2025). *Küresel Amaçlar*. <https://www.kureselamaclar.org/>
- Yükseköğretim Kurulu. (2025). *YÖK resmi web sitesi*. <https://www.yok.gov.tr/tr>