

DİJİTAL TOPLUMDA YAPAY ZEKÂ ve ETİK: SOSYAL ADALET PERSPEKTİFİ
ARTIFICIAL INTELLIGENCE and ETHICS in the DIGITAL SOCIETY: A SOCIAL
JUSTICE PERSPECTIVE

Emrah ARĞIN

Doç. Dr., Görsel İletişim Tasarımı ABD, Malatya Turgut Özal Üniversitesi
Assoc. Prof. Dr., Visual Communication Design Dept., Malatya Turgut Özal Univ.
emrah.argin@ozal.edu.tr

ORCID: <https://orcid.org/0000-0002-3619-8034>

Makale bilgisi | Article Information

Makale Türü / Article Type: Derleme Makalesi / Review Article

Geliş Tarihi / Date Received: 04 Aralık / 04 Dec

Kabul Tarihi / Date Accepted: 13 Aralık / 13 Dec

Yayın Tarihi / Date Published: 14 Aralık / 14 Dec

Yayın Sezonu / Pub Date Season: Aralık 2025 / Dec 2025

DOI: <https://doi.org/10.5281/zenodo.17924656>

Bu Makaleye Atıf İçin / To Cite This Article: Arğın, E. (2025). Dijital toplumda yapay zekâ ve etik: Sosyal adalet perspektifi. *IJESOS International Journal of Educational and Social Sciences*, 4(2), 248–266.
<https://doi.org/10.5281/zenodo.17924656>

İntihal: Bu makale turnitin.com yazılımı ile yazar tarafından taranmıştır.
İntihal tespit edilmemiştir.

Plagiarism: This article has been scanned with turnitin.com by writer. No
plagiarism detected.

İletişim:

e-posta: editor@ijesos.com
Web: www.ijesos.com

Öz: Bu makale, dijital toplumda yapay zekâ (YZ) teknolojilerinin yükselişiyle ortaya çıkan etik ikilemleri, sosyal adalet perspektifiyle incelemeyi amaçlamaktadır. YZ'nin toplumsal fayda potansiyeli kabul edilirken, tarihsel ve yapısal eşitsizlikleri yeniden üretebilme riski vurgulanmaktadır.

Çalışma dört bölümden oluşmaktadır. Giriş kısmı konunun önemini ve kapsamını belirlemektedir. Birinci bölüm, sorumlu ve güvenilir YZ kavramlarını açıklayarak algoritmik karar süreçlerinde etik ikilemler, veri mahremiyeti, şeffaflık ve hesap verebilirlik ihtiyacını tartışmakta; ulusal ve küresel etik çerçeveler değerlendirilmektedir.

İkinci bölüm, sosyal adalet bağlamında algoritmik ayrımcılık ve önyargıyı ele almakta, ceza adaleti, istihdam ve kamu hizmetleri gibi alanlarda YZ'nin dezavantajlı gruplar üzerindeki etkisini göstermektedir. Teknik çözümlerin sınırlılıklarına dikkat çekilerek feminist ve eleştirel yaklaşımların önemi vurgulanmaktadır.

Üçüncü bölüm, önyargı, opaklık, hesap verebilirlik açığı, mahremiyet erozyonu ve gözetim gibi etik sorunları sistematik olarak analiz etmekte ve bunların birbirine bağlı olduğunu göstermektedir. Makine öğrenimi yaşam döngüsünün bütüncül ele alınması gerekliliği savunulmaktadır.

Sonuç ve öneriler kısmında, önyargı azaltma teknikleri, açıklanabilir YZ, mevzuat güçlendirmesi, AB Yapay Zekâ Yasası ile uyum, etik kurullar ve toplumsal farkındalık önerilmektedir. Makale, etik YZ kullanımının disiplinlerarası iş birliği, şeffaf kamu diyalogu ve insan onurunu merkeze alan yönetimle mümkün olacağını sonucuna varmaktadır.

Anahtar Kelimeler: Yapay zekâ etiği, Sosyal adalet, Algoritmik önyargı, Dijital toplum, Sorumlu yapay zekâ, Hesap verebilirlik.

Abstract: This article aims to examine the ethical dilemmas emerging with the rise of artificial intelligence (AI) technologies in the digital society from a social justice perspective. While acknowledging AI's potential for societal benefit, it highlights the risk of its capacity to reproduce historical and structural inequalities.

The study consists of four sections. The introduction establishes the importance and scope of the subject. The first section elucidates the concepts of responsible and trustworthy AI, discussing ethical dilemmas in algorithmic decision-making processes, data privacy, and the need for transparency and accountability; national and global ethical frameworks are also evaluated.

The second section addresses algorithmic discrimination and bias within the context of social justice, demonstrating AI's impact on disadvantaged groups in areas such as criminal justice, employment, and public services. Noting the limitations of technical solutions, it emphasizes the importance of feminist and critical approaches. The third section provides a systematic analysis of ethical issues such as bias, opacity, the accountability gap, privacy erosion, and surveillance, demonstrating their interconnected nature. It advocates for a holistic approach to the entire machine learning lifecycle.

The conclusion and recommendations section proposes bias mitigation techniques, explainable AI (XAI), strengthened legislation, alignment with the EU AI Act, ethical review boards, and public awareness. The article concludes that ethical AI use is only possible through interdisciplinary collaboration, transparent public dialogue, and governance that centres human dignity.

Key Words: Artificial intelligence Ethics, Social justice, Algorithmic bias, Digital society, Responsible AI, Accountability.

Giriş

Dijital teknolojilerin ve yapay zekânın (YZ) toplumsal hayatın her alanına nüfuz etmesi, beraberinde önemli etik ve sosyal sorunları da getirmiştir. Yapay zekâ sistemleri, sağlıktan eğitime, adalet dağıtımından istihdama kadar kritik karar süreçlerinde giderek daha merkezi bir rol üstlenirken, bu sistemlerin nasıl tasarlandığı, eğitildiği ve kullanıldığı temel bir etik inceleme konusu haline gelmiştir (Dignum, 2019: 3). Bu bağlamda, dijital toplumda etik, salt teknik bir mesele olmaktan çıkmış; insan onuru, özerklik, adalet ve şeffaflık gibi temel değerler etrafında şekillenen sosyoteknik bir paradigmaya dönüşmüştür (Floridi ve Cowls, 2019: 2).

Sosyal adalet perspektifi, yapay zekâ etiği tartışmalarının merkezine eşitsiz güç dinamiklerini, ayrımcılığı ve sistematik önyargıyı yerleştirir. Araştırmacılar, büyük veriyle beslenen algoritmik karar verme mekanizmalarının, tarihsel ve toplumsal eşitsizlikleri yeniden üreterek ve hatta güçlendirerek "matematiksel yıkım silahlarına" dönüşebileceği konusunda uyardırmaktadır (O'Neil, 2016: 21). Benzer şekilde Crawford (2021), YZ'nin planetaryer maliyetine ve mevcut iktidar yapılarını nasıl pekiştirdiğine dikkat çekerek, bu sistemlerin nötr olmaktan uzak olduğunu vurgulamaktadır (s. 8).

Özellikle, hassas gruplar üzerindeki olumsuz etkiler dikkat çekicidir. Eubanks (2018), otomatikleştirilmiş kamu hizmeti sistemlerinin yoksulları nasıl gözetlediğini, profillediğini ve cezalandırdığını ortaya koyarken (s. 12), Noble (2018) arama motoru algoritmalarının ırkçı ve cinsiyetçi kalıp yargıları nasıl güçlendirdiğini belgeler (s. 24). Bu durum, Benjamin'in (2019) "Yeni Jim Kod" olarak adlandırdığı, teknolojinin ırksal hiyerarşileri nasıl yeniden inşa ettiği olgusuyla örtüşmektedir (s. 59).

Türkiye bağlamında ise, 6698 sayılı Kişisel Verilerin Korunması Kanunu (2016) ve Cumhurbaşkanlığı Dijital Dönüşüm Ofisi (2021) tarafından yayımlanan Ulusal Yapay Zekâ Stratejisi, konuya ilişkin hukuki ve stratejik çerçeveyi oluşturmaktadır. Ancak, algoritmik adalet (Yaşar Ümütlü, 2025), dijital mahremiyet (Tanışık, 2023) ve çalışma hayatında algoritmik ayrımcılık (Oğuz, 2024) gibi alanlarda yapılacak daha fazla araştırmaya ihtiyaç duyulmaktadır. Türkiye'deki bu araştırma boşluğu, aynı zamanda sosyal medya platformları ve dijital iletişim araçlarındaki algoritmik sistemlerin toplumsal cinsiyet, kimlik ve önyargı temsilleri üzerine yapılan çalışmalarla da paralellik göstermektedir (Argın, 2023; Biçer ve Şener, 2020).

Bu makale, dijital toplumda yapay zekâ ve etik ilişkisini sosyal adalet odağında incelemeyi amaçlamaktadır. Çalışma, öncelikle YZ ve etik ilişkisini kuramsal bir çerçeveye oturtacak, ardından sosyal adalet bağlamında YZ'nin rolünü tartışacak, karşılaşılan temel etik sorunları (önyargı, ayrımcılık, hesap

verebilirlik, şeffaflık) literatür ışığında analiz edecek ve sonuç olarak mevcut ve potansiyel çözüm önerilerini değerlendirecektir.

Bu çalışma, dijital toplumda yapay zekâ uygulamalarının etik boyutlarını sosyal adalet perspektifinden ele almayı amaçlamaktadır. Bu doğrultuda araştırma aşağıdaki sorulara yanıt aramaktadır:

1. Yapay zekâ sistemlerinin tasarım ve kullanım süreçlerinde ortaya çıkan temel etik sorunlar nelerdir?
2. Algoritmik karar verme mekanizmaları, toplumsal eşitsizlikleri ve ayrımcılığı hangi yollarla yeniden üretmektedir?
3. Sosyal adalet perspektifi, yapay zekâ etiği tartışmalarına hangi kavramsal ve eleştirel katkıları sunmaktadır?
4. Ulusal ve uluslararası etik çerçeveler (AB Yapay Zekâ Yasası, OECD ve UNESCO ilkeleri vb.), algoritmik önyargı ve adaletsizlik sorunlarını ne ölçüde ele almaktadır?
5. Dijital toplumda daha adil, şeffaf ve hesap verebilir yapay zekâ uygulamaları geliştirmek için hangi etik ve yönetim temelli yaklaşımlar öne çıkmaktadır?

Yöntem

Bu çalışma, nitel araştırma desenine dayalı sistematik bir literatür taraması olarak tasarlanmıştır. Araştırma kapsamında yapay zekâ etiği, algoritmik adalet ve sosyal adalet ilişkisini ele alan ulusal ve uluslararası akademik çalışmalar incelenmiştir. Literatür taraması, yapay zekâ etiği alanında temel kabul gören kuramsal çerçeveler ile eleştirel ve disiplinlerarası yaklaşımları kapsayacak şekilde yapılandırılmıştır.

Literatür tarama sürecinde; *"artificial intelligence ethics"*, *"algorithmic bias"*, *"social justice"*, *"responsible AI"*, *"algorithmic discrimination"* ve *"explainable AI"* gibi anahtar kavramlar kullanılmış; Springer, Taylor & Francis, Elsevier, ACM Digital Library, Google Scholar ve DergiPark gibi akademik veri tabanlarından yararlanılmıştır. Çalışmaların seçiminde hakemli dergilerde yayımlanmış olması, alanın temel eserleri arasında yer alması ve yapay zekânın toplumsal etkilerine eleştirel bir perspektif sunması temel ölçütler olarak benimsenmiştir.

Seçilen kaynaklar tematik analiz yöntemiyle incelenmiş; önyargı, ayrımcılık, şeffaflık, hesap verebilirlik, mahremiyet ve yönetim gibi ana temalar altında sınıflandırılmıştır. Bu temalar doğrultusunda literatürde öne çıkan kuramsal yaklaşımlar, etik çerçeveler ve politika önerileri karşılaştırmalı olarak değerlendirilmiştir.

Bu yöntemsel yaklaşım, yapay zekâ etiğine ilişkin mevcut akademik tartışmaları sistematik bir bütünlük içinde analiz etmeyi ve sosyal adalet perspektifinin bu tartışmalara sunduğu katkıları görünür kılmayı amaçlamaktadır.

1. Yapay Zekâ ve Dijital Toplumda Etik

Yapay zekânın (YZ) dijital toplumun ayrılmaz bir parçası haline gelmesi, insanlık tarihinde benzeri görülmemiş etik, hukuki ve sosyolojik soru(n)ları da beraberinde getirmiştir. Bu bölüm, YZ'nin dijital toplumdaki etik boyutunu, sorumlu YZ kavramından başlayarak, algoritmik karar verme süreçlerinin yarattığı etik ikilemlere, veri etiği ve mahremiyetine, şeffaflık ve hesap verebilirlik ihtiyacına ve nihayetinde bu sorunlara yönelik geliştirilen küresel ve ulusal etik çerçevelere kadar geniş bir yelpazede incelemeyi amaçlamaktadır.

1.1. Sorumlu Yapay Zekâ ve Temel Etik İlkeler

Dijital dönüşümün hızı, YZ sistemlerinin toplumsal kabulü ve sürdürülebilirliği için etik bir temelin zorunlu olduğu gerçeğini ortaya çıkarmıştır. Dignum (2019), "sorumlu yapay zekâ" kavramını merkeze alarak, YZ'nin geliştirilme ve kullanılma süreçlerinin yalnızca teknik değil, aynı zamanda hukuki, etik ve sosyo-ekonomik boyutlarıyla ele alınması gerektiğini vurgular (s. 15). Sorumlu YZ, sistemlerin insan refahına hizmet etmesini, insan özerkliğine saygı duymasını ve adil, şeffaf olmasını garanti altına alma taahhüdüdür. Floridi ve Cowls (2019), bu taahhüdü operasyonel hale getirmek için beş temel ilkeden oluşan birleştirici bir çerçeve önermiştir: *yardım etmemek (non-maleficence)*, *yarar sağlama (beneficence)*, *özerklik (autonomy)*, *adalet (justice)* ve *açıklanabilirlik (explicability)* (s. 4). Bu ilkeler, YZ etiği literatürünün temel taşlarını oluşturur ve çoğu ulusal ve uluslararası kılavuzun da omurgasını meydana getirir.

Avrupa Komisyonu Yapay Zekâ Üst Düzey Uzman Grubu (2019) ise "güvenilir YZ"yi dört ana bileşene dayandırmaktadır: (1) yasalılık (mevcut yasalara uyum), (2) etiklik (etik ilkelere uyum), (3) sağlamlık (teknik sağlamlık ve güvenlik) ve (4) insan merkezilik. Bu yaklaşım, YZ'nin yalnızca riskleri en aza indiren değil, aynı zamanda demokratik değerleri ve temel hakları aktif olarak destekleyen bir sistem olması gerektiğinin altını çizer. Benzer şekilde IEEE (2019) tarafından yayımlanan "Etik Amaçlı Tasarım" belgesi, otonom ve akıllı sistemlerin insan refahını nasıl önceliklendireceğine dair kapsamlı bir vizyon sunarak, insan haklarının, insan onurunun ve teknolojinin insanlığın yararına kullanılmasının altını çizer.

1.2. Algoritmik Karar Verme ve Etik İkilemler

YZ'nin dijital toplumdaki en belirgin etkisi, algoritmaların karar verme süreçlerinde giderek daha fazla rol almasıdır. Mittelstadt ve diğerleri (2016), algoritmaların etiği tartışmasını haritalandırdıkları çalışmalarında, bu sistemlerin yarattığı temel etik sorunları şu şekilde kategorize etmiştir: *önyargı ve ayrımcılık olasılığı (en: bias)*, *opaklık (en: opacity)*, *hesap verebilirlik (en: accountability)* ve *ölçeklenebilirlik (en: scalability)* ile gelen etki (s. 5). Algoritmalar, insan karar vericilerden farklı

olarak, çok daha büyük ölçekte ve yüksek hızda karar verebilir, ancak bu kararların arkasındaki mantık çoğu zaman bir "kara kutu" olarak kalır (Wachter, Mittelstadt ve Floridi, 2017: 78).

Bu opaklık, özellikle bireylerin kendilerini etkileyen otomatikleştirilmiş kararlara itiraz etme hakkını kullanmasını engeller. Selbst ve Powles (2017), GDPR'daki "açıklama hakkı"nın pratikteki sınırlarını tartışarak, yasal metinlerdeki "anamlı bilgi" kavramının teknik olarak nasıl karşılanabileceği sorusunu gündeme getirmiştir (s. 235). Doshi-Velez ve Kim (2017) ise, açıklanabilir makine öğreniminin (en: explainable ML) bilimsel olarak nasıl daha titiz bir hale getirilebileceğini araştırmış, sadece doğru sonuçlar üreten değil, aynı zamanda nasıl ve neden o sonuca vardığı anlaşılabilen sistemlerin geliştirilmesi gerektiğini savunmuştur.

1.3. Veri Etiği, Gözetim ve Mahremiyetin Dönüşümü

YZ sistemleri büyük veri kümeleri üzerine inşa edilir ve işler. Dolayısıyla, YZ etiği aynı zamanda bir veri etiği sorunudur. Zuboff (2019), "gözetim kapitalizmi" kavramıyla, kişisel deneyimlerin davranışsal verilere dönüştürülerek ticarileştirildiği, tahmin ve kontrol amacıyla kullanıldığı yeni bir ekonomik mantığı tanımlar (s. 15). Bu sistemde birey, bir vatandaş veya müşteri olmaktan ziyade, ham veri kaynağına indirgenir. Crawford (2021) ise bu durumu daha da genişleterek, YZ'nin sadece dijital bir sistem olmadığını, madencilikten üretim zincirlerine, içerik moderasyonundan askeri uygulamalara kadar uzanan ve gezegensel bir maliyeti olan bir altyapılar bütünü olduğunu iddia eder (s. 8). Bu perspektif, YZ etiği tartışmalarını salt yazılımsal bir mesele olmaktan çıkarıp, politik ekoloji ve küresel adalet perspektiflerini de içerecek şekilde genişletir.

Mahremiyet, bu bağlamda korunması gereken statik bir hak olmaktan çıkıp, kişisel veriler üzerindeki kontrol ve özerkliğin sürekli müzakere edildiği dinamik bir pratiğe dönüşmüştür. Avrupa Birliği'nin Genel Veri Koruma Tüzüğü (GDPR, 2016), otomatik karar verme süreçlerinde bireyin haklarını güçlendirmeye yönelik önemli düzenlemeler getirmiştir. Türkiye'de ise 6698 sayılı Kişisel Verilerin Korunması Kanunu (2016) benzer bir işlev görmekte, kişisel verilerin işlenmesinde açıklık, adillik ve hukuka uygunluk gibi ilkeleri şart koşmaktadır. Tanışık ve Bal (2024: 10), KVKK özelinde yaptığı değerlendirmede, dijital mahremiyetin korunmasında kurumsal sorumluluğun önemine işaret ederek, yasal uyumluluğun ötesine geçen etik bir yaklaşımın benimsenmesi gerektiğini savunur.x

1.4. Dijital Toplumda Etik için Yönetişim ve Düzenleyici Çerçeveler

YZ etiğine ilişkin teorik tartışmalar, somut yönetim ve düzenleme çerçeveleriyle desteklenmelidir. OECD (2019) Yapay Zekâ Tavsiyesi ve UNESCO (2021) Yapay Zekâ Etiği Tavsiyesi, üye ülkeler için insan merkezli, adil ve şeffaf bir YZ gelişimini teşvik eden uluslararası düzeyde kabul görmüş

ilkeler setleri sunar. Bu küresel çabaların yanı sıra, bölgesel düzeyde en kapsamlı düzenleme olan Avrupa Birliği Yapay Zekâ Yasası (AI Act, 2024) kabul edilmiştir. Bu yasa, YZ uygulamalarını kabul edilemez risk, yüksek risk, sınırlı risk ve minimal risk olarak kategorilere ayıran, risk temelli bir yaklaşım benimsemektedir. Yüksek riskli kategorideki uygulamalar (örn. kritik altyapı, eğitim, istihdam) için temel haklara saygı, veri kalitesi, risk yönetimi, insan gözetimi ve doğruluk gibi katı yükümlülükler öngörür.

Türkiye'de ise Cumhurbaşkanlığı Dijital Dönüşüm Ofisi (2021) tarafından yayımlanan "Ulusal Yapay Zekâ Stratejisi (2021-2025)", YZ'nin etik ve güvenli kullanımını beş ana hedefinden biri olarak belirlemiştir. Strateji belgesi, etik ilkelerin oluşturulması, mevzuat altyapısının güçlendirilmesi ve insan haklarına saygılı bir YZ ekosistemi geliştirilmesi taahhüdünde bulunur. Türkiye Yapay Zekâ İnisiyatifi (TRAI) (2024) de bu yöndeki çabaları destekleyen, etik ilkeler ve hukuki düzenlemeler üzerine raporlar yayımlayan sivil toplum tabanlı bir girişimdir.

Ancak, mevcut düzenleyici çerçeveler ve etik kılavuz ilkeler, hızla evrilen bu alanda her zaman yeterli kalmayabilir. Raji ve diğerleri (2020), YZ hesap verebilirliği açığını kapatmak için bağımsız denetim ve sertifikasyon mekanizmalarının kritik önem taşıdığını vurgulamaktadır. Suresh ve Guttat (2021) ise, makine öğrenimi yaşam döngüsünün her aşamasında (veri toplama, model seçimi, dağıtım, izleme) ortaya çıkabilecek zarar kaynaklarını anlamak için bütüncül bir çerçeve sunar. Bu yaklaşımlar, etiğin YZ sistemlerinin yaşam döngüsüne proaktif olarak entegre edilmesi gerektiğini gösterir.

1.5. Dijital Toplum İçin Etik Bir Temel Oluşturmak

Dijital toplumda yapay zekâ ve etik ilişkisi, disiplinlerarası, çok paydaşlı ve sürekli bir diyalog ve dikkatli bir düzenleme gerektiren karmaşık bir alandır. Teknik ilerleme, toplumsal değerler ve hukuki güvenceler arasında bir denge kurulması esastır. Floridi ve Cows (2019) tarafından öne sürülen beş ilke ve benzeri çerçeveler, bu dengenin sağlanması için değerli bir yol haritası sunar. Ancak, bu ilkelerin pratiğe geçirilmesi, algoritmik adalet (Yaşar Ümütlü, 2025), şeffaflık ve hesap verebilirlik mekanizmaları (Raji vd., 2020) ve insan hakları odaklı bir yaklaşım (Akıncılar Köseoğlu ve Çetin, 2024) olmadan mümkün değildir.

Dijital toplumun geleceği, YZ sistemlerinin ne kadar akıllı olduğundan ziyade, bu sistemleri nasıl bir etik temel üzerine inşa ettiğimiz, toplumsal adaleti ve insan onurunu nasıl gözettiğimiz ile şekillenecektir. Bu, yalnızca mühendislerin ve hukukçuların değil, aynı zamanda sosyal bilimcilerin, filozofların, sivil toplumun ve tüm vatandaşların içinde aktif bir rol alması gereken kolektif bir sorumluluktur.

2. Sosyal Adalet Bağlamında Yapay Zekâ

Yapay zekânın (YZ) toplumsal hayata entegrasyonu, sosyal adalet kavramını yeniden tanımlamakta ve derinlemesine bir sorgulama gerektirmektedir. Sosyal adalet, kaynakların, fırsatların ve toplumsal imtiyazların dağılımında eşitliği ve hakkaniyeti savunur. Ancak, YZ sistemleri, tarihsel ve yapısal eşitsizlikleri yeniden üretme, hatta ölçeklendirerek büyütme potansiyeli taşımaktadır. Bu bölüm, YZ'nin sosyal adalet üzerindeki etkisini, algoritmik ayrımcılık ve önyargı örnekleri üzerinden inceleyecek, bu sorunların farklı toplumsal gruplar ve sektörlerdeki yansımalarını analiz edecek ve adil YZ sistemlerine ulaşmak için öne sürülen teknik ve yapısal çözüm önerilerini değerlendirecektir.

2.1. Algoritmik Ayrımcılık ve Önyargının Anatomisi

Algoritmik ayrımcılık, bir YZ sisteminin adil olmayan, taraflı veya belirli bir gruba karşı ayrımcı sonuçlar üretmesi durumudur. Bu ayrımcılık, çoğu zaman kasıtlı olmaktan ziyade, sistemin tasarımından, eğitim verilerinden veya uygulama bağlamından kaynaklanır. Mittelstadt ve diğerleri (2016), algoritmaların "ölçeklenebilir" olma özelliğinin, var olan önyargıları çok daha geniş kitlelere çok daha hızlı bir şekilde yayma kapasitesi olduğuna dikkat çeker (s. 7). O'Neil (2016), bu tür sistemleri "Matematiksel Yıkım Silahları" (Weapons of Math Destruction - WMDs) olarak adlandırır ve onları şu özelliklere sahip olarak tanımlar: ölçeklenebilirlik, gizlilik (opaklık) ve yıkıcılık (s. 31). Bu silahlar, kredi skorlamadan işe alım süreçlerine, okul kabulünden polislik uygulamalarına kadar uzanan bir yelpazede, en savunmasız grupları hedef alır.

Önyargının temel kaynağı genellikle eğitim verisinin kendisidir. Tarihsel olarak ayrımcılığa uğramış gruplar, veri kümelerinde ya yeterince temsil edilmezler ya da bu gruplarla ilgili mevcut olumsuz kalıp yargılar ve ayrımcı kararlar veri setine işlenmiştir. Model, bu tarihsel adaletsizlikleri öğrenir ve onları gelecekteki kararlar için normatif bir gerçeklik haline getirir. Noble (2018: 24), arama motoru algoritmalarının yaptığı tam olarak budur; ırksal ve cinsiyetçi stereotipleri, "nesnel" bir algoritmik otorite görünümü altında güçlendirir ve meşrulaştırır. Örneğin, siyahi kadınların arama sonuçlarında nasıl olumsuz şekilde temsil edildiğini gösteren çalışması, teknolojinin tarafsız olmadığını, toplumsal iktidar ilişkilerini yansıttığını kanıtlar.

Benjamin (2019) ise bu fenomeni "Yeni Jim Kod" (New Jim Code) olarak adlandırır. Jim Crow yasaları, ırksal ayrımcılığı açıkça yasalaştırırken, Yeni Jim Kod, ırksal hiyerarşileri ve ayrımcılığı, teknolojik olarak "ileri", "veriye dayalı" ve bu nedenle de "nesnel" ve "tarafsız" görünen sistemler aracılığıyla yeniden üretir (s. 59). Bu, ayrımcılığın daha incelikli, teknik olarak meşrulaştırılmış ve dolayısıyla tespit edilmesi ve mücadele edilmesi daha zor bir forma bürünmesidir. Bu durum, sadece arama motorlarıyla sınırlı kalmayıp, popüler

kültür ve medya temsilleri aracılığıyla da toplumsal kalıp yargıların algoritmik sistemlere beslenmesine ve normalleştirilmesine neden olmaktadır (Noble, 2018; Şener, 2021).

2.2. Sektörel Örnekler: Adalet, İstihdam ve Kamu Hizmetleri

Algoritmik ayrımcılığın somut etkileri çeşitli sektörlerde açıkça görülebilir:

- **Ceza Adalet Sistemi:** Suç işleme riskini tahmin eden algoritmalar (recidivism prediction instruments), en çok tartışılan konulardan biridir. ProPublica'nın 2016'da yaptığı ünlü araştırmayı temel alan çalışmalar, COMPAS gibi araçların, siyahi sanıkların yeniden suç işleme riskini beyaz sanıklara kıyasla tutarsız bir şekilde yüksek tahmin etme eğiliminde olduğunu göstermiştir. Chouldechova (2017), bu tür araçlardaki yanlılığın istatistiksel olarak incelendiği çalışmasında, adil tahmin yapmanın istatistiksel açıdan içsel zorluklarını (trade-offs) ortaya koyar (s. 1400). Kleinberg, Mullainathan ve Raghavan (2017) ise, istatistiksel adalet ölçütleri (grupsal eşitlik, bireysel adillik vb.) arasında temel ve kaçınılmaz ödünleşimler (inherent trade-offs) olduğunu matematiksel olarak kanıtlayarak, "tamamen adil" bir algoritmanın teorik olarak imkansız olabileceğini öne sürer.
- **İstihdam ve İşe Alım:** Özgeçmiş tarama yazılımları, belirli anahtar kelimelere, okullara veya deneyimlere öncelik vererek, nitelikli olsalar bile geleneksel olarak dezavantajlı grupların (örneğin, belirli bir sosyo-ekonomik çevreden gelenler, kadınlar, azınlıklar) elenmesine neden olabilir. Oğuz (2024), çalışma hayatında algoritmik ayrımcılığı incelediği makalesinde, işe alım algoritmalarının cinsiyet, yaş ve engellilik durumu temelinde ayrımcılık yapma potansiyelini Türk hukukundaki eşitlik ilkeleri ışığında değerlendirir ve mevzuatın bu yeni ayrımcılık biçimlerine karşı güçlendirilmesi gerektiğini savunur.
- **Kamu Hizmetleri ve Refah Sistemleri:** Eubanks (2018), "Otomatikleştirilmiş Eşitsizlik" çalışmasında, ABD'de kamu konutları, gıda yardımı ve çocuk koruma hizmetlerinde kullanılan otomatik karar verme sistemlerinin, yoksulları ve işçi sınıfını nasıl sürekli olarak gözetlediğini, profillediğini ve cezalandırdığını güçlü vaka çalışmalarıyla ortaya koyar (s. 12). Bu sistemler, yoksulluğu bireysel bir başarısızlık veya ahlaki bir kusur olarak kodlayarak, yapısal sorunları görünmez kılar ve yardım almayı neredeyse imkânsız hale getiren bürokratik engeller yaratır.

Algoritmik adaletsizliğe yönelik teknik çözümler, genellikle "adalet" kavramının makine öğrenmesi modellerine nasıl entegre edileceği sorusuna odaklanır. Bu alanda çeşitli "adillik ölçütleri" (fairness metrics) geliştirilmiştir. Dwork ve diğerleri (2012), "farkındalık yoluyla adalet" (fairness through awareness) kavramını öne sürerek, hassas özellikler (ırk, cinsiyet) kullanılmadan da bireyleri benzerliklerine göre gruplayarak adil sonuçlar

retmenin yollarını arařtırmıřtır (s. 218). Bu, doęrudan ayrımcılıęı engellemeyi amalayan bir yaklařımdır.

Ancak, teknik özmlerin sınırlılıkları vardır. Selbst ve dięerleri (2019), adaletin sadece bir model sorunu olmadıęını, aynı zamanda bir "sosyoteknik sistem" sorunu olduęunu savunur. Modeli evreleyen veri toplama sreleri, insan etkileřimleri, kurumsal yapılar ve toplumsal baęlam, nihai sonucu derinden etkiler (s. 60). Sadece modeli "dzeltmeye" alıřmak, sorunun kkenindeki arpık veri yapılarını veya adaletsiz sosyal baęlamı grmezden gelmek anlamına gelebilir. Barocas ve Selbst (2016), byk verinin farklı etkisini (disparate impact) inceledikleri makalede, en iyi niyetle tasarlanmış sistemlerin bile, istatistiksel iliřkiler nedeniyle dolaylı ve istemsiz ayrımcılıęa yol aabileceęini detaylandırır (s. 675).

Sosyal adalet perspektifinden bakan eleřtirel ve feminist yaklařımlar, sorunu daha kkl bir řekilde ele alır. D'Ignazio ve Klein (2020), "Veri Feminizmi"nde, veri biliminin uzun sredir glnn perspektifiyle řekillendięini, ezilenlerin bilgisini ve deneyimini dıřladıęını savunur. Onlara gre zm, veriye ve veri uygulamalarına etik ve politik bir mercekten bakmak, erkek egemenlięini, beyaz stnlęn ve kapitalist tahakkm sorgulayan ilkeleri benimsemektir (s. 6). Bu, "adil bir algoritma" yazmaktan ziyade, verinin nasıl toplandıęı, kimin tarafından tanımlandıęı ve kimin yararına kullanıldıęı gibi temel soruları sormak anlamına gelir.

Bu yaklařım, katılımcı tasarım, topluluk denetimi ve algoritmik řeffaflık taleplerini ierir. Bu, Zuboff'un (2019) tanımladıęı gzetim kapitalizminin mantıęına doęrudan bir meydan okumadır. Crawford (2021) ise, YZ'nin gezegensel maliyetine (madencilik, enerji tketimi, elektronik atık) dikkat ekerek, sosyal adalet mcadelesinin evre adaleti mcadelesinden ayrılamayacaęını vurgular (s. 8). Gerek bir sosyal adalet, YZ'nin sadece ıktıları iin deęil, tm yařam dngs ve onu mmkn kılan maddi kořullar iin de sorumluluk almayı gerektirir.

Trkiye'de, YZ ve sosyal adalet kesiřimi henz emekleme ařamasındadır. Ulusal Yapay Zekâ Stratejisi (2021-2025), etik ve insan merkezcilik vurgusu yapmakla birlikte, sosyal adalet ve algoritmik ayrımcılıkla mcadeleye dair somut adımlar ve eylem planları konusunda daha detaylı bir ereveye ihtiya duymaktadır. Trkiye Yapay Zekâ İnisyatifi (TRAI) (2024) gibi sivil toplum giriřimleri, bu alandaki bořluęu doldurmak iin aba gstermektedir.

Trk mevzuatı, zellikle 6698 sayılı KVKK, kiřisel verilerin iřlenmesinde "hakkaniyet ve řeffaflık" ilkesini getirerek nemli bir koruma saęlar. Ancak, bu hkmlerin otomatik karar verme ve profil ıkarma gibi spesifik YZ uygulamaları karřısında nasıl etkin bir řekilde uygulanacaęı, henz netlik kazanmamıřtır. Tanıřık ve Bal (2025), KVKK'nın dijital mahremiyeti

korumadaki rolünü vurgularken, aynı zamanda kurumsal sorumluluğun önemine işaret eder (s. 12).

Sonuç olarak, sosyal adalet bağlamında YZ, teknik bir problem olmaktan çok, politik ve etik bir meseledir. Algoritmik sistemler, toplumdaki mevcut güç dinamiklerini yansıtır ve güçlendirir. Bu nedenle, adil, kapsayıcı ve hesap verebilir YZ sistemlerine ulaşmak, yalnızca daha iyi algoritmalar geliştirmekle değil, aynı zamanda daha adil bir toplum inşa etme çabasıyla mümkündür. Bu, veri feminizminin ilkelerini benimsemeyi, toplulukları sürece dahil etmeyi, güçlü denetim ve regülasyon mekanizmaları oluşturmayı ve nihayetinde, teknolojinin insanlığın ortak iyiliği için nasıl kullanılabileceğine dair kolektif ve demokratik bir tartışma yürütmeyi gerektirir.

3. Karşılaşılan Temel Etik Sorunlar

Yapay zekâ (YZ) sistemlerinin dijital toplumda yaygınlaşması, bir dizi karmaşık, iç içe geçmiş ve çok boyutlu etik sorunu beraberinde getirmiştir. Bu sorunlar, salt teknik hatalar olmanın ötesinde, toplumsal yapıları, temel hakları ve insani değerleri derinden etkileme potansiyeline sahiptir. Bu bölüm, YZ'nin neden olduğu veya şiddetlendirdiği temel etik problemleri – önyargı ve ayrımcılık, şeffaflık ve açıklanabilirlik eksikliği, hesap verebilirlik açığı, mahremiyet erozyonu ve gözetim – sistematik bir şekilde incelemeyi ve bu sorunların birbiriyle olan ilişkisini ortaya koymayı amaçlamaktadır.

3.1. Önyargı ve Ayrımcılığın Çok Boyutlu Doğası

Önyargı ve ayrımcılık, YZ etiği literatüründe en çok tartışılan ve en yaygın sorunlardan biridir. Bu sorun, tek bir nedene bağlı olmayıp, makine öğrenimi yaşam döngüsünün (ML) neredeyse her aşamasında ortaya çıkabilir. Suresh ve Guttag (2021), bu zarar kaynaklarını anlamak için bütüncül bir çerçeve sunar ve önyargının *veri temelli*, *ölçüm temelli*, *model temelli* ve *kullanım temelli* olmak üzere birden fazla kaynağı olabileceğini vurgular (s. 3).

- **Veri Temelli Önyargı:** Tarihsel veriler, toplumdaki mevcut eşitsizlikleri, ayrımcı uygulamaları ve sistematik önyargıları yansıtır. Bir model bu verilerle eğitildiğinde, geçmişteki adaletsizlikleri öğrenir ve gelecekteki kararlarda pekiştirir. Bu durum, Noble'un (2018: 24) arama motoru algoritmalarında gösterdiği gibi, ırkçı ve cinsiyetçi stereotipleri güçlendirir. Benzer şekilde, Eubanks (2018: 45), kamusal yardım sistemlerinde kullanılan verilerin, yoksulları "potansiyel dolandırıcılar" olarak damgalayan bir önyargıyı nasıl içerdiğini ve bu şüphenin algoritmalara nasıl kodlandığını ortaya koyar.
- **Kullanım Temelli Önyargı:** Bazen model teknik olarak "tarafsız" olsa bile, yanlış bağlamda veya yanlış amaçla kullanıldığında ayrımcı sonuçlar doğurabilir. Örneğin, bir hastane taburculuk riskini tahmin etmek için tasarlanmış bir model, sağlık hizmetlerine erişimde eşitsizliği tespit etmek yerine, bu eşitsizliği meşrulaştırmak için kullanılabilir. Selbst ve diğerleri

(2019), bu soruna dikkat çekerek, adaletin yalnızca algoritmanın kendisinde değil, onu çevreleyen sosyoteknik sistemin tamamında aranması gerektiğini savunur (s. 62).

Barocas ve Selbst (2016: 675), "farklı etki" (disparate impact) kavramını detaylandırarak, nötr görünen bir uygulamanın belirli bir grubu orantısız bir şekilde olumsuz etkileyebileceğini hukuki bir perspektifle inceler. Bu, kasıt gerektirmeyen ancak yine de ciddi ayrımcı sonuçlar doğurabilen bir durumdur.

3.2. Şeffaflık ve Açıklanabilirlik Eksikliği: "Kara Kutu" Problemi

Pek çok gelişmiş YZ modeli, özellikle derin öğrenme ağları, içsel karmaşıklıkları nedeniyle bir "kara kutu" olarak nitelendirilir. Yani, bir kararı *neye dayanarak* verdiğini anlamak ve yorumlamak son derece zordur. Bu opaklık, temel bir etik ve hukuki sorun teşkil eder.

Wachter, Mittelstadt ve Floridi (2017), GDPR'da sıklıkla atıfta bulunulan "açıklama hakkı"nın aslında metinde doğrudan ve mutlak bir hak olarak yer almadığını; bunun yerine, bireylere "anamlı bilgi" sağlama yükümlülüğünün bulunduğunu tartışır (s. 82). Ancak, anlamlı bilginin teknik olarak ne anlama geldiği belirsizliğini korumaktadır. Selbst ve Powles (2017) ise, bu hukuki belirsizliğin pratikteki zorluklarını vurgulayarak, yasal düzenlemelerin algoritmik açıklanabilirlik için yeterli çerçeveyi sunmakta nasıl yetersiz kaldığını ele alır (s. 238).

Açıklanabilirliğin önemi sadece bir hesap verebilirlik meselesi değil, aynı zamanda bir güven meselesidir. Doshi-Velez ve Kim (2017), açıklanabilir makine öğrenimini daha titiz bir bilim dalı haline getirmenin yollarını araştırır. Kullanıcıların, hastaların, vatandaşların ve diğer paydaşların, kendileri hakkında verilen ve hayatlarını derinden etkileyen kararların arkasındaki mantığı anlayabilmeleri, bu sistemlere olan güvenin temel taşıdır.

3.3. Hesap Verebilirlik Açığı: Sorun Çıktığında Sorumlu Kim?

Algoritmik karar verme süreçlerinin karmaşıklığı, geleneksel sorumluluk zincirlerini bulanıklaştırmaktadır. Bir algoritmik karar adil değilse veya bir zarara yol açıyorsa, sorumlu kimdir? Veriyi toplayan mı, modeli eğiten veri bilimci mi, modeli satın alıp kullanan kurum mu, yoksa algoritmanın kendisi mi? Bu "hesap verebilirlik açığı", YZ sistemlerinin denetimini ve regülasyonunu zorlaştıran temel engellerden biridir.

Raji ve diğerleri (2020), bu açığı kapatmanın yolunun, bağımsız denetim (auditing) ve sertifikasyon mekanizmalarını tanımlamak ve uygulamaktan geçtiğini savunur. Tıpkı finansal denetimler veya gıda güvenliği denetimleri gibi, YZ sistemlerinin de performanslarının, adilliklerinin ve güvenilirliklerinin düzenli olarak bağımsız kuruluşlar tarafından denetlenmesi gerektiğini önerirler (s. 5). Bu, sorumluluğu dağıtmak yerine, netleştirmeye ve kurumsal düzeyde yerleştirmeye yönelik bir çağrıdır.

3.4. Mahremiyetin Aşınması ve Gözetim Kapitalizmi

YZ, büyük veriyle beslenir ve bu verinin büyük kısmı kişisel, hatta hassas verilerden oluşur. Bu durum, bireylerin mahremiyetini ve kişisel verileri üzerindeki kontrolünü ciddi şekilde tehdit eder. Zuboff (2019), "gözetim kapitalizmi" kavramıyla, kişisel deneyimlerin davranışsal verilere dönüştürülerek ticarileştirildiği, tahmin ve kontrol amacıyla kullanıldığı yeni bir ekonomik mantığı tanımlar (s. 15). Bu sistemde, mahremiyet, korunması gereken bir insan hakkı olmaktan çıkıp, gözetim kapitalistleri için çözülmesi gereken bir teknik soruna indirgenir.

Crawford (2021) ise, bu gözetim ekonomisinin gezegensel boyutuna dikkat çeker. YZ sistemlerinin, veri merkezlerinden madencilik operasyonlarına kadar uzanan fiziksel altyapısı, gözetimin sadece dijital değil, aynı zamanda maddi bir boyutu olduğunu gösterir (s. 8). Bu, mahremiyet ihlallerinin çevresel adaletsizliklerle kesiştiği noktadır.

Türkiye bağlamında, 6698 sayılı Kişisel Verilerin Korunması Kanunu, kişisel verilerin işlenmesinde açıklık, adillik ve hukuka uygunluk gibi ilkeler getirerek önemli bir hukuki koruma sağlar. Tanışık ve Bal (2025: 12), KVKK'nın dijital mahremiyeti korumadaki rolünü vurgularken, yasal uyumluluğun ötesine geçen etik bir kurumsal sorumluluk anlayışının benimsenmesi gerektiğini savunur.

3.5. Diğer Önemli Etik Sorunlar

Yukarıdaki ana başlıkların yanı sıra, bir dizi başka kritik etik sorun da bulunmaktadır:

- **Özerklik ve İnsan Gözetimi:** YZ sistemleri insan karar verme süreçlerini ne ölçüde devralmalıdır? Floridi ve Cowls (2019), insan özerkliğinin korunmasını temel ilkelerden biri olarak belirler (s. 5). Kritik kararlarda nihai yetkinin ve sorumluluğun her zaman bir insanda olması gerektiği, "insan gözetimi" (human-in-the-loop) kavramıyla vurgulanır.
- **İşin Geleceği ve Ekonomik Eşitsizlik:** Otomasyonun yaygınlaşması, kitlesel işsizlik ve gelir eşitsizliği korkularını beraberinde getirmiştir (Demirel vd., 2023: 81). O'Neil (2016), YZ sistemlerinin sadece bireyleri değil, bütün meslek gruplarını ve sektörleri nasıl dönüştürdüğünü ve bu dönüşümün sosyo-ekonomik maliyetlerinin nasıl adaletsiz dağıldığını tartışır (s. 125).
- **Güvenlik ve Sosyal Manipülasyon:** Otonom silah sistemleri veya derin sahtekarlık (deepfake) teknolojileri gibi kötü niyetle kullanıma açık YZ uygulamaları, fiziksel güvenliği ve toplumsal güveni tehdit etmektedir. Dignum (2019: 89), sorumlu YZ geliştirmenin, sistemlerin kötüye kullanımını engellemek için gerekli teknik ve yönetim önlemlerini almayı da içerdiğini belirtir.

Karşılaşılan bu temel etik sorunlar birbirinden bağımsız değildir; aksine, derinden iç içe geçmiş ve birbirini besleyen bir yapı oluşturur. Opaklık, hesap verebilirliği engeller. Hesap verebilirliğin olmaması, ayrımcılığın denetlenememesine ve sürmesine yol açar. Ayrımcı veri kümeleri, gözetim kapitalizminin temelini oluşturur. Bu nedenle, bu sorunlara parçacıl ve izole çözümler aramak yetersiz kalacaktır.

Suresh ve Gutttag'ın (2021) önerdiği gibi, makine öğrenimi yaşam döngüsünün tamamını kapsayan bütüncül bir yaklaşım benimsemek esastır. Bu, etik değerlendirmenin sadece nihai üründe değil, veri toplama, etiketleme, model seçimi, dağıtım ve izleme aşamalarının her birinde proaktif olarak yapılması gerektiği anlamına gelir. D'Ignazio ve Klein (2020)'ın veri feminizmi ilkeleri, gücü merkeze alan bu sorgulamanın nasıl yapılabileceğine dair kıymetli bir rehber sunar.

Nihayetinde, bu etik sorunların üstesinden gelmek, yalnızca teknik bir müdahaleyi değil, aynı zamanda hukuki, kurumsal, toplumsal ve felsefi bir dönüşümü gerektirir. İnsan onurunu, adaleti ve refahını merkeze alan bir YZ ekosistemi, disiplinlerarası iş birliği, şeffaf kamu diyalogu ve güçlü yasal çerçevelerle mümkün olabilecektir.

4. Sonuç ve Öneriler

Bu makale, dijital toplumda yapay zekâ (YZ) ve etik ilişkisini, sosyal adalet perspektifini merkeze alarak incelemiştir. Yapılan analizler, YZ teknolojilerinin toplumsal fayda ve ilerleme vaadini taşımakla birlikte, beraberinde önyargı, ayrımcılık, opaklık, hesap verebilirlik açığı ve mahremiyet erozyonu gibi derin ve iç içe geçmiş etik sorunlar getirdiğini ortaya koymuştur. Bu sorunlar, teknolojinin nötr bir araç olmadığını; aksine, mevcut toplumsal eşitsizlikleri, güç dinamiklerini ve ayrımcılık biçimlerini yeniden üreten ve ölçeklendiren bir yapıya sahip olduğunu göstermektedir (O'Neil, 2016: 21; Benjamin, 2019: 59). Bu nedenle, YZ'nin etik bir temelde geliştirilmesi ve kullanılması, teknik bir mesele olmaktan çok, sosyoteknik bir sistemler bütününe dönüştürmeyi gerektiren politik ve toplumsal bir projedir.

Çalışmanın ana bulguları şu şekilde özetlenebilir:

- **Etik, Tekniğin Önünde Gelmelidir:** Sorumlu ve güvenilir YZ, ancak insan onuru, özerklik, adalet, şeffaflık ve fayda sağlama gibi temel etik ilkelerin (Floridi ve Cowls, 2019: 4; Dignum, 2019:15) tasarım ve uygulama süreçlerine proaktif bir şekilde entegre edilmesiyle mümkündür. Etik düşünce, teknik geliştirmeyi takip eden bir son düşünce olmamalı, onu yönlendirmelidir.
- **Sosyal Adalet Merkezde Olmalıdır:** YZ sistemleri, tarihsel verilerde kodlanmış ve kurumsallaşmış ayrımcılıkları öğrenerek "Yeni Jim Kod"a (Benjamin, 2019) veya "Matematiksel Yıkım Silahları"na (O'Neil, 2016) dönüşebilmektedir. Adalet, eğitim, istihdam, kredi erişimi ve ceza hukuku

gibi alanlarda (Eubanks, 2018; Noble, 2018; Oğuz, 2024) bu sistemlerin en savunmasız gruplar üzerindeki orantısız etkisi mutlaka değerlendirilmeli ve hafifletilmelidir.

- **Sorunlar Birbirinden Bağımsız Değildir:** Opaklık ("kara kutu" problemi), hesap verebilirliği imkânsız kılar; hesap verebilirliğin olmaması ise ayrımcılığın denetlenememesine ve sürmesine yol açar (Wachter vd., 2017; Raji vd., 2020). Bu iç içe geçmiş yapı, sorunlara parçalı çözümlerin yetersiz kalacağını, bütüncül ve sistemik bir yaklaşım gerektirdiğini göstermektedir (Suresh ve Gutttag, 2021).
- **Bağlam ve Güç İlişkileri Kritik Öne Sahnedir:** Algoritmalar, içine yerleştirildikleri sosyal, ekonomik ve politik bağlamdan bağımsız değerlendirilemez. Gözetim kapitalizminin (Zuboff, 2019) ekonomik mantığı ve veri feminizminin (D'Ignazio & Klein, 2020) işaret ettiği güç eşitsizlikleri, YZ etiği tartışmalarının merkezine alınmalıdır.

Bu bulgular ışığında, dijital toplumda daha adil, şeffaf ve sorumlu bir YZ ekosistemi inşa etmek için aşağıdaki öneriler sunulabilir:

- **Önyargı Azaltma Tekniklerinin Benimsenmesi:** Makine öğrenimi yaşam döngüsünün (veri toplama, temizleme, model eğitimi, değerlendirme, dağıtım) her aşamasında önyargıyı tespit ve azaltmaya yönelik teknikler (Dwork vd., 2012; Chouldechova, 2017) standart uygulama haline getirilmelidir.
- **Şeffaflık ve Açıklanabilirliğin Artırılması:** "Anlamlı bilgi" sunmayı sağlayacak açıklanabilir YZ (Explainable AI - XAI) yöntemleri geliştirilmeli ve kullanılmalıdır (Doshi-Velez ve Kim, 2017). Kullanıcıların, kendileri hakkında verilen kararların mantığını anlayabilmeleri sağlanmalıdır.
- **Bağımsız Denetim ve Sertifikasyon:** YZ sistemleri, performans, adillik, güvenlik ve gizlilik açısından düzenli olarak bağımsız kuruluşlar tarafından denetlenmeli ve sertifikalandırılmalıdır (Raji vd., 2020). Bu, hesap verebilirlik açığını kapatmaya yönelik kritik bir adımdır.
- **Mevcut Mevzuatın Güçlendirilmesi:** Türkiye'de 6698 sayılı KVKK ve ilgili mevzuat, otomatik karar verme ve profil çıkarma gibi spesifik YZ uygulamaları karşısında daha net hükümlere kavuşturulmalıdır. Tanışık ve Bal'ın (2025) da vurguladığı gibi, kurumsal sorumluluklar açıkça tanımlanmalıdır.
- **AB Yapay Zekâ Yasası ile Uyum ve Ulusal Stratejinin Güncellenmesi:** Türkiye'nin, Avrupa Birliği'nin kabul ettiği Yapay Zekâ Yasası (2024) gibi kapsamlı düzenlemeleri yakından takip etmesi ve Ulusal Yapay Zekâ Stratejisi'ni (Cumhurbaşkanlığı Dijital Dönüşüm Ofisi, 2021) bu doğrultuda, sosyal adalet ve algoritmik denetim vurgusunu güçlendirerek güncellemesi hayati önem taşımaktadır.

- **Algoritmik Ayrımcılıkla Mücadele:** Oğuz'un (2024) önerisi doğrultusunda, İş Kanunu ve ilgili mevzuatta, işe alım ve çalışma hayatında algoritmik ayrımcılığı açıkça yasaklayan ve caydırıcı yaptırımlar öngören düzenlemeler yapılmalıdır.
- **Etik Kurullar ve Etik İlkelerin İçselleştirilmesi:** YZ geliştiren ve kullanan tüm kurum ve kuruluşlarda, disiplinlerarası üyelerden oluşan etik kurullar kurulmalı ve bu kurullar karar mekanizmalarında etkin bir role sahip olmalıdır. IEEE (2019) ve Avrupa Komisyonu (2019) tarafından önerilen etik ilkeler somut proje değerlendirme kriterlerine dönüştürülmelidir.
- **Etik Okuryazarlık ve Kamusal Diyalogun Güçlendirilmesi:** Toplumun her kesiminde YZ etiği okuryazarlığını geliştirmeye yönelik eğitim programları desteklenmeli, konuya ilişkin şeffaf ve kapsayıcı bir kamusal diyalog teşvik edilmelidir. Vatandaşların, kendilerini etkileyen teknolojilerin geliştirilme süreçlerine katılımı için kanallar açılmalıdır.
- **Sivil Toplum ve Akademinin Rolünün Desteklenmesi:** Türkiye Yapay Zekâ İnisiyatifi (TRAI, 2024) gibi sivil toplum kuruluşları ile Yaşar Ümütlü (2025), Akıncılar Köseoğlu ve Çetin (2024) gibi akademisyenlerin bu alandaki çalışmaları desteklenmeli, bu aktörlerin politika yapımcılar ve endüstri ile iş birliği yapması teşvik edilmelidir.

Yapay zekâ, toplumumuzun karşı karşıya olduğu en önemli dönüşümlerden birinin merkezinde yer almaktadır. Bu dönüşümün yönü, teknolojik kaçınılmazlıklar tarafından değil, yapacağımız kolektif tercihler ve kuracağımız etik, hukuki ve toplumsal çerçeveler tarafından belirlenecektir. İnsanı, onurunu ve refahını merkeze alan; çeşitliliği, kapsayıcılığı ve adaleti gözeten; şeffaf ve hesap verebilir bir YZ geleceği inşa etmek, hepimizin sorumluluğudur. Bu, mühendislerin, hukukçuların, sosyal bilimcilerin, politikacıların ve nihayetinde tüm vatandaşların katılımıyla ilerleyecek, zorlu ama zorunlu bir yolculuktur. Bu yolculukta, eleştirel düşünce, etik refleks ve sosyal adalet arayışı, en az teknik yenilik kadar değerli rehberlerimiz olacaktır.

KAYNAKÇA

- Akıncılar Köseoğlu, N., & Çetin, B. (2024). Avrupa'da Dijital Etik, İnsan Hakları Bağlamında Yapay Zekâ ve Algoritmik Ayrımcılık. *Kastamonu Üniversitesi İktisadi ve İdari Bilimler Fakültesi Dergisi*, 26(1), 69-83. <https://doi.org/10.21180/iibfdkastamonu.1384167>
- Argın, Y. (2023). Türk Sinemasında Kadın ve Akıl Hastalığı Temsilleri: Beyza'nın Kadınları Filminin Eleştirel Söylem Çözümlemesi. Y. Arğın (Ed.), *Güncel Yaklaşımlarla Geleneksel ve Yeni Medyada Beden* içinde (s. 214). Nobel Bilimsel.
- Barocas, S., & Selbst, A. D. (2016). *Big data's disparate impact*. *California Law Review*, 104(3), 671-732. <https://doi.org/10.2139/ssrn.2477899>
- Benjamin, R. (2019). *Race after technology: Abolitionist tools for the New Jim Code*. Polity Press.

- Biçer, S., & Şener, Y. (2020). Castells'in İzinden Mısır ve Hong Kong Protestoları Örneğiyle Dijital Aktivizm ve Yeni Toplumsal Hareketler. A. S. İgit & Ö. Sayılğan (Eds.), *Dijital İletişim - Kuram ve Araştırmaları* içinde (s. 320). Nobel Akademik Yayıncılık.
- Chouldechova, A. (2017). Fair prediction with disparate impact: A study of bias in recidivism prediction instruments. *Big Data*, 5(2), 153–163. <https://doi.org/10.1089/big.2016.0047>
- Crawford, K. (2021). *Atlas of AI: Power, politics, and the planetary costs of artificial intelligence*. Yale University Press.
- Cumhurbaşkanlığı Dijital Dönüşüm Ofisi. (2021). *Ulusal Yapay Zekâ Stratejisi (2021–2025)*. <https://cbddo.gov.tr>
- Demirel, Y. T., & Arıkan, N. İ. (2023). Yapay zekânın afet bölgelerinde kullanımı. *International Journal of Educational and Social Sciences*, 2(2), 77–82.
- D'Ignazio, C., & Klein, L. F. (2020). *Data feminism*. MIT Press.
- Dignum, V. (2019). *Responsible artificial intelligence: How to develop and use AI in a responsible way*. Springer. <https://doi.org/10.1007/978-3-030-30371-6>
- Doshi-Velez, F., & Kim, B. (2017). Towards a rigorous science of interpretable machine learning. *arXiv*. <https://arxiv.org/abs/1702.08608>
- Dwork, C., Hardt, M., Pitassi, T., Reingold, O., & Zemel, R. (2012). Fairness through awareness. In *Proceedings of the 3rd Innovations in Theoretical Computer Science Conference* (pp. 214–226). ACM. <https://doi.org/10.1145/2090236.2090255>
- Eubanks, V. (2018). *Automating inequality: How high-tech tools profile, police, and punish the poor*. St. Martin's Press.
- European Commission High-Level Expert Group on Artificial Intelligence. (2019). *Ethics guidelines for trustworthy AI*. European Commission. <https://digital-strategy.ec.europa.eu>
- European Union. (2016). Regulation (EU) 2016/679 (General Data Protection Regulation). *Official Journal of the European Union*, L 119, 1–88.
- European Union. (2024). Regulation (EU) 2024/1689 (Artificial Intelligence Act). *Official Journal of the European Union*.
- Floridi, L., & Cows, J. (2019). A unified framework of five principles for AI in society. *Harvard Data Science Review*, 1(1). <https://doi.org/10.1162/99608f92.8cd550d1>
- IEEE. (2019). *Ethically aligned design: A vision for prioritizing human well-being with autonomous and intelligent systems* (1st ed.). IEEE Standards Association.
- Kleinberg, J., Mullainathan, S., & Raghavan, M. (2017). Inherent trade-offs in the fair determination of risk scores. In *Proceedings of the 8th Innovations in Theoretical Computer Science Conference* (pp. 43:1–43:23). Schloss Dagstuhl. <https://doi.org/10.4230/LIPIcs.ITCS.2017.43>
- Mittelstadt, B. D., Allo, P., Taddeo, M., Wachter, S., & Floridi, L. (2016). The ethics of algorithms: Mapping the debate. *Big Data & Society*, 3(2), 1–21. <https://doi.org/10.1177/2053951716679679>
- Noble, S. U. (2018). *Algorithms of oppression: How search engines reinforce racism*. NYU Press.
- O'Neil, C. (2016). *Weapons of math destruction: How big data increases inequality and threatens democracy*. Crown.
- OECD. (2019). *Recommendation of the Council on Artificial Intelligence*. OECD. <https://legalinstruments.oecd.org>
- Oğuz, Ö. (2024). Çalışma Hayatında Algoritmik Ayrımcılık. Süleyman Demirel Üniversitesi Hukuk Fakültesi Dergisi, 14(2), 1851–1886. <https://doi.org/10.52273/sduhfd..1581436>
- Raji, I. D., Smart, A., White, R., Mitchell, M., Gebru, T., Hutchinson, B., Smith-Loud, J., Theron, D., & Barnes, P. (2020). Closing the AI accountability gap: Defining auditing and accounting for AI systems. In *Proceedings of the 2020*

- Conference on Fairness, Accountability, and Transparency (FAccT '20)* (pp. 33–44). ACM. <https://dl.acm.org/doi/10.1145/3351095.3372873>
- Selbst, A. D., & Powles, J. (2017). Meaningful information and the right to explanation. *International Data Privacy Law*, 7(4), 233–242. <https://doi.org/10.1093/idpl/ix022>
- Selbst, A. D., Boyd, D., Friedler, S. A., Venkatasubramanian, S., & Vertesi, J. (2019). Fairness and abstraction in sociotechnical systems. In *Proceedings of the 2019 Conference on Fairness, Accountability, and Transparency (FAccT '19)* (pp. 59–68). ACM. <https://doi.org/10.1145/3287560.3287598>
- Suresh, H., & Gutttag, J. (2021). A framework for understanding sources of harm throughout the machine learning life cycle. In *Proceedings of the 2021 ACM Conference on Fairness, Accountability, and Transparency (FAccT '21)* (pp. 113–123). ACM. <https://doi.org/10.1145/3465416.3483305>
- Şener, Y. (2021). Sosyal Medya ve İdeal Bedenin İnşası: Diyet İçerikli Instagram Sayfalarında İdealize Edilmiş Kadın Bedenleri Üzerine Göstergebilimsel Bir Analiz. A. Karabulut (Ed.), *Dijital Yozlaşma ve Etik* içinde (s. 303). LİTARATÜRK.
- T.C. Resmî Gazete. (2016, 7 Nisan). 6698 sayılı Kişisel Verilerin Korunması Kanunu (Sayı: 29677). <https://www.mevzuat.gov.tr>
- Tanişik, S., & Bal, S. (2024). Dijital Mahremiyet ve Kurumsal Sorumluluk: Kişisel Verilerin Korunmasında İletişim Teknolojilerinin Kamusal Rolü. *Yeni Medya*(16), 268–285. <https://doi.org/10.55609/yenimedya.1424182>
- Türkiye Yapay Zekâ İnisiyatifi (TRAI). (2024). *Yapay zekâ etik ilkeleri ve hukuki düzenlemeler raporu*. <https://turkiye.ai/wp-content/uploads/2024/06/TRAI-Yapay-Zeka-Etik-Ilkeleri-ve-Hukuki-Duzenlemeler-Raporu-Mayis-2024-5.pdf>
- UNESCO. (2021). *Recommendation on the ethics of artificial intelligence*. UNESCO. <https://unesdoc.unesco.org>
- Wachter, S., Mittelstadt, B., & Floridi, L. (2017). Why a right to explanation of automated decision-making does not exist in the GDPR. *International Data Privacy Law*, 7(2), 76–99. <https://doi.org/10.1093/idpl/ix005>
- Yaşar Ümütlü, A. (2025). Algoritmik Adalet: Uluslararası Hukukta Yapay Zeka Hakimliği. Selçuk Üniversitesi Hukuk Fakültesi Dergisi, 33(1), 777–815. <https://doi.org/10.15337/suhfd.1637446>
- Zuboff, S. (2019). *The age of surveillance capitalism: The fight for a human future at the new frontier of power*. PublicAffairs.

GENİŞLETİLMİŞ ÖZET

Bu makale, dijital toplumda hızla yaygınlaşan yapay zekâ (YZ) teknolojilerinin doğurduğu etik sorunları sosyal adalet perspektifinden incelemektedir. YZ'nin kamu hizmetlerinden istihdama, ceza adaletinden sağlık sistemlerine kadar pek çok alanda karar süreçlerini otomatikleştirmesi hem yeni fırsatlar hem de önemli riskler yaratmaktadır. Çalışma, bu risklerin özellikle dezavantajlı toplumsal gruplar üzerinde nasıl eşitsiz sonuçlar doğurduğunu tartışmakta ve algoritmik önyargı, ayrımcılık, şeffaflık eksikliği, hesap verebilirlik açığı ve mahremiyet ihlalleri gibi temel etik sorunları derinlemesine ele almaktadır.

Makalenin ilk bölümünde, sorumlu ve güvenilir YZ kavramları kuramsal bir çerçevede tartışılmakta; insan onuru, yararlılık, zarar vermeme, adalet, özerklik ve açıklanabilirlik ilkelerinin YZ sistemlerinin tasarım ve kullanımında neden kritik olduğu açıklanmaktadır. Avrupa Birliği Yapay Zekâ Yasası, OECD ve UNESCO etik çerçeveleri gibi uluslararası düzenlemeler analiz edilerek, etik yönetim modellerinin dijital toplumun sürdürülebilirliği için önemi vurgulanmaktadır.

İkinci bölüm, sosyal adalet odağında algoritmik ayrımcılığı irdelemekte, yapısal eşitsizlikleri veri setleri üzerinden yeniden üreten YZ sistemlerinin nasıl "Yeni Jim Kod" (Benjamin, 2019) ya da "Matematiksel Yıkım Silahları" (O'Neil, 2016) haline gelebildiğini örneklerle göstermektedir. Ceza adaletinde risk tahmin algoritmalarının ırksal yanlılık üretmesi, işe alım algoritmalarının kadınlar ve engelliler aleyhine

sonular retmesi ve sosyal yardım sistemlerinde otomatik karar verme srelerinin yoksulları cezalandırabilmesi, ayrıntılı şekilde tartıřılan rnekler arasındadır.

nc blmde opaklık, hesap verebilirlik aıėı, mahremiyet erozyonu ve gzetim kapitalizmi gibi gncel etik sorunlar ele alınmaktadır. zellikle veri temelli iktidar iliřkileri, byk teknoloji řirketlerinin toplumsal eřitsizlikleri pekiřtiren ekonomik modelleri ve bireylerin dijital zerkliėinin ařınması zerinde durulmaktadır.

Sonu olarak makale, adil ve etik YZ kullanımının teknik czmlerle sınırlı olamayacaėını, sosyoteknik bir dnřm gerektirdiėini savunmaktadır. neriler arasında aıklanabilir YZ yntemleri, baėımsız algoritma denetimleri, Trkiye'nin KVKK ile uyumlu ama daha kapsayıcı dzenlemeler geliřtirmesi, AB Yapay Zek Yasası ile uyum saėlanması, etik kurulların kurumsal yapılara entegre edilmesi ve toplumda YZ okuryazarlıėının artırılması yer almaktadır. alıřma, dijital toplumun ancak insan onurunu, toplumsal adaleti ve hesap verebilirliėi nceleyen yntemlerle srdrlebilir olabileceėi sonucuna varmaktadır.

EXTENDED ABSTRACT

This article examines the ethical challenges generated by the rapid integration of artificial intelligence (AI) technologies into the digital society from a social justice perspective. As AI increasingly shapes decision-making in public services, employment, criminal justice, and healthcare, it produces both significant opportunities and substantial risks. The study focuses on how these risks disproportionately affect marginalized groups and analyzes major ethical concerns such as algorithmic bias, discrimination, opacity, accountability gaps, and privacy erosion.

The first section provides the theoretical foundation for responsible and trustworthy AI, emphasizing the relevance of ethical principles such as human dignity, beneficence, non-maleficence, justice, autonomy, and explicability. International governance frameworks—such as the EU Artificial Intelligence Act, OECD AI Principles, and UNESCO Ethics of AI Recommendation—are reviewed to highlight the importance of robust regulatory structures for the digital society.

The second section explores algorithmic discrimination within the broader context of social justice, demonstrating how AI systems can reproduce and amplify structural inequalities embedded in historical data. Concepts like the “New Jim Code” (Benjamin, 2019) and “Weapons of Math Destruction” (O’Neil, 2016) are used to illustrate how algorithmic processes may institutionalize racial, gender-based, or socio-economic disparities. Real-world cases—including biased recidivism prediction tools, discriminatory hiring algorithms, and automated welfare systems that penalize the poor—are examined in detail.

The third section discusses systemic ethical issues such as opacity (“black box” models), accountability gaps, surveillance capitalism, and the erosion of individual autonomy in the digital environment. The article argues that these problems are interconnected: opacity undermines accountability, insufficient accountability allows discriminatory outcomes to persist, and data-driven economic structures reinforce existing power imbalances.

In conclusion, the study asserts that achieving ethical and socially just AI cannot be handled solely through technical interventions but requires a comprehensive sociotechnical transformation. The recommendations include the adoption of explainable AI practices, independent auditing mechanisms, more inclusive legal and regulatory frameworks aligned with the EU AI Act, institutional AI ethics committees, and the promotion of AI ethics literacy among the public. The article ultimately argues that a sustainable digital society depends on governance models that prioritize human dignity, fairness, transparency, and collective accountability.