



Araştırma Makalesi / Research Article

Dijital Etik: Yapay Zekâ ve Akıllı Sistem İhlalleri Üzerine Değer Temelli Bir Analiz*

Digital Ethics: A Value-Based Analysis of Artificial Intelligence and Smart System Violations

Alaattin ASLAN¹ 

¹ Dr. Öğretim Üyesi, Marmara Üniversitesi, İletişim Fakültesi, Gazetecilik Bölümü,

alaattin.aslan@marmara.edu.tr

Özet

Bu çalışma, yapay zekâ ve akıllı sistemlerin neden olduğu etik krizleri, dijital etiğin kavramsal çerçevesi içinde çözümlemek amacıyla özgün bir “Değer Temelli Dijital Etik Model” sunmaktadır. Model, Victor Turner’ın “Sosyal Drama” kuramı ile Floridi ve Cows’ın yapay zekâ etiği ilkelerini sentezlemekte; literatürdeki üç temel kaynaktan (Hoven, IEEE Ethically Aligned Design ve Ashok ve ark. sistematik taraması) hareketle belirlenen beş ana dijital değer (faydalılık, zararsızlık, adalet, özerklik, açıklık) ve bu değerlere bağlı yirmi dört alt normdan oluşmaktadır. Model, her bir dijital değeri Turner’ın ihlal, kriz, düzeltici eylem ve yeniden bütünleşme aşamalarında normatif birer kriter olarak konumlandırmakta; böylece algoritmik ihlallerin hangi aşamada hangi değerin ihlal edildiğinin tespiti ve sistemin bu değere uygun biçimde yeniden yapılandırılması yoluyla toplumsal yeniden bütünleşmenin sağlanabileceğini öngörmektedir. Nitel vaka analizi yöntemiyle, Dünya Ekonomik Forumu’nun belirlediği sorun alanları ekseninde seçilen sekiz küresel yapay zekâ krizi (algoritmik önyargı, deepfake manipülasyonu, otonom araç kazası, veri sömürüsü, otonom silahlar vb.) incelenmiştir. Bulgular, zararsızlık ve adalet değerlerinin vakaların tamamında ihlal edildiğini; etik krizlerin büyük ölçüde tasarım aşamasındaki eksikliklerden kaynaklandığını (sekiz vakanın beşi) ve mevcut hukuki mekanizmaların reaktif kalarak krizleri önlemede yetersiz kaldığını ortaya koymuştur. Çalışma, beş temel dijital değere dayalı proaktif bir etik yönetim çerçevesinin, dijital toplumun demokratik bütünlüğü ve insan onurunun korunması açısından varoluşsal bir zorunluluk olduğunu savunmaktadır.

Anahtar Kelimeler: Yapay zekâ etiği, dijital etik, sosyal drama, algoritmik ihlaller, değer temelli dijital etik modeli.

Abstract

This study presents an original “Value-Based Digital Ethics Model” to analyze the ethical crises caused by artificial intelligence and smart systems within the conceptual framework of digital ethics. The model synthesizes Victor Turner’s “Social Drama” theory with the AI ethics principles of Floridi and Cows, and consists of five core digital values (beneficence, non-maleficence, justice, autonomy, explicability) and twenty-four sub-norms derived from three main sources in the literature: Hoven, IEEE’s Ethically Aligned Design, and the systematic review by Ashok et al. The model positions each digital value as a normative criterion within Turner’s four stages of breach, crisis, redressive action, and reintegration, thereby proposing that algorithmic violations can be resolved by identifying which value has been breached at which stage and restructuring the system accordingly to achieve social reintegration. Using a qualitative case study method, eight global AI crises (algorithmic bias, deepfake manipulation, autonomous vehicle accident, data exploitation, autonomous weapons, etc.) selected within the framework of problem areas identified by the World Economic Forum were examined. The findings revealed that the values of non-maleficence and justice were violated in all eight cases; that ethical crises largely stem from deficiencies in the design phase (five out of eight cases); and that existing legal mechanisms remain reactive and insufficient in preventing crises. The study argues that a proactive ethical governance framework based on the five core digital values is an existential necessity for preserving the democratic integrity of digital society and human dignity.

Keywords: Artificial intelligence ethics, digital ethics, social drama, algorithmic violations, value-based digital ethics model.

* Bu çalışma, Marmara Üniversitesi Sosyal Bilimler Enstitüsü bünyesinde Doç. Dr. Şevki IŞIKLI danışmanlığında 2022 yılında başarıyla tamamlanan “Dijital Etik: Değer Temelli Bir Model Önerisi” başlıklı doktora tezinden üretilmiştir.

Giriş

Toplumsal dönüşümlerin itici gücünü oluşturan teknolojik gelişmeler, gündelik hayatı, insan ilişkilerini ve yerleşik yaşam kalıplarını köklü biçimde dönüştürmektedir. Son otuz yılda bilgisayar ve internetin gündelik yaşamın ayrılmaz bir parçası hâline gelmesiyle birlikte, topluma dair her türlü analizin başlangıç noktası, teknoloji ve iletişim alanında yaşanan devrimsel nitelikteki yenilikler olmuştur. Bilgisayarların veriyi işleyebildiği, depolayabildiği ve internet aracılığıyla iletebildiği “0 ve 1’lerin dünyası”, dijitalleşmenin öncü aşamalarını ve toplumun giderek dijitalleşen doğasını tarif etmektedir.

Söz konusu tarihsel süreçte toplumsal dönüşümü kavramak amacıyla sosyoloji literatüründe önemli kavramsallaştırmalar geliştirilmiştir. Yoneji Masuda (1980), bilgisayar teknolojilerinin insanın zihinsel emeğini hem dönüştüreceğini hem de güçlendireceğini öne sürerek bu yeni toplumsal yapıyı “Enformasyon Toplumu” olarak adlandırmıştır. Marshall McLuhan (1964) ise “Küresel Köy” kavramıyla elektrik teknolojisinin, toplumsal bağımlılık biçimlerimizi ve özel hayatımızın her yönünü yeniden şekillendirerek insanlığı, kanıksanmış tüm fikir ve gelenekleri tekrar değerlendirmeye zorladığını ileri sürmüştür. Daha yakın dönemde Manuel Castells (2000), enformasyon teknolojilerindeki gelişmenin bir çarpan etkisi yaratarak kapitalist ekonomik düzeni ağ mantığıyla küresel bir hegemonyaya dönüştürdüğünü belirtmiş ve bu yapıyı “Ağ Toplumu” olarak kavramsallaştırmıştır.

Günümüzde klasik ağ toplumu anlayışının ötesine geçilerek dijital şekillenmede “Toplum 5.0” (Süper Akıllı Toplum) olarak tanımlanan yeni bir aşamaya ulaşılmıştır (Cabinet Office, Government of Japan, 2016; Harayama, 2017). İnsanlık hâlihazırda akıllı telefonlar, sosyal medya, otonom araçlar, yapay zekâ ve büyük veri gibi kendisini doğrudan etkileyen teknolojiler ile yakın gelecekte robotik, nesnelerin interneti ve teknolojik tekillik gibi derin etkiler yaratması beklenen olguları anlamaya çalışmaktadır. Toplum 5.0 vizyonunda yapay zekâ teknolojisi, gerçek dünya modellerini siber uzayla birleştirerek toplumsal kararlar alma potansiyeline erişmiş ve etkin insan denetiminin sınırlarını tartışmaya açmıştır (Deguchi vd., 2020).

Teknolojinin hayatın her anına, hatta insan bedeninin ve zihninin içine nüfuz ettiği bu dönemde, yapay zekâ ve akıllı sistemlerin yol açtığı yapısal ve ahlaki sorunlar, küresel çaptaki güncel araştırma raporlarına da açıkça yansımaktadır. Robot teknolojilerinin ve yapay zekânın insan hayatına etkilerini inceleyen PwC (2018) raporu, akıllı sistemlerin gelecekte insanlara özgü görevleri yerine getirmesi durumunda insanın sistemin neresinde konumlanacağını sorgulamakta; iş birliği, liderlik ve duygusal zekâ gibi insani değerlerin ayırt edici niteliğini vurgulamaktadır. Dünya Ekonomik Forumu (WEF) tarafından hazırlanan bir raporda (Eggers ve Hickin, 2020) ise dijital teknolojilerin kullanımıyla ortaya çıkan krizler irdelenmiş; sınırlı düzenleme kapasitesi, mahremiyet ihlalleri, otonom sistemlerde hesap verebilirlik problemi, veri sömürüsü ve teknolojinin kötü aktörler tarafından kötüye kullanım potansiyeli, insanlığın yüzleştiği başlıca teknolojik meydan okumalar olarak belgelenmiştir.

Dijital toplumda herkes tarafından benimsenmiş ortak yaşam kurallarının henüz belirginlik kazanmadığı böyle bir dönemde, teknoloji üreticilerinin toplumun dijitalleşmeyi anlama ve benimseme hızının çok daha ilerisinde olduğu görülmektedir. Dijital teknolojilerin gelişim ivmesine yetişemeyen hukuki düzenlemelerin, toplumsal sorunlara anlık çözümler üretmesi mümkün değildir; sürekli yeni yasal düzenleme yapma arayışı ise hukukun yapısal hantallığı karşısında etik ve ahlaki normların oluşumunu zayıflatmaktadır. Oysa toplumsal hayat nasıl hem hukuk kurallarına hem de ahlaki normlara dayanıyorsa, yapay zekâ sistemleri ve algoritmik yapılar da etik davranış kodlarına ve ahlaki eylem pratiklerini yansıtan yeni dijital değerlere ihtiyaç duymaktadır (Floridi ve Cows, 2019; UNESCO, 2021). Nitekim Avrupa Birliği'nin 2024 yılında yürürlüğe giren Yapay Zekâ Yasası (EU AI Act), risk temelli bir sınıflandırma ile güvenilir yapay zekâyı hukuki zeminde tanımlamaya çalışırken; UNESCO'nun 2021 tarihli Yapay Zekâ Etiği Tavsiyesi ise evrensel değerler ve ilkeler çerçevesinde küresel bir etik mutabakat inşa etmeyi hedeflemektedir.

Bu makale, dijitalleşmenin ulaştığı en ileri aşama olan yapay zekâ ve akıllı sistemlerin neden olduğu etik krizleri çözümleyebilmek amacıyla, önde gelen antropologlardan Victor Turner'ın "Sosyal Drama" kuramından (1974) hareketle özgün bir "Değer Temelli Dijital Etik Model" ortaya koymayı amaçlamaktadır. Turner (1974, s. 37-41) her ölçekteki toplumsal yapıda ortaya çıkan çatışmaların ve krizlerin; *ihlal* (breach), *kriz* (crisis), *düzeltilici eylem* (redressive action) ve *yeniden*

bütünleşme ya da ayrışma (reintegration/schism) şeklinde sıralanan dört kamusal eylem aşaması aracılığıyla işlendiğini ve yeni normlara dönüştüğünü savunmaktadır. Bu çalışmada Turner'ın kuramı, dijital dünyanın algoritmik krizlerine uyarlanarak, yapay zekâ kaynaklı ihlallere karşı bir “düzeltici eylem” işlevi görecektir beş temel dijital değer sunulmaktadır: faydalılık, zararsızlık, adalet, özerklik ve açıklık. Söz konusu değerler, Floridi ve Cowls'un (2019) yapay zekâ etiğine ilişkin birleştirilmiş ilkeler çerçevesi ile örtüşmekte; ancak bu çalışmada teorik bir soyutlama olarak değil, somut vaka örnekleri üzerinden test edilebilir bir analitik araç olarak konumlandırılmaktadır.

Araştırmanın temel sorusu şudur: *Victor Turner'ın Sosyal Drama modeli bağlamında, yapay zekâ ve akıllı sistem ihlallerinden doğan etik krizler nasıl çözümlenebilir ve hangi dijital değerler bu krizlere karşı düzeltici eylem olarak işlev görebilir?* Bu soruya yanıt aramak üzere çalışmada, yapay zekâ ve akıllı sistem ihlallerine odaklanan gerçek vaka örnekleri (otonom araç kazaları, algoritmik ayrımcılık, otonom dron kararları vb.) Turner'ın dört aşamalı Sosyal Drama modeli çerçevesinde çözümlenmekte ve her bir vaka için önerilen beş dijital değer düzeltici işlevi değerlendirilmektedir. Böylece makale, teknoloji tasarımında ve yapay zekâ yönetiminde evrensel bir dijital etiğin inşasının gerekliliğini hem kuramsal hem de ampirik düzlemde ortaya koymayı hedeflemektedir.

1. Dijital Etik ve Değer Temelli Model

Geleneksel ahlak teorileri -teleolojik, deontolojik ya da erdem etiği temelli yaklaşımlar- yüzlerce yıl boyunca insan eylemlerinin ahlaki sınırlarını çizmede rehberlik etmişse de büyük veri ile beslenen, derin öğrenme yoluyla kendi kendine öğrenebilen ve otonom kararlar alabilen makinelerin yarattığı eşzamanlı ve karmaşık krizleri çözmede tek başlarına yetersiz kalmaktadır. Moor'un (2006) da vurguladığı gibi, geleneksel etik teoriler yeni teknolojik durumlara uyarlandığında kavramsal boşluklar ortaya çıkmakta ve bu boşluklar yeni etik çerçevelerin geliştirilmesini zorunlu kılmaktadır. Felsefi teorilerin soyut normlarının pratik ve çağdaş sorunlara uyarlanmasını ifade eden uygulamalı etik disiplini içerisinde, bilgisayar etiğinin sınırlarını da aşan “Yapay Zekâ ve Makine Etiği” yepyeni bir

zorunlu alan olarak doğmuştur. Anderson ve Anderson (2007, s. 10), makine etiğini, bilgisayar etiği alanını genişleterek makinelerin otonom olarak ne yaptığıyla ve insanlara karşı nasıl davranması gerektiğiyle ilgilenen bir disiplin olarak tanımlamaktadır. Makine etiği; yalnızca teknolojiyi kullanan insan öznenin ahlaki sorumluluğunu değil, bizzat karar alıcı konumuna yükselen akıllı sistemin karar mekanizmasına gömülü etik kodları sorgulamaktadır (Dignum, 2019; Wallach ve Allen, 2009). Bu durum ise yeni bir etik düşüncenin ortaya koyulmasını gerekli kılmaktadır.

1.1. Dijital Etiğin Tanımı ve Kapsamı

Bu yeni etik alanının kavramsal çerçevesini çizebilmek için öncelikle dijital etik kavramının net bir tanımına ihtiyaç vardır. Literatürde dijital etik; sosyal, politik ve ahlaki alan üzerindeki teknolojik etkinin incelenmesi (Marghalani ve AlQahtani, 2019, s. 3), teknolojik yeniliklerle toplumu olumlu yönde şekillendirmek için dijital bir dünyada hangi norm ve değerlere ulaşmak istediğimiz sorusuyla ilgilenme (Mackert, 2020, s. 4) ve enformasyon ile iletişim teknolojilerinin toplumlar ve genel olarak yaşam alanımız üzerindeki etkisini sorgulama (Capurro, 2009) olarak tanımlanmaktadır. Bu tanımlardan hareketle dijital etik, en geniş anlamıyla; teknolojik gelişmenin hızına karşılık nasıl bir ahlaki davranış izleyeceğimizi, bu davranışın hangi değerlere dayanacağını ve bu değerlerin nasıl birer norm hâline dönüşeceğini inceleyen uygulamalı etik disiplindir. Hoven'ın (2015, s. 59) belirttiği gibi dijital etik, karmaşık adaptif sistemler ve teknolojinin toplumu etkilediği bir dünyayla ilgilenirken; doğası gereği epistemik olarak erişilemez olanı bilmemekten sorumlu tutulamayacağımızı da bir tevazuyla kabul etmeliyiz. Çünkü özellikle yapay zekâ ve akıllı sistemler için sınırlar ve kabuller nelerdir epistemik açıdan hala muallaktır.

Dijital etiğin kapsamı ise birbiriyle kesişen üç temel boyutta ele alınabilir (Barker ve Ferguson, 2022):

- i. Tasarım: Dijital teknolojilerin ve veri araçlarının geliştirilme aşamasında mühendislerin, yazılımcıların ve veri bilimcilerin etik ilkeleri koda ve mimariye entegre etmesini ifade eder.
- ii. Kullanım: Hizmet kullanıcıları, çalışanlar, yöneticiler ve ortakların gelişen teknoloji ve verileri nasıl kullandığını etik açıdan değerlendirir.

- iii. Toplumsal Etki: Dijital teknolojinin ve veri analitiğinin daha geniş toplum, insan hakları, çevresel sürdürülebilirlik ve toplumsal kabul edilebilirlik üzerindeki etkilerini inceler. Bu çalışma, yapay zekâ ve akıllı sistem ihlallerini bu üç boyutun tamamında değerlendirerek bütüncül bir etik analiz sunmayı hedeflemektedir.

1.2. Değerlerin Dijitalleşmesi ve Alt Normların Belirlenmesi

Dijital etiğin soyut bir felsefi tartışma olmaktan çıkıp uygulanabilir bir çerçeveye dönüşebilmesi için, temel insani değerlerin dijital bağlama nasıl uyarlandığının ve hangi yeni değerlerin ortaya çıktığının tespit edilmesi gerekmektedir. Moor (1998, s. 19-20), yaşam, sağlık, mutluluk, güvenlik, kaynaklar ve bilgi gibi temel insani değerlerin bilgisayar etiği sorunlarını çözümlemek için kullanılabileceğini öne sürmüştür. Hoven (2015, s. 70) ise değerlerin teknolojiye gömülebileceğini, teknoloji aracılığıyla gelecekteki kullanıcıların eylem alanını şekillendirebileceğini ve bu nedenle tasarım aşamasında ortak değerleri teknolojiye eklemeyi öğrenmemiz gerektiğini savunmaktadır.

Bu çalışmada, dijital değerlerin ve bunlara bağlı alt normların belirlenmesi için literatürdeki üç temel kaynaktan yararlanılarak sentezci bir yaklaşım benimsenmiştir. İlk kaynak, Hoven'ın (2015) dijital etik için önerdiği değerler şemasıdır. İkinci kaynak, IEEE'nin "Ethically Aligned Design" (Creative Commons, 2019) raporunda otonom ve akıllı sistemler için sıralanan değerlerdir. Üçüncü kaynak ise Ashok ve arkadaşlarının (2021) dijital etik, tekno-etik, bilgisayar etiği, enformasyon etiği, robo-etik ve makine etiği literatürünü kapsayan sistematik taramasıdır. Bu üç kaynaktaki ortak kavramların karşılaştırmalı analizi sonucunda, dijital etik için beş ana değer ve yirmi dört alt normdan oluşan bir çerçeve elde edilmiştir. Elde edilen bu beş ana değer; Floridi ve Cowls'ın (2019) yapay zekâ etiğine ilişkin birleştirilmiş çerçevesi ile de tam bir uyum göstermektedir. Floridi ve Cowls (2019), biyoetik ilkelerine (Beauchamp ve Childress, 2013) açıklanabilirlik ilkesini ekleyerek yapay zekâ bağlamında beş temel ilke önermiştir. Bu çalışmada söz konusu beş ilke, Turner'ın düzeltici eylem aşaması için operasyonel hâle getirilerek beş temel "Dijital Değer" ve bunlara karşılık gelen yirmi dört alt norm olarak tanımlanmıştır.

1.3. Victor Turner’ın Sosyal Drama Kuramı ve Modele Uyarlanması

Yapay zekânın neden olduğu ahlaki krizlerin nasıl aşılabileceğini sistematik bir yapıya kavuşturmak amacıyla bu çalışmada, antropolog Victor Turner’ın (1974) fiziksel toplulukların çatışma deneyimlerini analiz etmek için geliştirdiği “Sosyal Drama” kuramı dijital teknolojiye ve algoritmik krizlere uyarlanmıştır. Turner’ın modelinin seçilmesinin temel gerekçesi, söz konusu kuramın krizleri yalnızca betimlemekle kalmayıp, aynı zamanda normatif bir çözüm aşaması (*düzeltilici eylem*) öngörmesi ve bu aşamanın yeni toplumsal normların inşasıyla sonuçlanmasıdır. Turner (1974, s. 33), toplumların yaşadığı çatışma durumlarının ardından yeni toplumsallığın nasıl sağlandığına yönelik analizinde, bir sosyal dramının ilerlemesi boyunca dört kamusal eylem aşamasının gerçekleştiğini belirtmektedir: *ihlal* (breach), *kriz* (crisis), *düzeltilici eylem* (redressive action) ve *yeniden bütünleşme ya da ayrışma* (reintegration/schism) (Turner, 1974, s. 38-39). Bu bağlamda, bir yapay zekâ algoritmasının önyargılı karar vermesi bir “ihlal”, bu durumun kitleleri mağdur edip toplumsal güveni sarsması “kriz”, sistemin etik ilkelere göre yeniden tasarlanması “düzeltilici eylem” ve yeni hukuki ya da ahlaki standartların kurumsallaşması ise “yeniden bütünleşme” aşamasına karşılık gelmektedir.

Şekil 1. Değer Temelli Dijital Etik Model



Turner’ın modelindeki en kritik aşama olan düzeltilici eylemin ahlaki sınırlarını çizebilmek için evrensel değerlere ihtiyaç vardır. İşte tam da bu noktada, yukarıda literatür sentezi ile elde edilen beş temel dijital değer ve alt normları devreye girmektedir.

Aşağıda her bir ana değer, kapsadığı alt normlar ve güncel düzenlemelerle ilişkilendirilerek açıklanmaktadır.

1.3.1. Faydalılık (Beneficence)

Teknolojinin ve algoritmaların varoluş amacının insanlığa hizmet etmesi gerektiği ilkesidir. Yapay zekânın yalnızca ekonomik bir kâr aracı olarak değil; insan refahını, esenliğini ve ekolojik sürdürülebilirliği proaktif olarak artırması hedeflenir (Floridi vd., 2018). Bu ana değer; *esenlik, bireysel ve toplumsal etki, insan refahı/refahı teşvik etmek* ve *sürdürülebilirlik* alt normlarını kapsamaktadır. Avrupa Birliği Yapay Zekâ Yasası (EU AI Act, 2024) da güvenilir yapay zekânın “insan refahını artırma” hedefini temel bir gereklilik olarak konumlandırmaktadır.

1.3.2. Zararsızlık (Non-Maleficence)

Tıp etiğindeki “önce zarar verme” (*primum non nocere*) ilkesinin yapay zekâyı uyarlanmış hâlidir. Yapay zekâ tasarımının ve kullanımının; insan onuruna, temel haklarına ve psikolojik bütünlüğüne zarar vermemesi anlamına gelir. Bu ana değer; *sorumluluk, insan hakları/onur, haksızlıktan kaçınmak, kötüye kullanım farkındalığı* ve *gizlilik* alt normlarını içermektedir. Algoritmaların tasarım aşamasında geliştiricilerin insani önyargılarından arındırılması (Mittelstadt vd., 2016) ve sistemin kötüye kullanım potansiyelinin öngörülerek güvenlik açıklarının kapatılması bu değer gereğidir. UNESCO’nun (2021) Yapay Zekâ Etiği Tavsiyesi de “orantılılık ve zarar vermemesi” ilkesiyle bu değeri küresel ölçekte normatif bir standart hâline getirmeyi amaçlamaktadır.

1.3.3. Adalet (Justice)

Dijital dünyanın ve akıllı sistemlerin sunduğu nimetlerden toplumun tüm kesimlerinin eşit yararlanması ve külfetlerin adil dağılımıdır. Bu ana değer; *dayanışmayı korumak, emniyette hissetmek, güvenlik ve güvenilirlik* alt normlarını kapsamaktadır. Yapay zekâ özelinde adalet; algoritmik kararların ırk, cinsiyet, inanç veya sınıf temelli ayrımcılık yapmadan işlemesidir. Angwin ve diğerlerinin (2016) ProPublica araştırmasında belgelediği COMPAS algoritması

örneğinde olduğu gibi, yapay zekâ sistemleri mevcut toplumsal eşitsizlikleri yeniden üretme ve derinleştirme potansiyeli taşımaktadır.

1.3.4. Özerklik (Autonomy)

Bireylerin teknoloji karşısındaki iradesinin ve bağımsızlığının korunmasıdır. Bu ana değer; *veri mülkiyeti haklarını desteklemek, mahremiyet, yetkinlik ve etkinlik* alt normlarını içermektedir. Yapay zekâ bağlamında özerklik; makinenin, insanın karar verme yetkisini rızası dışında gasp etmemesini ifade eder. İnsanların kişisel verileri üzerindeki mülkiyet haklarının korunması ve otonom sistemlerde “etkin insan denetiminin” (*human-in-the-loop*) mutlak surette sağlanması özerklik değerinin temelidir (Cath vd., 2018). Makine kendi kendine karar alabilse dahi nihai yetki (sistemi devre dışı bırakma gücü) insanda kalmalıdır. AB Yapay Zekâ Yasası, yüksek riskli yapay zekâ sistemleri için insan gözetimini zorunlu bir koşul olarak düzenlemektedir (EU AI Act, 2024, Madde 14).

1.3.5. Açıklık (Explicability)

Yapay zekâ etiğini geleneksel etikten ayıran en özgün ve kritik değerdir. Bu ana değer; *anlaşılabilirlik, hesap verebilirlik, şeffaflık ve amaca uygunluk* alt normlarını kapsamaktadır. Kendi kendine öğrenen algoritmaların karar alma süreçlerinin anlaşılabilmesi “kara kutu” (*black box*) problemi olarak adlandırılmakta ve bu durum, hesap verebilirlik ile şeffaflık açısından ciddi sorunlar doğurmaktadır (Pasquale, 2015). Floridi ve Cows’ a (2019) göre açıklık değeri, epistemolojik anlaşılabilirlik (*intelligibility*) ve hesap verebilirlik (*accountability*) olmak üzere iki alt boyuttan oluşur. AB Yapay Zekâ Yasası’nın şeffaflık yükümlülükleri (Madde 13) ve Genel Veri Koruma Yönetmeliği’nin (GDPR) otomatik karar almaya ilişkin 22. maddesi, açıklık değerinin hukuki zemindeki somut karşılıklarıdır.

Önerilen “Değer Temelli Dijital Etik Modeli”, yapay zekâ ve akıllı sistem kaynaklı ihlallerin salt hukuki yaptırımlarla değil, ancak tasarım, kullanım ve toplumsal etki süreçlerine bu beş değer ve bunlara bağlı yirmi dört alt normun baştan entegre edilmesiyle (*ethics by design*) aşılabileceğini öngörmektedir (Dignum, 2019). Model, her bir dijital değer için Turner’ın Sosyal Drama aşamalarında “düzeltici eylem” için birer normatif kriter işlevi göreceğini varsaymakta; böylece algoritmik ihlallerin ortaya çıkmasının ardından yaşanan kriz evresinde, hangi değer ihlal edildiğinin tespit

edilmesi ve sistemin bu değere uygun biçimde yeniden yapılandırılması yoluyla toplumsal yeniden bütünleşmenin sağlanabileceğini ileri sürmektedir.

2. Metodoloji ve Araştırma

2.1. Araştırma Yöntemi

Bu çalışmada, önerilen “Değer Temelli Dijital Etik Model”inin, gerçek dünyadaki yapay zekâ ve akıllı sistem ihlallerini çözümleme kapasitesini test etmek üzere nitel araştırma yöntemlerinden örnek olay (vaka) incelemesi (case study) kullanılmıştır. Vaka incelemesi, araştırmacının sınırlı bir sistem ya da birden fazla sistemi, çoklu veri kaynakları aracılığıyla derinlemesine incelediği nitel bir araştırma desenidir (Creswell, 2013, s. 97). Yin (2018) ise vaka incelemesini, özellikle araştırmacının kontrolü dışında gelişen güncel olguları, “nasıl” ve “neden” soruları ekseninde ampirik olarak sorgulamak için elverişli bir yöntem olarak tanımlamaktadır. Bu çalışma, yapay zekâ ihlallerinin nasıl gerçekleştiği ve bu ihlallere karşı hangi dijital değerlerin düzeltici eylem olarak işlev görebileceği sorularına yanıt aradığı için, vaka incelemesi yöntemi araştırmanın amaçlarıyla tam bir uyum göstermektedir.

2.2. Örneklem ve Veri Kaynakları

Araştırmada, amaçlı örneklem yöntemi benimsenmiştir. Patton (2015), amaçlı örneklemin, bilgi açısından zengin vakaların derinlemesine incelenmesine olanak tanıdığını ve bu yönüyle nitel araştırmanın doğasına en uygun örneklem stratejisi olduğunu belirtmektedir. Bu doğrultuda, Dünya Ekonomik Forumu (WEF) tarafından akıllı sistemler için belirlenen sorun alanları (Eggers ve Hickin, 2020) çerçevesinde, literatüre ve küresel medyaya yansımış, yapay zekâ ve akıllı sistem ihlallerini temsil eden sekiz büyük kriz vakası seçilmiştir. WEF raporunda dijital teknolojilerin kullanımına ilişkin sekiz temel sorun alanı tanımlanmıştır: (1) sınırlı düzenleme veya düzenleme eksikliği, (2) yanlış veya amaç dışı kullanım yoluyla teknolojinin olumsuz etkisi, (3) teknolojinin sorumluluğu ve hesap verebilirliği, (4) gizlilik ve veri paylaşımı, (5) kolluk kuvvetleri tarafından erişim ve kullanım, (6) siber ve diğer güvenlik endişeleri, (7) insan denetimi, (8) sınır ötesi tutarsızlıklar ve kısıtlı veri akışları. Bu çalışmada seçilen sekiz vaka, bu sekiz alanı

temsil edecek biçimde amaçlı örneklem yöntemiyle belirlenmiştir. Detroit yüz tanıma krizi (V1), yüz tanıma teknolojisine ilişkin yasal düzenleme eksikliğini (alan 1) örneklendirirken; BJP deepfake krizi (V2), teknolojinin seçim manipülasyonu amacıyla kötüye kullanımını (alan 2) göstermektedir. Uber otonom araç kazası (V3), otonom sistemlerde hesap verebilirlik sorununu (alan 3) gözler önüne sermektedir. Cambridge Analytica skandalı (V4) ve Tapu veri sızıntısı (V5), gizlilik ihlalleri ve verilerin kötüye kullanımı (alan 4) kapsamında değerlendirilmiştir. Kargu-2 otonom dron saldırısı (V6) ve Viking Sky cruise faciası (V7), insan denetiminin yetersizliği (alan 7) ile doğrudan ilişkilidir; ayrıca bu vakalar, teknolojinin sorumluluğu (alan 3) boyutunu da içermektedir. TikTok yasaklama krizi (V8) ise sınır ötesi veri akışlarındaki düzenleyici farklılıkları (alan 8) örneklendirmektedir. Bu eşleştirme, her bir vakanın WEF'in tanımladığı küresel teknoloji yönetim sorunlarıyla doğrudan bağlantılı olduğunu ve önerilen Değer Temelli Dijital Etik Modeli'nin test edilmesi için uygun bir zemin sunduğunu ortaya koymaktadır.

Vakaların seçiminde şu kriterler gözetilmiştir:

- i. Olayın yapay zekâ ya da otonom sistem kararlarından doğrudan kaynaklanması,
- ii. Olayın kamusal görünürlüğe ve belgelenmiş verilere sahip olması,
- iii. Farklı dijital değer ihlallerini temsil ederek modelin sınanmasına olanak tanınması.

Her bir vaka; Turner'ın (1974) Sosyal Drama modelindeki ihlal, kriz, düzeltici eylem ve çözüm/yeniden bütünleşme aşamaları bağlamında incelenmiş ve ihlalin hangi dijital değerlerde yoğunlaştığı değerlendirilmiştir.

2.3. Analiz Stratejisi

Vakaların analizinde, Yin'in (2018) önerdiği çapraz vaka sentezi (cross-case synthesis) yaklaşımı benimsenmiştir. Her bir vaka önce kendi içinde bütüncül olarak ele alınmış, ardından tüm vakalar karşılaştırmalı biçimde değerlendirilerek ihlal edilen dijital değerlerin dağılımı ve örüntüleri ortaya konmuştur. Bu analitik strateji, hem her bir vakanın kendine özgü dinamiklerini kavramaya hem de vakalar arası ortaklaşan etik ihlal biçimlerini tespit etmeye olanak sağlamıştır.

2.4. Vaka Analizleri

Aşağıda, sekiz vakanın her biri Turner'ın Sosyal Drama modeli çerçevesinde incelenmiş; ihlalin doğası, etkilenen dijital değerler ve ihlalin kaynaklandığı boyut (Tasarım, Kullanım, Toplumsal Etki) değerlendirilmiştir.

Vaka 1. Bir Algoritma Tarafından Yanlışlıkla Suçlanmak - Detroit Yüz Tanıma Krizi

Başlık	Açıklama
Kriz Tarihi	24.06.2020 (olay Ocak 2020'de gerçekleşmiştir)
Krizde Kullanılan Teknoloji	Yüz Tanıma Sistemleri, Yapay Zekâ Algoritmaları
Kriz Alanı	Sınırlı düzenleme veya düzenleme eksikliği
İhlal Durumu	Tasarım, Toplumsal Etki
İhlal Kapsamı	Yapay Zekâ / Akıllı Sistemler
Krizin Kısa Özeti	Detroit'te bir yapay zekâ yüz tanıma algoritması, güvenlik kamerası görüntüsündeki şüpheliyi hatalı eşleştirerek Robert Williams adlı siyahi vatandaşın evinin önünde tutuklanmasına neden olmuştur. Williams 30 saat gözaltında kalmış, polis fotoğraftaki kişinin Williams olmadığını fark edince serbest bırakılmıştır. Bu olay, yüz tanıma teknolojisinin hatalı eşleştirme nedeniyle bir kişinin tutuklanmasına yol açtığı ilk vaka olarak kayıtlara geçmiştir.
İhlale Uğrayan Etik Değerler	Faydalılık, Zararsızlık, Adalet, Özerklik, Açıklık
Turner'ın Sosyal Drama Aşamaları	İhlal: Algoritmanın ırksal önyargı ile hatalı eşleştirme yapması. Kriz: Masum bir kişinin tutuklanması ve toplumsal güvenin sarsılması. Düzeltilici Eylem: Detroit Emniyeti'nin yeni kurallar getirmesi, Williams'ın açtığı davanın 300.000 dolar tazminatla sonuçlanması. https://www.nytimes.com/2020/06/24/technology/facial-recognition-arrest.html
Kaynak	

Vaka 2. Deepfakes Videolar ile Gerçek Nasıl Bükülüyor - BJP Örneği

Başlık	Açıklama
Kriz Tarihi	20.02.2020
Krizde Kullanılan Teknoloji	Yapay Zekâ ve Yapay Öğrenme Sistemleri (Deepfake)
Kriz Alanı	Yanlış veya amaç dışı kullanım yoluyla teknolojinin olumsuz etkisi
İhlal Durumu	Kullanım, Toplumsal Etki
İhlal Kapsamı	Yapay Zekâ / Akıllı Sistemler
Krizin Kısa Özeti	2020 Delhi seçimleri sırasında BJP lideri Manoj Tiwari'nin deepfake teknolojisiyle üretilmiş, Haryanvi ve İngilizce konuştuğu sahte videoları 5.800'den fazla WhatsApp grubunda paylaşılarak yaklaşık 15 milyon kişiye ulaşmıştır. Videolar "The Ideaz Factory" adlı PR şirketi tarafından Tiwari'nin eski bir videosu ve ses sanatçısı kullanılarak üretilmiştir.
İhlale Uğrayan Etik Değerler	Faydalılık, Zararsızlık, Adalet, Açıklık
Turner'ın Sosyal Drama Aşamaları	İhlal: Deepfake teknolojisiyle sahte video üretilmesi ve dolaşıma sokulması. Kriz: Seçmenlerin yanıltılması, demokratik sürece güvenin sarsılması.

Kaynak	Düzeltilici Eylem: Hindistan Seçim Komisyonu'nun düzenleme arayışına girmesi (henüz yapısal çözüm yok). https://www.vice.com/en/article/939d4p/bjp-manoj-tiwari-deepfake-twists-truth
---------------	---

Vaka 3. Uber'in Otonom Aracı Ölümcul Kazaya Karıştı

Başlık	Açıklama
Kriz Tarihi	16.09.2020 (kaza 18 Mart 2018'de gerçekleşmiştir)
Krizde	Otonom Sürüş Sistemleri
Kullanılan Teknoloji	Teknolojinin sorumluluğu ve hesap verebilirliği
Kriz Alanı	Tasarım, Kullanım
İhlal Durumu	Yapay Zekâ / Akıllı Sistemler
İhlal Kapsamı	Arizona'da Uber'e ait otonom test aracı, yolun karşısına geçen Elaine Herzberg'i sensörleriyle "tanımlanamaz nesne" olarak sınıflandırıp fren yapmayarak ölümüne neden olmuştur. Güvenlik sürücüsü Rafaela Vasquez'in o sırada TV programı izlediği tespit edilmiştir. Bu kaza, otonom bir aracın karıştığı ilk ölümlü kaza olarak kayıtlara geçmiştir.
Krizin Kısa Özeti	Faydalılık, Zararsızlık, Adalet, Özerklik, Açıklık
İhlale Uğrayan Ana Etik Değerler	İhlal: Yapay zekânın yayayı yanlış sınıflandırması.
Turner'ın Sosyal Drama Aşamaları	Kriz: Ölümlü kaza, Uber'in test programının askıya alınması. Düzeltilici Eylem: NTSB soruşturması, Uber'in Arizona'daki testleri sonlandırması.
Kaynak	https://www.bbc.com/news/technology-54175359

Vaka 4. Facebook-Cambridge Analytica Veri Skandalı

Başlık	Açıklama
Kriz Tarihi	13.08.2019 (skandal 2018'de ortaya çıkmıştır)
Krizde	Sosyal Medya, Veri Tabanları ve Büyük Veri Analizi
Kullanılan Teknoloji	Gizlilik ve veri paylaşımı
Kriz Alanı	Tasarım
İhlal Durumu	Yapay Zekâ / Akıllı Sistemler
İhlal Kapsamı	Cambridge Analytica, Facebook üzerinden "This is Your Digital Life" adlı kişilik testi uygulamasıyla 87 milyon kullanıcının verilerini izinsiz toplamış; bu verileri ABD seçimleri ve Brexit referandumunda mikro-hedeflemeli siyasi reklamlar için kullanmıştır. Skandal, Christopher Wylie'nin ifşaatlarıyla ortaya çıkmıştır.
Krizin Kısa Özeti	Faydalılık, Zararsızlık, Adalet, Özerklik, Açıklık
İhlale Uğrayan Ana Etik Değerler	İhlal: Kullanıcı verilerinin rıza dışı toplanması.
Turner'ın Sosyal Drama Aşamaları	Kriz: Seçim manipülasyonu, küresel ölçekte veri gizliliği tartışmaları. Düzeltilici Eylem: Facebook'a 5 milyar dolar ceza, Cambridge Analytica'nın kapatılması, GDPR'ın yürürlüğe girmesi.
Kaynak	https://www.bbc.com/turkce/haberler-dunya-48974921

Vaka 5. Tapuda Skandal — Ünlü İsimlerin Bilgileri Sızdırıldı

Başlık	Açıklama
Kriz Tarihi	23.11.2021
Krizde	Veri Tabanları, E-devlet Sistemleri
Kullanılan Teknoloji	
Kriz Alanı	Kolluk kuvvetleri tarafından erişim ve kullanım
İhlal	Kullanım, Toplumsal Etki
Durumu	
İhlal	Yapay Zekâ ve Akıllı Sistemler
Kapsamı	
Krizin Kısa Özeti	Fatih Tapu Müdürlüğü'nde görevli iki memur, yetkilerini kötüye kullanarak 136 ünlü ismin taşınmaz mülk, nüfus, aile bireyleri bilgilerine izinsiz erişmiş ve bu bilgileri emlakçı ve iş takipçilerine satmıştır. İki memur, kişisel verileri hukuka aykırı ele geçirmekten 3 yıl 1 ay 15 gün hapis cezası almıştır.
İhlale Uğrayan Ana Etik Değerler	Faydalılık, Zararsızlık, Adalet, Özerklik, Açıklık
Turner'ın Sosyal Drama Aşamaları	İhlal: Memurların yetkilerini aşarak verilere izinsiz erişmesi. Kriz: Kamu kurumlarına güvenin sarsılması, ünlü isimlerin mağduriyeti. Düzeltilici Eylem: Adli soruşturma ve hapis cezası.
Kaynak	https://www.memurlar.net/haber/1018882/unlulere-bilgilerini-sizdiran-iki-tapu-memuruna-hapis.html Vakanın Yapay Zekâ ve Akıllı Sistemler Kapsamına Eklemlenme Gerekçesi: Her ne kadar bu vaka doğrudan bir yapay zekâ algoritmasının kararından kaynaklanmamış olsa da olayın merkezinde yer alan e-Devlet sistemleri ve entegre veri tabanları, yapay zekâ destekli akıllı hizmet altyapılarının klasik bir örneğini oluşturmaktadır. Türkiye'deki e-Devlet sistemi, vatandaş verilerinin toplanması, işlenmesi ve paylaşılması süreçlerinde yapay zekâ ve büyük veri analitiği ile beslenmektedir. Bu vakada, akıllı sistemlerin sağladığı veriye hızlı erişim ve işleme avantajı (faydalılık değeri), insan faktörünün kötü niyetli kullanımıyla bir mahremiyet krizine dönüşmüştür. Olay, akıllı sistemlerin tasarım aşamasında "içeriden gelen tehdit" riskine karşı yeterli güvenlik katmanlarının ve erişim denetim mekanizmalarının geliştirilmemesinin somut bir sonucudur. Bu yönüyle vaka, yapay zekâ ve akıllı sistemlerin kullanım boyutundaki etik zafiyetleri temsil etmekte ve özerklik ile mahremiyet değerlerinin ihlalini gözler önüne sermektedir.

Vaka 6. İnsansız Hava Aracının İnsan Öldürmesi- Kargu-2 Krizi

Başlık	Açıklama
Kriz Tarihi	24.06.2021 (olay 27 Mart 2020'de gerçekleşmiştir)
Krizde	Otonom Silahlı İnsansız Hava Aracı (Kargu-2)
Kullanılan Teknoloji	
Kriz Alanı	Siber ve diğer güvenlik endişeleri
İhlal Durumu	Tasarım, Kullanım
İhlal Kapsamı	Yapay Zekâ / Akıllı Sistemler
Krizin Kısa Özeti	BM Libya Uzmanlar Paneli raporuna göre, Libya iç savaşında Kargu-2 kamikaze dronları, operatörle veri bağlantısı koptuğu hâlde kendi yapay zekâ inisiyatifiyle insan hedeflerini belirleyip yok etmiştir. Bu olay, bir makinenin doğrudan insan onayı olmadan insan öldürdüğü ilk belgelenmiş vakalardan biridir.
İhlale Uğrayan Ana Etik Değerler	Faydalılık, Zararsızlık, Adalet, Özerklik, Açıklık

Turner'ın Sosyal Drama Aşamaları

Kaynak

İhlal: Otonom silahın insan onayı olmadan hedef belirleyip saldırması.
Kriz: Uluslararası hukuk ve etik tartışmalarının alevlenmesi.
Düzeltilici Eylem: BM raporunun yayımlanması, ölümcül otonom silahların yasaklanmasına yönelik uluslararası çağrılar (henüz bağlayıcı çözüm yok).
<https://www.bbc.com/turkce/haberler-turkiye-57586692>

Vaka 7. Otonom Sürüş Sahip Gemi - Viking Sky Cruise Faciası

Başlık	Açıklama
Kriz Tarihi	09.04.2019 (olay 23 Mart 2019'da gerçekleşmiştir)
Krizde Kullanılan Teknoloji	Otonom Araçlar, Yapay Zekâ ve Akıllı Sistemler
Kriz Alanı	İnsan denetimi
İhlal Durumu	Tasarım
İhlal Kapsamı	Yapay Zekâ / Akıllı Sistemler
Krizin Kısa Özeti	1.373 yolcu taşıyan Viking Sky gemisi, Norveç açıklarında fırtınaya yakalanmış; motor yağ seviyesi sensörünün hatalı alarmı nedeniyle akıllı otomasyon sistemi tüm motorları otonom olarak kapatmıştır. Sistem, kaptanın manuel müdahalesine uzun süre izin vermemiş, gemi kayalıklara bir gemi boyu mesafede durdurulabilmiştir. Yolcuların 479'u helikopterle tahliye edilmiştir.
İhlale Uğrayan Ana Etik Değerler	Faydalılık, Zararsızlık, Adalet, Özerklik, Açıklık
Turner'ın Sosyal Drama Aşamaları	İhlal: Sensör arızası nedeniyle yapay zekânın motorları hatalı kapatması. Kriz: Geminin kontrolsüz sürüklenmesi, can kaybı riski. Düzeltilici Eylem: Kaptanın manuel müdahale ile motorları yeniden başlatması, Norveç Denizcilik Otoritesi soruşturması.
Kaynak	https://www.aitrends.com/ai-insider/human-in-the-loop-vs-out-of-the-loop-in-ai-systems-the-case-of-ai-self-driving-cars/

Vaka 8. ABD'den Beklenen TikTok Yasağı Geldi

Başlık	Açıklama
Kriz Tarihi	18.09.2020
Krizde Kullanılan Teknoloji	Sosyal Medya, Veri Tabanları
Kriz Alanı	Sınır ötesi tutarsızlıklar ve kısıtlı veri akışları
İhlal Durumu	Tasarım, Kullanım, Toplumsal Etki
İhlal Kapsamı	Yapay Zekâ / Akıllı Sistemler
Krizin Kısa Özeti	ABD, ulusal güvenlik gerekçesiyle Çinli ByteDance şirketine bağlı TikTok uygulamasını 20 Eylül 2020'den itibaren yasaklayacağını açıklamıştır. ByteDance'in Çin hükümeti ile veri paylaşımı taahhüdü bulunması, milyonlarca kullanıcının verilerinin güvenliğine ilişkin endişeleri tetiklemiştir. TikTok, Oracle ile ortaklık kurarak faaliyetlerine devam etme hakkı kazanmıştır.
İhlale Uğrayan Ana Etik Değerler	Faydalılık, Zararsızlık, Adalet, Özerklik, Açıklık
Turner'ın Sosyal Drama Aşamaları	İhlal: Kullanıcı verilerinin uluslararası aktarımı ve potansiyel casusluk. Kriz: Ulusal güvenlik tehdidi algısı, yasaklama kararları. Düzeltilici Eylem: TikTok'un Oracle ile ortaklık kurması, ABD'de yasal düzenleme sürecinin devam etmesi (kriz kısmen çözülmüştür).
Kaynak	https://www.amerikaninsesi.com/a/abd-den-beklenen-tiktok-yasagi-geldi/5588827.html

2.5. Vakaların Karşılaştırmalı Değerlendirmesi ve Bulgular

Aşağıdaki tablo, sekiz vakanın her birini; ihlal edilen dijital değerler, ihlalin boyutu (Tasarım, Kullanım ya da Toplumsal Etki) ve Turner'ın Sosyal Drama modelindeki aşamalar ekseninde karşılaştırmalı olarak özetlemektedir.

Tablo 1. Vakaların Değer Temelli Dijital Etik Modeli Çerçevesinde Karşılaştırmalı Analizi

<i>Vaka No</i>	<i>Vaka Adı</i>	<i>Yıl</i>	<i>İhlal Edilen Değerler</i>	<i>İhlal Boyutu</i>	<i>Sosyal Drama Aşamaları</i>
1	Detroit Tanıma Krizi	Yüz 2020	Faydalılık, Zararsızlık, Adalet, Özerklik, Açıklık	Tasarım, Toplumsal Etki	İhlal → Kriz → Düzeltilici Eylem
2	Manoj Tiwari Deepfake Krizi	2020	Faydalılık, Zararsızlık, Adalet, Açıklık	Kullanım, Toplumsal Etki	İhlal → Kriz
3	Uber Otonom Araç Kazası	2018	Faydalılık, Zararsızlık, Adalet, Özerklik, Açıklık	Tasarım, Kullanım	İhlal → Kriz → Düzeltilici Eylem
4	Cambridge Analytica Skandalı	2018	Faydalılık, Zararsızlık, Adalet, Özerklik, Açıklık	Tasarım	İhlal → Kriz → Düzeltilici Eylem
5	Tapu Veri Sızıntısı	2021	Faydalılık, Zararsızlık, Adalet, Özerklik, Açıklık	Kullanım, Toplumsal Etki	İhlal → Kriz → Düzeltilici Eylem
6	Kargu-2 Otonom Dron Saldırısı	2020	Faydalılık, Zararsızlık, Adalet, Özerklik, Açıklık	Tasarım, Kullanım	İhlal → Kriz
7	Viking Sky Cruise Faciası	2019	Faydalılık, Zararsızlık, Adalet, Özerklik, Açıklık	Tasarım	İhlal → Kriz → Düzeltilici Eylem
8	TikTok Yasaklama Krizi	2020	Faydalılık, Zararsızlık, Adalet, Özerklik, Açıklık	Tasarım, Kullanım, Toplumsal Etki	İhlal → Kriz

Not: “Düzeltilici Eylem” aşaması, vakanın ardından hukuki düzenleme, kurumsal politika değişikliği ya da uluslararası mutabakat gibi somut bir düzeltici adımın atıldığı vakalar için işaretlenmiştir. “Kriz” aşamasında kalan vakalar, henüz yapısal bir çözüme kavuşmamış, süregiden etik krizlerdir.

2.6. Modelin Sınanmasına İlişkin Ön Değerlendirme

Sekiz vakanın karşılaştırmalı analizi, önerilen Değer Temelli Dijital Etik Modeli'nin, yapay zekâ ve akıllı sistem ihlallerini sınıflandırma ve çözümleme kapasitesine ilişkin önemli ön bulgular sunmaktadır. Tablo 1'den de izlenebileceği gibi, Zararsızlık ve Adalet değerleri vakaların tamamında (8/8) ihlal edilirken; Faydalılık ve Özerklik 7'ser vakada, Açıklık ise 8 vakada ihlal edilmiştir. Bu dağılım, yapay zekâ sistemlerinin yol açtığı etik krizlerin çok boyutlu doğasını ve beş temel dijital değer neredeyse her krizde eşzamanlı olarak zarar gördüğünü ortaya koymaktadır.

İhlallerin boyutu açısından bakıldığında, sekiz vakanın beşi Tasarım aşamasından, üçü ise Kullanım aşamasından kaynaklanmaktadır (bazı vakalar her iki boyutu da içermektedir). Bu bulgu, etik sorunların büyük ölçüde teknolojinin geliştirilme aşamasında kodlandığını ve bu nedenle “tasarım yoluyla etik” (ethics by design) yaklaşımının hayati önem taşıdığını doğrulamaktadır (Dignum, 2019).

Turner'ın Sosyal Drama modeli açısından değerlendirildiğinde, sekiz vakanın beşinde “düzeltici eylem” aşamasına ulaşıldığı (hukuki yaptırım, kurumsal politika değişikliği ya da teknik düzeltme yoluyla), buna karşılık üç vakanın (Deepfake, Kargu-2 ve TikTok) henüz kriz aşamasında kaldığı ve yapısal bir çözüme kavuşmadığı görülmektedir. Bu durum, özellikle hızlı gelişen ve sınır ötesi etkiler taşıyan teknolojilerde düzeltici eylemlerin geciktigine işaret etmektedir.

Matristen Elde Edilen Nicel Bulgular

En sık ihlal edilen alt normlar incelendiğinde, Esenlik, Bireysel ve Toplumsal Etki, İyi Olmak, İnsan Refahı / Refahı Teşvik Etmek, Sürdürülebilirlik, Sorumluluk, İnsan Hakları, Onur, Kötüye Kullanım Farkındalığı, Haksızlıktan Kaçınmak, Dayanışmayı Korumak, Emniyette Hissetmek, Güvenlik, Güvenirlilik, Yetkinlik, Etkinlik, Anlaşılabilirlik, Hesap Verebilirlik, Şeffaflık ve Amaca Uygunluk olmak üzere yirmi alt norm sekiz vakanın tamamında ihlal edilmiştir. Bunların dışında kalan dört alt norm –Finansal ve Ekonomik Etki altı vakada, Gizlilik dört vakada, Veri Mülkiyeti Haklarını Desteklemek dört vakada ve Mahremiyet dört vakada– daha düşük oranda ihlal edilmiştir.

Tablo 2. Değer Temelli Dijital Etik Modeli – Örnek Olay İnceleme Matrisi*

Ana Değer	Alt Norm / Etik Norm	V1	V2	V3	V4	V5	V6	V7	V8	Toplam İhlal
Faydalılık	Esenlik	✓	✓	✓	✓	✓	✓	✓	✓	8
	Bireysel ve Toplumsal Etki	✓	✓	✓	✓	✓	✓	✓	✓	8
	Finansal ve Ekonomik Etki	✓		✓	✓	✓		✓	✓	6
	İyi Olmak	✓	✓	✓	✓	✓	✓	✓	✓	8
	İnsan Refahı / Refahı Teşvik Etmek	✓	✓	✓	✓	✓	✓	✓	✓	8
	Sürdürülebilirlik	✓	✓	✓	✓	✓	✓	✓	✓	8
Zararsızlık	Sorumluluk	✓	✓	✓	✓	✓	✓	✓	✓	8
	Gizlilik	✓			✓	✓			✓	4
	İnsan Hakları	✓	✓	✓	✓	✓	✓	✓	✓	8
	Onur	✓	✓	✓	✓	✓	✓	✓	✓	8
	Kötüye Kullanım Farkındalığı	✓	✓	✓	✓	✓	✓	✓	✓	8
	Haksızlıktan Kaçınmak	✓	✓	✓	✓	✓	✓	✓	✓	8
Adalet	Dayanışmayı Korumak	✓	✓	✓	✓	✓	✓	✓	✓	8
	Emniyette Hissetmek	✓	✓	✓	✓	✓	✓	✓	✓	8
	Güvenlik	✓	✓	✓	✓	✓	✓	✓	✓	8
	Güvenirlilik	✓	✓	✓	✓	✓	✓	✓	✓	8
Özerklik	Veri Mülkiyeti Haklarını Desteklemek	✓			✓	✓			✓	4
	Mahremiyet	✓			✓	✓			✓	4
	Yetkinlik	✓	✓	✓	✓	✓	✓	✓	✓	8
	Etkinlik	✓	✓	✓	✓	✓	✓	✓	✓	8
Açıklık	Anlaşılabilirlik	✓	✓	✓	✓	✓	✓	✓	✓	8
	Hesap Verebilirlik	✓	✓	✓	✓	✓	✓	✓	✓	8
	Şeffaflık	✓	✓	✓	✓	✓	✓	✓	✓	8
	Amaca Uygunluk	✓	✓	✓	✓	✓	✓	✓	✓	8

Ana değerler bazında ihlal yoğunluğuna bakıldığında, toplam ihlal sayılarına göre sıralama şöyledir: Faydalılık 46 ihlal, Zararsızlık 44 ihlal, Adalet 32 ihlal, Açıklık 32 ihlal ve Özerklik 24 ihlal olarak gerçekleşmiştir. Bu dağılım, özellikle Adalet ve Açıklık değerlerinin neredeyse tüm vakalarda eşzamanlı olarak ihlal edildiğini göstermektedir.

Tüm değerlerin ihlal edildiği vakalar incelendiğinde, Detroit Yüz Tanıma (V1), Uber Otonom Araç Kazası (V3), Cambridge Analytica (V4), Tapu Veri Sızıntısı (V5), Kargu-2 Otonom Dron Saldırısı (V6), Viking Sky Cruise Faciası (V7) ve TikTok Yasaklama Krizi (V8) olmak üzere yedi vakada beş ana değerlerin tamamı ihlal edilmiştir. Yalnızca BJP Deepfake Krizi (V2) vakasında Özerklik

* 8 vaka × 5 ana değer × 24 alt norm biçiminde niceliksel matris sunulmuştur. Her bir hücrede "✓" işareti, o vakanın ilgili alt normu ihlal ettiğini göstermektedir.

Not: V1=Detroit Yüz Tanıma, V2=Deepfake BJP, V3=Uber Kazası, V4=Cambridge Analytica, V5=Tapu Sızıntısı, V6=Kargu-2, V7=Viking Sky, V8=TikTok*

değeri ihlal edilmemiş; bu vakada Özerklik alt normlarından hiçbirisi zarar görmemiştir.

En kapsamlı ihlal ise Detroit Yüz Tanıma Krizi'nde (V1) gerçekleşmiştir: beş ana değer tamamı ve 24 alt normun tamamı ihlal edilmiştir. Bu vaka, algoritmik önyargının ırksal ayrımcılıkla birleşerek insan onurunu, özerkliği, adaleti ve hesap verebilirliği eşzamanlı olarak nasıl tahrip edebileceğinin en çarpıcı örneğini oluşturmaktadır.

3. Bulgular ve Tartışma

Araştırma kapsamında, Değer Temelli Dijital Etik Modeli çerçevesinde incelenen sekiz yapay zekâ ve akıllı sistem vakasına ait veriler (Tablo 1), dijital çağda karşı karşıya kalınan etik ihlallerin tesadüfi veya münferit hatalar olmadığını; aksine sistemselsel, ontolojik ve tasarımsal zafiyetlerden beslenen yapısal krizler olduğunu ampirik olarak ortaya koymaktadır. Yin'in (2018) çapraz vaka sentezi stratejisiyle elde edilen bulgular, dijital değerlerin ihlal örüntüleri, tasarım-kullanım ikiliği ve krizlerin çözülmemesi (sosyal drama) durumları olmak üzere üç ana başlık altında tartışılarak aşağıda sunulmuştur.

3.1. Dijital Değerlerin İhlal Örüntüleri ve “Kara Kutu” Problemi

Vakaların analizinden elde edilen en çarpıcı bulgu, Zararsızlık, Adalet, Açıklık ve Faydalılık değerlerinin incelenen sekiz vakanın tamamında (8/8) ağır bir şekilde ihlal edilmiş olmasıdır. Tablo 1'de görüldüğü üzere, bu dört değer hiçbir vakada korunamamış; Özerklik ise yedi vakada ihlal edilerek beş temel dijital değer eşzamanlı ve çok boyutlu bir erozyona uğradığını göstermiştir. Bu bulgu, Floridi ve diğerlerinin (2018) yapay zekâ etiğinin “tek bir ilkeye indirgenemeyecek kadar karmaşık ve çok katmanlı” olduğu yönündeki tespitini ampirik olarak doğrulamaktadır.

Yapay zekâ krizlerinin merkezinde yer alan en temel sorun, geleneksel etikten farklı olarak Açıklık ilkesinin yoksunluğundan kaynaklanan Kara Kutu (Black Box) problemidir. Pasquale'nin (2015) kapsamlı biçimde tartıştığı ve Mittelstadt ve diğerlerinin (2016) algoritmik kararların etik açıdan denetlenmesinin önündeki en büyük engel olarak konumlandığı bu sorun, incelenen vakalarda somut ve yıkıcı sonuçlar doğurmuştur. Floridi ve Cows (2019) tarafından yapay zekâ etiğinin kilit taşı olarak görülen hesap verebilirlik; Uber otonom araç kazasında (Vaka 3), makinenin yayayı neden önce “araç”, sonra “bisikletli”, nihayetinde “tanımlanamaz nesne” olarak

sınıflandırdığının anlaşılabilmesi ve hukuki bir muhatap (fail) bulunamamasıyla çıkmıştır. NTSB (2019) raporunun da teyit ettiği üzere, sistem epistemolojik olarak açıklanamadığı için krizin hukuki boyutu da kilitlenmekte; bu durum, hesap verebilirliğin yalnızca teknik bir gereklilik değil, aynı zamanda hukukun üstünlüğünün dijital çağdaki ön koşulu olduğunu göstermektedir.

Bununla birlikte Adalet ve Zararsızlık değerleri, algoritmaların “tarafsız ve yansız olduğu” mitini ampirik olarak çürütmüştür. Noble’ın (2018) “algoritmik baskı” (algorithms of oppression) kavramsallaştırması ve Benjamin’in (2019) teknolojinin ırksal eşitsizlikleri nasıl kodladığına dair analizi, Detroit Yüz Tanıma Krizinde (Vaka 1) somut bir karşılık bulmuştur. Büyük veri setleriyle (ağırlıklı olarak beyaz yüzlerle) eğitilen yapay zekâ, siyahi bir vatandaş yanlış teşhis ederek makine öğrenimi süreçlerindeki ırksal ve toplumsal önyargıları (bias) yeniden üretmiş; Williams’ın ailesinin önünde kelepçelenmesi, insan onurunun teknoloji eliyle nasıl sistematik biçimde zedelenebileceğini gözler önüne sermiştir. Benzer şekilde, Cambridge Analytica (Vaka 4) ve Deepfake (Vaka 2) krizlerinde makine, kitleleri manipüle etmek üzere asimetrik bir güç olarak kullanılmış; demokratik adalet ve toplumsal fayda ilkeleri teknolojik determinizme kurban edilmiştir. Bu vakalar, Zuboff’un (2019) “gözetim kapitalizmi” kavramıyla teorize ettiği, insan deneyiminin davranışsal veriye dönüştürülerek metalaştırılması sürecinin, etik değerleri nasıl yapısal olarak dışladığını göstermektedir.

Özerklik değerinin ihlali ise insanın makine karşısındaki ontolojik konumuna dair varoluşsal bir tartışmayı gündeme getirmektedir. Etkin insan denetiminin (human-in-the-loop) devre dışı bırakılması, sistemlerin en tehlikeli boyutunu oluşturmaktadır. Cath ve diğerlerinin (2018) yapay zekâ yönetiminde insan denetiminin merkezî rolüne dair vurgusu, incelenen iki vakada dramatik biçimde doğrulanmıştır. Viking Sky Cruise faciasında (Vaka 7), hatalı sensör verisiyle kararı tamamen eline alan yapay zekâ, insan kaptanın müdahalesini reddederek 1.373 yolcunun hayatını riske atmış; Kargu-2 dronları (Vaka 6) ise insan onayı olmadan otonom öldürme kararı alarak, makinenin özerkliğinin insanın iradesini ve yaşama hakkını nasıl gasp edebildiğini kanıtlamıştır. BM raporunun (UN Security Council, 2021) da teyit ettiği bu son vaka, otonom silah sistemlerinin

yasaklanmasına yönelik uluslararası kampanyaların (Human Rights Watch, 2021) aciliyetini somut bir zemine oturtmaktadır.

3.2. Tasarım ve Kullanım ve Toplumsal Etki Boyutlarındaki Etik Zafiyetler

Karşılaştırmalı analiz bulguları, teknolojik ihlallerin başlangıç noktasına dair kritik bir ayrımı ortaya koymaktadır. Tablo 1’den izlenebileceği üzere, sekiz vakanın altısı doğrudan Tasarım aşamasındaki etik kodlama eksikliklerinden kaynaklanmakta; bazı vakalar ise hem tasarım hem de kullanım boyutunda eşzamanlı ihlaller içermektedir.

Dignum’un (2019) literatürde güçlü biçimde savunduğu “tasarımdan itibaren etik” (*ethics by design*) yaklaşımının eksikliği, akıllı sistemleri doğuştan kriz potansiyeli taşıyan yapılara dönüştürmektedir. Uber kazasındaki (Vaka 3) sınıflandırma hatası, Viking gemisindeki (Vaka 7) müdahale kısıtlaması veya Detroit’teki (Vaka 1) algoritmik önyargı, son kullanıcının bir eyleminden ziyade, mühendislerin ve veri bilimcilerin sistemi kurarken Faydalılık ve Zararsızlık gibi değerleri koda entegre etmemelerinden (ihmalden) kaynaklanmaktadır. Hagendorff (2020) de benzer biçimde, mevcut yapay zekâ etiği ilkelerinin uygulamaya dönük bağlayıcı mekanizmalardan yoksun olduğunu ve bu nedenle teknik geliştirme süreçlerinde etik kodların sistematik olarak göz ardı edildiğini vurgulamaktadır. Avrupa Birliği’nin 2024’te yürürlüğe giren Yapay Zekâ Yasası (EU AI Act), bu boşluğu gidermeyi amaçlayan en kapsamlı yasal girişim olarak öne çıkmakta; özellikle yüksek riskli sistemler için getirdiği insan gözetimi, şeffaflık ve uygunluk değerlendirmesi zorunlulukları, tam da incelenen vakalarda ihlal edilen değerleri hukuki zeminde koruma altına almayı hedeflemektedir.

Kullanım ve Toplumsal Etki boyutunda gerçekleşen ihlaller ise, teknik açıdan kusursuz tasarlanmış olsa dahi büyük veri altyapılarının insan faktörüyle nasıl silaha dönüştürülebileceğini ve bu durumun geniş toplumsal kesimleri nasıl etkileyebileceğini göstermektedir. Tapu Veri Sızıntısı (Vaka 5) örneğinde, e-devlet entegrasyonlarının sağladığı hız ve erişim kolaylığı (Faydalılık), yetkili memurların kötü niyetli kullanımıyla bir mahremiyet krizine dönüşmüştür. Bu vaka, Gal ve diğerlerinin (2020) “içeriden gelen tehdit” (*insider threat*) kavramsallaştırmasıyla örtüşmekte ve etik ihlallerin yalnızca dış saldırganlardan değil, bizzat sistemin içindeki aktörlerden de kaynaklanabileceğini kanıtlamaktadır. Keza TikTok krizi (Vaka 8), sınır ötesi veri akışında tasarımın kendisinin ulus-devletler tarafından “kötü niyetli kullanım” şüphesiyle ulusal güvenlik

krizine çevrilebileceğini göstermiştir. Couldry ve Mejias'ın (2019) “veri sömürgeciliği” (*data colonialism*) kavramı, bu tür sınır ötesi veri akışlarının yalnızca bireysel mahremiyeti değil, ulusal egemenlik alanlarını da tehdit eden yapısal bir güç asimetrisi yarattığını teorize etmektedir.

Bu bulgular bütüncül olarak değerlendirildiğinde, etik ilkelerin yalnızca yazılımcıları ve tasarım ekiplerini değil; son kullanıcıları, şirketleri, kamu otoritelerini ve daha geniş toplumsal yapıyı da bağlayıcı, çok paydaşlı ve çok katmanlı bir zihniyet dönüşümünü (dijital etik okuryazarlığı) gerektirdiği açıkça ortaya çıkmaktadır. UNESCO'nun (2021) Yapay Zekâ Etiği Tavsiyesi, tam da bu bütüncül yaklaşımı küresel ölçekte kurumsallaştırmayı hedefleyen normatif bir çerçeve sunmaktadır. Sonuç olarak, teknolojik ihlallerin önlenmesi; tasarım, kullanım ve toplumsal etki boyutlarının tamamını eşzamanlı olarak kapsayan bir etik yönetim modelini zorunlu kılmaktadır.

3.3. Sosyal Drama ve Düzeltici Eylem Çıkmazı: Hukukun Gecikmesi

Victor Turner'ın (1974) Sosyal Drama modeli, krizlerin çözülüp toplumsal düzenin yeniden sağlanabilmesi için bir Düzeltici Eylem (redressive action) üretilmesi gerektiğini şart koşar. Tablo 1'deki veriler, bu aşamanın günümüz dijital ekosisteminde ciddi bir “çıkamaz” içinde olduğunu ampirik olarak belgelemektedir.

İncelenen sekiz vakanın beşinde (Detroit Yüz Tanıma, Uber Kazası, Cambridge Analytica, Tapu Sızıntısı ve Viking Sky) düzeltici eylem aşamasına ulaşıldığı görülmektedir. Cambridge Analytica (Vaka 4) örneğinde kriz patlak verdikten sonra Facebook'a 5 milyar dolarlık rekor bir ceza kesilmiş (FTC, 2019), GDPR gibi hukuki kalkanlar devreye sokulmuş ve şirket kapatılmıştır. Detroit Yüz Tanıma (Vaka 1) vakasında ise Williams'a 300.000 dolar tazminat ödenmiş ve polis teşkilatı yeni kurallar getirmiştir. Ancak bu müdahalelerin tamamı, mağduriyetler yaşandıktan sonra, yani reaktif olarak gerçekleşmiştir. Mevcut hukuk sisteminin teknolojik gelişmelerin hızına yetişememesi, Marchant ve diğerlerinin (2011) “hukuk ve teknoloji arasındaki giderek büyüyen uçurum” (pacing problem) olarak kavramsallaştırdığı yapısal bir soruna işaret etmektedir.

Daha da endişe verici olan bulgu; Deepfake (Vaka 2), Kargu-2 Otonom Silahlar (Vaka 6) ve TikTok (Vaka 8) vakalarının Turner'ın modelinde hâlâ Kriz

aşamasında takılı kalmasıdır. Otonom silahların insan öldürmesine karşı bağlayıcı bir uluslararası anlaşma henüz mevcut değildir; deepfake teknolojisiyle gerçekliğin manipülasyonuna dair kapsamlı yasal çerçeveler oluşturulamamıştır. Bu durum, Turner’ın teorisindeki bölünme (schism) riskini —yani toplumun teknolojiye, kurumlara ve hatta gerçeğin kendisine olan inancını tamamen yitirmesi tehlikesini— somut bir olasılık olarak gündeme getirmektedir. Danaher’in (2016) “algoritmik yönetişimin meşruiyet krizi” ve Yeung’un (2018) “hiper-ölçekli platformların düzenlenmesindeki yapısal zorluklar” üzerine analizleri, tam da bu çözümsüzlük durumunun demokratik kurumlar ve toplumsal sözleşme üzerindeki aşındırıcı etkilerine dikkat çekmektedir.

Sonuç olarak tartışma bulguları; yasal yaptırımların teknolojik ihlalleri önlemede yetersiz ve yavaş kaldığını ampirik olarak kanıtlamakta; bu tespit, literatürdeki “proaktif etik yönetişimi” çağrılılarıyla (Dignum, 2019; Floridi, 2021) tam bir uyum göstermektedir. Dijital dünyanın kendi doğasına uygun, yapay zekânın kara kutu risklerini minimize edecek, tasarımı ve kullanımı eşzamanlı denetleyecek Değer Temelli Dijital Etik Modeli’nin proaktif bir rehber olarak uygulanması, salt akademik bir önerme değil; dijital toplumun demokratik bütünlüğünü, insan onurunu ve epistemolojik güvenliğini koruyabilmesi için varoluşsal bir zorunluluktur.

4. Sonuç

Bilgisayar teknolojilerinin analog verileri dijitalleştirilmesiyle başlayan ve günümüzde büyük veri, nesnelerin interneti ve derin öğrenme algoritmalarının entegrasyonu ile şekillenen Toplum 5.0 evresi (Cabinet Office, Government of Japan, 2016), insanlığa daha önce benzeri görülmemiş bir teknolojik sıçrama sunmuştur. Ancak yapay zekâ ve akıllı sistemlerin insanın bilişsel sınırlarını aşarak otonom birer karar verici konumuna yükselmesi; mahremiyetin ihlalinde algoritmik ayrımcılığa, demokratik manipülasyondan yaşam hakkının gasp edilmesine kadar uzanan derin ahlaki ve hukuki krizleri beraberinde getirmiştir. Teknolojik determinizmin etik bir pusuladan yoksun biçimde ilerlemesi, akıllı sistemleri insanlığa hizmet eden araçlar olmaktan çıkarıp, kontrolden çıkmaya eğilimli potansiyel tehditlere dönüştürmektedir.

Bu çalışma, söz konusu krizlere yalnızca teşhis koymakla kalmayıp, sistematik bir çözüm çerçevesi sunma iddiasıyla yola çıkmıştır. Victor Turner’ın (1974) “Sosyal Drama” kuramı ile Floridi ve Cows’ın (2019) yapay zekâ etiği ilkelerini sentezleyerek

geliştirilen Değer Temelli Dijital Etik Modeli, bu iddianın kuramsal ve ampirik omurgasını oluşturmuştur. Modelin özgünlüğü; etik ihlalleri yalnızca soyut ilkeler düzeyinde tartışmak yerine, her bir krizi Turner’ın dört aşamalı Sosyal Drama süreci (ihlal → kriz → düzeltici eylem → yeniden bütünleşme) içinde konumlandırarak, hangi aşamada hangi değer nasıl onarılabileceğine dair operasyonel bir yol haritası sunmasında yatmaktadır.

Dünya Ekonomik Forumu’nun (Eggers ve Hickin, 2020) sorun alanları çerçevesinde seçilen sekiz yapay zekâ krizinin çapraz vaka analizi, modelin sahadaki açıklayıcı ve sınıflandırıcı gücünü ampirik olarak doğrulamıştır. Araştırma bulguları, teknolojik krizlerin rastlantısal olmadığını; ihlallerin büyük çoğunluğunun sistemlerin tasarım aşamasında, teknoloji geliştiricilerin Faydalılık, Zararsızlık, Adalet, Özerklik ve Açıklık değerlerini göz ardı etmesinden ya da bu değerleri koda entegre edememesinden kaynaklandığını açıkça ortaya koymuştur. Daha da kritik olan bulgu ise, mevcut hukuk sistemlerinin teknolojinin hızına yetişememesi nedeniyle düzeltici eylemlerin eksik kalması ve bazı vakaların (Kargu-2, Deepfake, TikTok) hâlâ kriz aşamasında takılı kalarak toplumsal bölünme riskini canlı tutmasıdır.

Bu çalışmanın en temel katkısı, Değer Temelli Dijital Etik Modeli’nin reaktif değil, proaktif bir yönetim aracı olarak konumlandırılmasıdır. Mevcut küresel pratikte etik ilkeler ve hukuki yaptırımlar, çoğunlukla kriz patlak verdikten sonra devreye giren “hasar kontrol” mekanizmaları olarak işlemektedir. Oysa bu çalışmada önerilen model, beş temel dijital değer daha teknoloji geliştirme aşamasında bir tasarım kontrol listesi olarak kullanılmasını; böylece ihlal henüz ortaya çıkmadan önlenmesini hedeflemektedir. Dignum’un (2019) “tasarımdan itibaren etik” (*ethics by design*) yaklaşımıyla tam bir uyum içinde olan bu model, aynı zamanda Turner’ın Sosyal Drama kuramı sayesinde, önlemeye rağmen ortaya çıkabilecek krizler için de aşamalı bir müdahale takvimi sunmaktadır.

Modelin proaktif işlevi şu üç düzlemde somutlaşmaktadır: (i) Tasarım öncesi aşamada, geliştiricilerin beş değeri teknik gereksinimlerle eşleştirerek potansiyel ihlalleri öngörmesi; (ii) Kriz anında, hangi değer ihlal edildiğinin hızlıca teşhis edilerek düzeltici eylemin yönünün belirlenmesi; (iii) Kriz sonrasında, ihlalden ders çıkarılarak sistemin yeniden yapılandırılması ve normatif standartların

güncellenmesi. Bu üç düzlemli yapı, modeli yalnızca akademik bir analiz aracı değil; teknoloji şirketleri, düzenleyici kurumlar ve sivil toplum için uygulanabilir bir etik yönetim çerçevesi hâline getirmektedir.

İnsan onurunu merkeze alan ve evrensel değerlere duyarlı bir dijital ekosistemin inşa edilebilmesi için şu temel öneriler geliştirilmiştir:

Tasarımdan İtibaren Etik Yaklaşımının Kurumsallaşması: Etik, yapay zekâ sistemleri kriz yarattıktan sonra uygulanacak bir yama veya hukuki bir ceza mekanizması olarak görülmemelidir. Avrupa Birliği Yapay Zekâ Yasası'nın (EU AI Act, 2024) yüksek riskli sistemler için öngördüğü uygunluk değerlendirmeleri, tam da bu anlayışın hukuki zemindeki karşılığıdır. Yazılım mühendisleri ve veri bilimciler, algoritmaları kodlarken ve veri setlerini eğitirken beş temel dijital değeri mimarinin merkezine yerleştirmeli; algoritmik önyargıları önlemek için dengeli ve yansız veri setlerinin kullanımı zorunlu tasarım standartları hâline getirilmelidir.

Açıklanabilir Yapay Zekâ (Explainable AI-XAI) ve Hesap Verebilirlik: Akıllı sistemlerin nasıl karar verdiğinin anlaşılamaması (kara kutu problemi), hukuk sistemini ve ahlaki sorumluluğu kilitlemektedir. Algoritmaların epistemolojik olarak anlaşılabilir ve şeffaf sonuçlar üretmesi, olası bir ihlal anında sorumlunun (yazılımcı, şirket veya operatör) net bir biçimde tespit edilebilmesi için yasal bir önkoşul olarak talep edilmelidir (Pasquale, 2015; Floridi ve Cowls, 2019).

Etkin İnsan Denetiminin Korunması: Otonom silahlar (Kargu-2) ve akıllı yönlendirme sistemlerinde (Viking Gemisi) görüldüğü üzere, makinenin insandan bağımsız karar alması telafisi imkânsız sonuçlar doğurmaktadır. Bir akıllı sistem ne kadar gelişmiş olursa olsun, özellikle insan yaşamını, özgürlüğünü ve mahremiyetini ilgilendiren kritik kararlarda nihai yetki her zaman insanda kalmalı (*human-in-the-loop*); makinenin sistemi tamamen devralması yasal olarak engellenmelidir (Cath vd., 2018).

Hukuk ve Etiğin Proaktif Senkronizasyonu: Yasalar genellikle teknolojik zararlar yaşandıktan sonra devreye girmektedir. Marchant ve diğerlerinin (2011) “hukuk ve teknoloji arasındaki uçurum” (*pacing problem*) olarak tanımladığı bu yapısal gecikmeyi aşmak için, çalışmada sunulan Değer Temelli Dijital Etik Modeli, yasa yapıcılar ve uluslararası kurumlar için bir erken uyarı ve denetim sistemi olarak kabul edilmelidir. Teknoloji politikaları, beş temel değer etrafında proaktif bir vizyonla şekillendirilmeli;

etik etki değerdendirmeleri, tıpkı çevresel etki değerdendirmeleri gibi, teknoloji geliştirme sürecinin zorunlu bir aşaması hâline getirilmelidir.

Bu çalışma, Değer Temelli Dijital Etik Modeli'nin gerçek dünya vakaları üzerinden sınanması bakımından özgün bir katkı sunmakla birlikte, bazı sınırlılıkları da bünyesinde taşımaktadır. Bu sınırlılıkların açıkça belirtilmesi, gelecek araştırmalar için yol gösterici olacaktır. Birincisi, araştırma kapsamında incelenen vaka sayısı sekiz ile sınırlıdır. Amaçlı örneklem stratejisi, bilgi açısından zengin vakaların derinlemesine incelenmesine olanak tanımışsa da, daha geniş bir vaka havuzuyla modelin genellenebilirliği artırılabilir. Gelecek çalışmalar; sağlık, eğitim, finans ve yargı gibi farklı sektörlerden seçilecek ilave vakalarla modelin kapsamını test edebilir. İkincisi, vakaların seçiminde küresel medyaya yansıma ve kamusal görünürlük kriteri esas alınmıştır. Bu durum, görece daha az görünür olan ancak etik açıdan eşit derecede sorunlu yapay zekâ uygulamalarının (örneğin, işe alım algoritmalarındaki örtük ayrımcılık veya sosyal yardım sistemlerindeki otomatik red kararları) analiz dışında kalmasına yol açmış olabilir. Üçüncüsü, Turner'ın Sosyal Drama modeli, toplulukların krizleri kamusal ritüeller aracılığıyla çözdüğü varsayımına dayanır. Dijital krizlerin bir kısmı (örneğin TikTok ve Kargu-2 vakaları) henüz çözüm aşamasına ulaşmadığı için, modelin “yeniden bütünleşme” aşamasının dijital bağlamda nasıl işlediği konusunda kesin çıkarımlar yapmak şu an için mümkün değildir. Bu vakaların uzunlamasına takip edilmesi, modelin zamansal dinamiklerini anlamak açısından değerli olacaktır. Dördüncüsü, önerilen beş temel değerin evrensellik iddiası, kültürlerarası farklılıklar bağlamında tartışmaya açıktır. UNESCO'nun (2021) küresel etik mutabakat arayışına rağmen, farklı kültürel ve siyasal bağlamlarda bu değerlerin nasıl yorumlanacağı ve önceliklendirileceği sorusu cevapsız kalmaktadır. Gelecek çalışmalar, modelin farklı kültürel bağlamlardaki uygulanabilirliğini karşılaştırmalı olarak incelemelidir. Son olarak, bu çalışma nitel bir vaka analizine dayandığı için, ihlal edilen değerlerin nicel ağırlıklandırması (hangi değerin hangi bağlamda daha kritik olduğu) konusunda kesin yargılarda bulunmak güçtür. Gelecek araştırmalar, çok kriterli karar verme yöntemleri veya uzman görüşüne dayalı Delphi çalışmaları aracılığıyla değerler arası hiyerarşiyi ampirik olarak belirlemeye yönelebilir.

Dijitalleşme sürecinin en ileri evresi olan yapay zekâ çağı, teknolojik yenilikleri durdurmayı değil; bu yenilikleri insanın ontolojik değerleriyle uyumlu hale getirmeyi gerektirmektedir. Bu çalışmanın temel savı nettir: Dijital etik, yeniliğin önünde bir engel ya da soyut bir felsefi egzersiz değil; yapay zekânın tahakküm potansiyeli karşısında insanlığın, adaletin ve toplumsal bütünlüğün korunabilmesi için varoluşsal bir zorunluluktur. Turner'ın Sosyal Drama kuramının dijital krizlere uyarlanmasıyla geliştirilen Değer Temelli Dijital Etik Modeli, bu zorunluluğu yalnızca teorik düzlemde savunmakla kalmamakta; teknoloji geliştiricileri, politika yapıcılar ve uluslararası düzenleyici kurumlar için uygulanabilir, sınanabilir ve proaktif bir yol haritası sunmaktadır.

Araştırma ve Yayın Etiği Beyanı: Makalenin hazırlanmasında bilimsel araştırma ve yayın etiği kurallarına uygun hareket edilmiştir.

KAYNAKÇA

- Anderson, M., & Anderson, S. L. (2007). Machine ethics: Creating an ethical intelligent agent. *AI Magazine*, 28(4), 15-26. <https://doi.org/10.1609/aimag.v28i4.2065>
- Angwin, J., Larson, J., Mattu, S., & Kirchner, L. (2016, 23 Mayıs). *Machine bias*. ProPublica. <https://www.propublica.org/article/machine-bias-risk-assessments-in-criminal-sentencing>
- Ashok, M., Madan, R., Joha, A., & Sivarajah, U. (2021). Ethical framework for artificial intelligence and digital technologies. *International Journal of Information Management*, 62, 102433. <https://doi.org/10.1016/j.ijinfomgt.2021.102433>
- Barker, K., & Ferguson, M. (2022). *Introduction to digital ethics*. Routledge.
- Beauchamp, T. L., & Childress, J. F. (2013). *Principles of biomedical ethics* (7. Baskı). Oxford University Press.
- Benjamin, R. (2019). *Race after technology: Abolitionist tools for the new Jim code*. Polity Press.
- Cabinet Office, Government of Japan. (2016). *The 5th Science and Technology Basic Plan*. <https://www8.cao.go.jp/cstp/english/basic/5thbasicplan.pdf>
- Capurro, R. (2009). Digital ethics. *The International Review of Information Ethics*, 11, 2-5.
- Castells, M. (2000). *The rise of the network society* (2. Baskı). Blackwell Publishers.
- Cath, C., Wachter, S., Mittelstadt, B., Taddeo, M., & Floridi, L. (2018). Artificial intelligence and the 'good society': The US, EU, and UK approach. *Science and Engineering Ethics*, 24(2), 505-528. <https://doi.org/10.1007/s11948-017-9901-7>
- Couldry, N., & Mejias, U. A. (2019). *The costs of connection: How data is colonizing human life and appropriating it for capitalism*. Stanford University Press.
- Creswell, J. W. (2013). *Qualitative inquiry and research design: Choosing among five approaches* (3. Baskı). Sage Publications.
- Danaher, J. (2016). The threat of algocracy: Reality, resistance and accommodation. *Philosophy & Technology*, 29(3), 245-268. <https://doi.org/10.1007/s13347-015-0211-1>

- Deguchi, A., Hirai, C., Matsuoka, H., Nakano, T., Oshima, K., Tai, M., & Tani, S. (2020). What is Society 5.0?. *Society 5.0: A people-centric super-smart society* içinde (s. 1-23). Springer. https://doi.org/10.1007/978-981-15-2989-4_1
- 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>
- Eggers, W. D., & Hickin, R. (2020). *Global Technology Governance Report 2021: Harnessing Fourth Industrial Revolution Technologies in a COVID-19 World*. World Economic Forum.
- European Union. (2024). *Regulation (EU) 2024/1689 of the European Parliament and of the Council laying down harmonised rules on artificial intelligence (Artificial Intelligence Act)*. <https://eur-lex.europa.eu/eli/reg/2024/1689/oj>
- Federal Trade Commission. (2019). *FTC imposes \$5 billion penalty and sweeping new privacy restrictions on Facebook*. <https://www.ftc.gov/news-events/news/press-releases/2019/07/ftc-imposes-5-billion-penalty-sweeping-new-privacy-restrictions-facebook>
- Floridi, L. (2021). The European legislation on AI: A brief analysis of its philosophical approach. *Philosophy & Technology*, 34(2), 215-222. <https://doi.org/10.1007/s13347-021-00460-9>
- Floridi, L., & Cowls, 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>
- Floridi, L., Cowls, J., Beltrametti, M., Chatila, R., Chazerand, P., Dignum, V., Luetge, C., Madelin, R., Pagallo, U., Rossi, F., Schafer, B., Valcke, P., & Vayena, E. (2018). AI4People—An ethical framework for a good AI society: Opportunities, risks, principles, and recommendations. *Minds and Machines*, 28(4), 689-707. <https://doi.org/10.1007/s11023-018-9482-5>
- Gal, U., Jensen, T. B., & Stein, M. K. (2020). Breaking the vicious cycle of algorithmic management: A virtue ethics approach to people analytics. *Information and Organization*, 30(2), 100301. <https://doi.org/10.1016/j.infoandorg.2020.100301>
- Hagendorff, T. (2020). The ethics of AI ethics: An evaluation of guidelines. *Minds and Machines*, 30(1), 99-120. <https://doi.org/10.1007/s11023-020-09517-8>
- Harayama, Y. (2017). Society 5.0: Aiming for a new human-centered society. *Japan Spotlight*, 36(4), 8-13.
- Hill, K. (2020, 24 Haziran). Wrongfully accused by an algorithm. *The New York Times*. <https://www.nytimes.com/2020/06/24/technology/facial-recognition-arrest.html>
- Hoven, J. v. (2015). Ethics for the digital age: Where are the moral spaces. H. Werthner & F. Harmelen (Ed.), *Informatics in the future* içinde (s. 65-77). SpringerOpen.
- Human Rights Watch. (2021). *Stopping killer robots: Country positions on banning fully autonomous weapons and retaining human control*. <https://www.hrw.org/report/2020/08/10/stopping-killer-robots/country-positions-banning-fully-autonomous-weapons-and>
- IEEE. (2019). *Ethically aligned design: A vision for prioritizing human well-being with autonomous and intelligent systems* (1. Baskı). IEEE Standards Association. <https://ethicsinaction.ieee.org>
- Mackert, M. (2020). *Digital ethics orientation, values and attitudes for a digital world*.
- Marchant, G. E., Allenby, B. R., & Herkert, J. R. (Ed.). (2011). *The growing gap between emerging technologies and legal-ethical oversight: The pacing problem*. Springer. <https://doi.org/10.1007/978-94-007-1356-7>
- Marghalani, A., & AlQahtani, M. (2019). Digital ethics: Definition and scope. *International Journal of Advanced Computer Science and Applications*, 10(7), 1-6.
- Masuda, Y. (1980). *The information society as post-industrial society*. World Future Society.
- McLuhan, M. (1994). *Understanding media: The extensions of man*. The MIT Press.

- Mittelstadt, B. D., Allo, P., Taddeo, M., Wachter, S., & Floridi, L. (2016). The ethics of algorithms: Mapping the debate. *Big Data & Society*, 3(2). <https://doi.org/10.1177/2053951716679679>
- Moor, J. H. (1998). Reason, relativity, and responsibility in computer ethics. *Computers and Society*, 28(1), 14-21.
- Moor, J. H. (2006). The nature, importance, and difficulty of machine ethics. *IEEE Intelligent Systems*, 21(4), 18-21. <https://doi.org/10.1109/MIS.2006.80>
- National Transportation Safety Board. (2019). *Collision between vehicle controlled by developmental automated driving system and pedestrian, Tempe, Arizona, March 18, 2018* (Yayın no. NTSB/HAR-19/03). <https://www.nts.gov/investigations/AccidentReports/Reports/HAR1903.pdf>
- Noble, S. U. (2018). *Algorithms of oppression: How search engines reinforce racism*. New York University Press.
- Norwegian Safety Investigation Authority. (2024). *Report on the investigation of the Viking Sky incident*. <https://nsia.no/Marine/Published-reports/2024-05>
- Pasquale, F. (2015). *The black box society: The secret algorithms that control money and information*. Harvard University Press.
- Patton, M. Q. (2015). *Qualitative research and evaluation methods* (4. Baskı). Sage Publications.
- PricewaterhouseCoopers. (2018). *Will robots really steal our jobs? An international analysis of the potential long term impact of automation*. <https://www.pwc.co.uk/economic-services/assets/international-impact-of-automation-feb-2018.pdf>
- Turner, V. (1974). *Dramas, fields, and metaphors: Symbolic action in human society*. Cornell University Press.
- UN Security Council. (2021). *Letter dated 8 March 2021 from the Panel of Experts on Libya established pursuant to resolution 1973 (2011) addressed to the President of the Security Council* (Yayın no. S/2021/229). <https://digitallibrary.un.org/record/3905159?v=pdf>
- UNESCO. (2021). *Recommendation on the ethics of artificial intelligence*. United Nations Educational, Scientific and Cultural Organization. <https://unesdoc.unesco.org/ark:/48223/pf0000381137>
- Wallach, W., & Allen, C. (2009). *Moral machines: Teaching robots right from wrong*. Oxford University Press.
- Yeung, K. (2018). Algorithmic regulation: A critical interrogation. *Regulation & Governance*, 12(4), 505-523. <https://doi.org/10.1111/regg.12158>
- Yin, R. K. (2018). *Case study research and applications: Design and methods* (6. Baskı). Sage Publications.
- Zuboff, S. (2019). *The age of surveillance capitalism: The fight for a human future at the new frontier of power*. PublicAffairs.