

YAPAY ZEKÂYA DAİR ETİK SORUNLAR

Nazan YEŞİLKAYA

Şırnak Üniversitesi, İlahiyat Fakültesi, Felsefe ve Din Bilimleri Bölümü
Şırnak University, Faculty of Theology, Department of Philosophy and Religious Studies,
nkorkunc14@gmail.com, <https://orcid.org/0000-0002-9628-3492>

Article Types / Makale Türü: Research Article / Araştırma Makalesi
Received / Makale Geliş Tarihi:15/10/2022, **Accepted / Kabul Tarihi:** 31/12/2022
<https://doi.org/10.26791/sarkiat.1189864>

YAPAY ZEKÂYA DAİR ETİK SORUNLAR

ÖZ

Bilim kurgunun karanlık vizyonlarını bilim gerçekliğine dönüştürebilen insan-dışı zekânın etik sorunları, akıllı makineler yaygın hâle geldikçe çoğalmakta ve popüler ilgi odağı haline gelmektedir. Yapay zekâ (YZ) algoritmaları ya da düşünen makineler, modern yaşamın ciddi yükümlülüklerini üstlenerek sosyal yaşamın vazgeçilmez bir evresi olmaya doğru yol alırken bir takım etik sorunları da beraberinde getirmektedir. Dönüştürücü teknoloji olarak lanse edilen YZ destekli sistemlerle ilgili tematik sorunların yer aldığı bu çalışmada, YZ'nin sorumluluk alanlarının ve YZ ile ilgili etik kaygı ve sorunların neler olduğunun netleştirilmesi konu edilmiştir. YZ'nin geleneksel ahlak kavramlarımızı, etik yaklaşımlarımızı ve ahlak teorilerimizi yeniden tanımlama potansiyeline sahip olduğuna dikkat çekilen makalede, YZ tasarımında etik ilkelere duyulan ihtiyaç, teknoloji alanındaki güncel araştırmalardan örneklerle ortaya konuyor. Yeni teknolojilerin ve teknolojik süreçlerin tasarım, dağıtım ve kontrolünün makinelere aktarıldığı bir döneme hızla yaklaştığımızı vurgulayan bu çalışmanın amacı, inovasyon hızına bağlı olarak oluşabilecek teknolojik başıboşluğun önüne geçmektir. Makale hem politika yapıcıları hem de yasa koyucuları, YZ'de ortaya çıkan etik sorunları çözmek ve endişeleri hafifletmek için felsefenin rehberliğinde etik değerleri ve ilkelere titizlikle YZ'ye işlemeye çağırılmaktadır.

Anahtar Kelimeler: Felsefe Tarihi, Etik, Yapay Zekâ Etiği, Yapay Zekânın Etik Sorunları, Yapay Zekâya Dair Kaygılar.

ETHICAL ISSUES REGARDING ARTIFICIAL INTELLIGENCE

ABSTRACT

Ethical issues regarding non-humanoid intelligence that can turn the dark visions of science fiction into science fact are multiplying and becoming the focus of popular attention as intelligent machines become almost ubiquitous. Artificial intelligence (AI) algorithms or thinking machines bring with them a number of ethical problems as they take on serious obligations of modern life and move towards becoming an indispensable stage of social life. In this study, which includes thematic problems related to artificial intelligence supported systems, which are launched as transformative technology, it is aimed to clarify the responsibility areas of artificial intelligence and what the ethical concerns and problems about artificial intelligence are. In the article, which draws attention to the fact that artificial intelligence has the potential to redefine our traditional moral concepts, ethical approaches and moral theories, the need for ethical principles in artificial intelligence design is demonstrated with examples from current research in the field of technology. Emphasizing that we are rapidly approaching a period in which the design, distribution and control of new technologies and technological processes are transferred to machines, the aim of this study is to avoid any technological straying that may occur depending on the speed of innovation. The article calls on both policy makers and legislators to meticulously process ethical values and principles into AI under the guidance of philosophy in order to resolve the ethical problems that arise in AI and to alleviate concerns.

Keywords: History of Philosophy, Ethics, Ethics of Artificial Intelligence, Ethical Problems of Artificial Intelligence, Concerns about Artificial Intelligence.

GİRİŞ

Son yıllarda, etten ve çelikten yapılmış transhümanları ve siborgları daha kapsamlı bir ifadeyle post-hümanları düşünüyoruz. İnsanların geçmişte sadece techné yani sanatla üretilebilen makineler olduğu görüşünü geride bırakarak insanların makinelere karışacağı fikrine yöneliyoruz.¹ Yeni teknolojilerin ve teknolojik süreçlerin tasarım, dağıtım ve kontrolünün makinelere devredildiği bir döneme hızla yaklaşıyoruz. George Grant'ın belirttiği gibi, “teknoloji çağın ontolojisi”² söylemi, arkaik dünya görüşünü geride bırakan yeni dijital dönüşüm çağına geçişi belki de en iyi şekilde özetliyor. Hayatımızın birçok yönüne entegre olarak küresel düzeyde önemli değişikliklere neden olan YZ uygulamaları; sağlıktan ulaşım, üretimden güvenliğe, eğitimden eğlenceye kadar birçok sektörü dönüştürmektedir. YZ algoritmaları ya da düşünen makineler, modern yaşamın ciddi yükümlülüklerini üstlenerek sosyal yaşamın vazgeçilmez bir evresi olmaya doğru yol alırken bir takım etik sorunları da beraberinde getirmektedir.

Yeni teknolojilerin günlük hayata entegre edilmesi, ahlaki bütünlükle ilgili sorunların ortaya çıkmasını da beraberinde getirmektedir. Yapay zeka kullanımı yaygınlaştıkça, yapılan tereddütler ve itirazlarla öne çıkan karmaşık etik sorunlar gündeme gelmektedir. Örneğin, YZ ile kendi kendini süren arabalar veya insan müdahalesi olmadan öldürmeye kodlanan otonom silahlar, “katil robotlar”³, YZ uygulamalarında dün olduğu gibi bugün de süregelen tartışmaların konusudur. YZ’de etiğe odaklı tartışmaların çoğu, ahlaki açıdan ele alınması gereken sorunlu hususları kapsamaktadır. Bu meselelerden en belirginini, YZ’ye sahip makinelerin insanlara veya diğer varlıklara zarar verip vermeyeceği yönündeki etik tartışmalardır.

Etik değerlere hâkim insansı robot üretmeye doğru ilerlenen şu günlerde bir diğer gündem maddesi ise düşünen makinelerin “ahlaki statü”leriyle⁴ ilgilidir. YZ, hayatımızın çeşitli yönlerinde karar vermede önem kazanırken insansı olmayan bu yapıların, “etik” olup olmadığı ve adaletin normatif beklentilerine bağlı kalarak insanlara fayda sağlayıp sağlamadığı konusunda inceleme yapılması gerekmektedir. Çünkü makineler ve bilgisayarlar toplumsal rolleri itibarıyla inisiyatif ele alarak daha da karmaşık hâle geliyor. İnsan yetenekleriyle eşleşebilecek veya insan yerine ikame edilebilecek YZ’ye sahip makinelerin ortaya çıkışı, dünyadaki en yüksek ahlaki statüye sahip varlıklar olarak insanların geleneksel benlik duygusuna büyük bir meydan okuma olarak yorumlanabilir.

YZ’nin geleneksel ahlaki kavramlarımızı, etik yaklaşımlarımızı ve ahlak kuramlarımızı yeniden tanımlama potansiyeline sahip olduğuna dikkat çekersek, makinelerin etik güvence olmadan özerk bir şekilde çalışma ihtimali, çoğumuzu rahatsız edecek bir durum olarak kabul edilebilir. Bu nedenle, “YZ, doğrudan insanla ve toplumsal yaşamla ilişkili olması bakımından etik çerçevede incelenmesi” gereken bir konudur.⁵ Bilindiği üzere etik, hayatımızın her alanında verdiğimiz kararları yöneten ahlaki yargıya rehberlik etmeyi amaçlarken değer ise iyi bir yaşamın ne olduğu, insan dışı varlıklar dahil olmak üzere başkalarına ve doğaya ne borçlu olduğumuzu tanımlayan şeydir. Dolayısıyla yüksek verimlilik ve üretkenlik seviyesi vaat eden YZ’nin, insanların daha iyi hayatlar sürdürmesine katkıda bulunması ve refah düzeyini artırması etik bir değer olarak kabul edilebilir.

Bu bağlamda, YZ destekli sistemlerin sunduğu sayısız faydayı vurgulamakta fayda vardır. Analitik yetenekleri göz önüne alındığında YZ; insanların basitçe işleyemedikleri verileri hızlı bir şekilde sonuçlandırabilir ve bağlayabilir, insanlardan daha tutarlı olup onları sıkıcı veya tekrarlayan görevlerden kurtardığı için çoğu zaman hayatı kolaylaştırır. Öte yandan, son yıllarda bilimsel literatür ve medyanın gündeminde yer alan insan zekâsının dışsallaştırılması olarak nitelendirilen YZ’nin bazen kasıtsız olarak önyargıları yaydığı, sorumluluktan kaçındığı ve ayrımcılık yaptığı yönünde eleştiriler de hüküm sürdürmektedir. YZ’nin rastgele bir kullanıcı grubunu bir bilgisayar sistemindeki diğerlerine göre ayrıcalıklı kıldığı, haksız sonuç-

1 Bkz. Nazan Yeşilkaya, “Yeni Bir Varlık Türü Olarak Biyoteknolojik Varlıklar ve Adalet Sorunu”, *Şırnak Üniversitesi İlahiyat Fakültesi Dergisi* Sayı 26 (2021), 233-256.

2 George Grant, *Technology and Justice* (Concord, Ontario: House of Anansi, 1986): 32.

3 Bkz. Richard G. Epstein, “The Case of The Killer Robot (part 2),” *Acm Sigcas Computers and Society* 24.4 (1994): 12-32; Robert Sparrow, “Killer Robots”, *Journal of Applied Philosophy*, 24(1), (2007) 62-77.

4 Bkz. John P. Sullins, “When Is a Robot a Moral Agent?”, *International Review of Information Ethics* Vol. 6 (12/2006), 23-30.

5 Arslan Topakkaya – Yağmur Eyibaş, “Yapay Zekâ ve Etik İlişkisi”, *Felsefe Dünyası*, Sayı: 70, 81-99 (2019), 86.

lar ürettiği, sistematik ve tekrarlanabilir hataları tespit edip sürdürdüğü, algoritmik önyargı ve adaletsizliği artırdığı akademik dünyada yaygın olarak tartışılmaktadır.

Bu noktada belirtilmelidir ki teknolojik gelişmeler ne kadar karmaşıkça ortaya çıkacak etik sorunlar da o kadar karmaşıklaşır. Büyük gücün büyük sorumluluk getirdiği düşünüldüğünde, büyük bir güç olarak lanse edilen YZ'ye sahip makinelerin insanlara veya diğer varlıklara zarar verip vermeyeceği ve sorumluluk konusu önemli etik tartışmalar arasındadır. Örneğin insan müdahalesi olmadan öldürebilen “ölümcül otonom silah sistemleri (LAWS)”⁶, “otonom arabalar”⁷ ve “öngörücü polislik”⁸ gibi gelişmelerle öne çıkan YZ uygulamaları, karmaşık etik sorunların merkezinde yer almaktadır. Bu nedenle YZ'nin, insanlarla güven ve adalet içinde çalışması ancak şeffaflık, sorumluluk ve hesap verebilirlik içeren etik bir YZ'nin inşa edilmesi ile mümkün kılınabilir.

Yukarıda belirtilen faktörler nedeniyle, insan-dışı zeki yapıların sorumluluk alanlarının felsefi ve etik analizini gerektiren tematik sorunları ve etik kaygıları değerlendirmek ve inovasyon hızını göz önünde bulundurarak herhangi bir teknolojik başıboşluktan kaçınmak için YZ'de etiğin kaçınılmazlığını vurgulamak önemlidir. İnsanlığın ve hatta gezegenin ortak yararını hedef alarak, dönüştürücü teknoloji olarak lanse edilen YZ destekli sistemlerin inşa edilebilmesi için etik sorunların ve etik kaygıların şimdiden açıklığa kavuşturulması gerekliliği bu makalenin konusudur. YZ'nin geleneksel ahlaki kavramlarımızı, etik yaklaşımlarımızı ve ahlak kuramlarımızı yeniden tanımlama potansiyeline sahip olduğuna dikkat çeken bu çalışmada, YZ'nin gelişimi ve kullanımından kaynaklanan etik sorunlar felsefi bir bakış açısıyla ele alınmaktadır. Ayrıca YZ etiğinin önemini, YZ'de etik ilkelere duyulan ihtiyacı ve kaygıları kapsayarak, alandaki güncel araştırmalardan örneklere yer vermektedir.

1. ETİK YAPAY ZEKÂYI TANIMLAMA

Yunanca “alışkanlık, gelenek, örf, adet, karakter” anlamına gelen “ethos” kelimesinden türetilen etik, felsefenin “doğru ve yanlış davranış kavramlarını sistematikleştirmeyi, savunmayı ve önermeyi içeren” bir dalıdır.⁹ Ahlaki yargılarımızın rasyonel gerekçesini inceleyen etik; ahlaki olarak doğru ya da yanlış, haklı ya da haksız olanı inceleme konusu yapar. İnsanların doğa ve diğer insanlarla etkileşimlerinde özgürlük, sorumluluk ve adaletin bir yansıması olan etiğin genel olarak insanlarla dünya arasındaki ilişkiye odaklandığını belirtebiliriz.

İnsan eşdeğeri veya üstün performansla sahip mevcut YZ algoritmaları, yalnızca tek ve sınırlı bir alanda özel olarak programlanmış bir yetkinlik ile karakterize edilir. YZ sistemlerinin insan onurunu koruması ve hiçbir şekilde insana zarar vermemesi, “etik YZ” olarak tanımlanmaktadır ve konusu bakımından uygulamalı etik alanına girmektedir. Teknoloji felsefesinin de en önemli konularından biri olan etik YZ; “bi-reysel haklar, mahremiyet, ayrımcılık yapmama ve manipülasyon yapmama gibi temel değerlerle ilgili iyi tanımlanmış etik yönergelerle bağlı kalan YZ'dir.”¹⁰ YZ etiği, YZ teknolojisinin geliştirilmesini ve sorumlu kullanımını bilgilendirmeyi amaçlayan bir ahlaki ilkeler ve teknikler sistemidir.¹¹

Etik YZ, YZ'nin meşru ve gayri meşru kullanımını belirlemede etik hususları önceler. Etik YZ uygulayan kuruluşlar, bu yönergelerle uyulmasını sağlamak için açıkça belirtilen politikalara ve iyi tanımlanmış inceleme süreçlerine sahiptir. Uygulamalı etik çatısı altında konumlanan tekno-etik olarak da bilinen teknoloji etiği, gelişen teknolojilerle bağlantılı olarak teknoloji çağının etik sorunlarını ele almaktadır. Yeni

6 LAWS (Lethal autonomous weapons) olarak anılan “ölümcül otonom silah sistemleri”, barut ve nükleer silahların ardından savaşta “Üçüncü Devrim” olarak lanse edilmektedir. Bkz. Rebecca Crotoof, “The Killer Robots Are Here: Legal and Policy Implications”, *36 Cardozo L. Rev.*: 1837 (2015).

7 Bk. Sven Nyholm, “The Ethics of Crashes with Self-Driving Cars: A Roadmap, I”, *Philosophy Compass* Vol. 13 (2018), 2-10.

8 R. Richardson vd., “Dirty Data, Bad Predictions: How Civil Rights Violations Impact Police Data, Predictive Policing Systems, and Justice”, *NY Univ Law Rev.* Online 192, (2019) <https://ssrn.com/abstract=3333423>. Erişim 10 Eylül 2022.

9 Ethics, Britannica, <https://www.britannica.com/topic/ethics-philosophy> Erişim: 20 Kasım 2022.

10 Ethical AI, *Glossary*, <https://c3.ai/glossary/artificial-intelligence/ethical-ai/> Erişim: 10 Mayıs 2022.

11 George Lawton, “AI Ethics (AI Code of Ethics)”, TechTarget, <https://www.techtarget.com/whatis/definition/AI-code-of-et-hics> Erişim: 12 Mayıs 2022.

teknolojileri kullanılmanın doğuracağı etik sorunlar nelerdir? İnsan yaşamını hem iyileştiren hem de bozan YZ teknolojilerinin dönüştüren ve geliştiren gücü, toplumsal alanda derin etkilere neden olurken bu durum, ilkeler ve değerler bakımından ne tür tartışmalara yol açmaktadır? YZ, sorumluluk alabilir mi? Etik bir YZ olası mıdır? YZ, empati kurabilir mi? YZ, özgürce karar alabilir mi? Robotların hakları olmalı mı? Akıllı makineler, insan zekâsından daha iyi düşünebilir mi? Karar vermede insan yargısının bazı unsurları vazgeçilmez midir? Bu ve benzeri sorular, uygulamalı etiğin cevabını aradığı konular arasında yer almaktadır.

Öz farkındalığı olmayan ve dolayısıyla etiğin temeli olan “empati” denilen şeye sahip olmayan YZ’nin, nasıl etik olabileceği merak konusudur.¹² YZ’nin var olan tek ahlaki pusulası, doğru ve yanlış olanın çıtasını belirleyen geliştiricisine aittir. Tasarımcısının yüksek ya da düşük bir ahlaki pusulaya sahip olması YZ’ya aynı paralelde nitelikler yüklemesiyle sonuçlanabilir. YZ’de etik açıdan sorumlu inovasyonun geliştirilmesinde YZ’nin “etik” olması gerektiği konusunda açık bir fikir birliği olsa da hem etik YZ’yi neyin oluşturduğu hem de bunun gerçekleşmesi için hangi etik gerekliliklerin, teknik standartların ve en iyi uygulamaların gerekli olduğu konusunda tartışma sona ermemiştir.¹³ YZ için etik bir çerçeve oluşturmak, YZ destekli sistemlerin ve araçların risklerini, faydalarını aydınlatmak ve sorumlu kullanımları için kılavuzlar oluşturmak önemlidir.

YZ’yi sorumlu bir şekilde kullanmak, bizi insan yapan şeyin ne olduğu sorusunu incelemeyi gerektirir. Stahl’in de belirttiği üzere etiği anlamanın formülü insanlık durumunu anlamaktır. Etik ilkeler geliştirir ve yaşamımızda uygularız. Çünkü bizler, bizimkine benzer korkuları ve umutları olan başkalarıyla empati kurabilen bedensel, savunmasız, ölümlü varlıklarız. Bu özelliğimiz sosyal doğamızın ve dolayısıyla ahlakımızın temelidir. YZ’nin ahlaki olarak sorumlu ve etik bir temsilci olması için bu özellikleri paylaşması gerekir. Şu an için hiçbir yapay sistem empati kurma özelliğine haiz değildir. İnsana dair bu özellik YZ’nin insanı aşan ve bir süredir başarmış olduğu hesaplama yeteneğiyle ilgili olmayıp YZ’nin bizimle aynı kategoride olmadığı gerçeğiyle alakalıdır.¹⁴

2. YAPAY ZEKÂDA ETİĞİN ÖNEMİ

YZ için etiğin neden gerekli olduğu üzerinde çalışan YZ araştırmacılarının ve düşünürlerinin gündemlerinde öne çıkan konulardan biri otonom araçlara etik bir boyut katmaktır. YZ’ya insana has zihinsel özelliklerin merceğinden bakmak, onu insan zihninin bir tür kopyasına indirgemek, YZ’nin ortaya çıkardığı sorunların kusurlu ve nihayetinde sınırlı bir etik analizine yol açar. YZ’de etiğin önemi, özellikle birtakım sorularla daha görünür olabilir: Herhangi bir YZ sisteminde yanlış giden bir durum söz konusu olduğunda sonuçlarından kim ya da kimler sorumlu olacaktır? Olumsuz bir sonuçla karşılaştığımız zaman, YZ’yi suçlamayı mı tercih edeceğiz yoksa bu durumdan sorumlu şirketleri denetlemeyen bürokratları ya da bürokrasiyi mi suçlayacağız? Bir YZ sistemi, kendisine verilen görevde başarısız olduğunda suçu kim üstlenir, YZ’yi tasarlayan veya kodlayan programcılar mı yoksa son kullanıcılar mı sorumlu tutulacaktır?¹⁵ Kısacası YZ sistemleri, insanların karar alma sürecine dahil edilecek mi, insan yerine kendileri mi karar verecek yoksa melez bir yapı olarak mı değerlendirilecek? Bu minvaldeki sorular, insan-dışı zeki yapıların sorumluluk alanlarının etik analizini gerektiren tematik sorunlar olarak karşımıza çıkmaktadır.

İnsanoğlu doğru ile yanlış ayırt edebilme ve empati kurabilme yeteneğine sahipken YZ veya otonom sistemler gibi algoritmik unsurlarda benzeri nitelikler söz konusu değildir. Bu sistemlerin sadece doğru ve yanlış etiketlerle kodlanmış olması bir takım etik yarıklara neden olmaktadır.¹⁶ Teknoloji şirketlerinin kendi çıkarları doğrultusunda etik ilkeler uyguladığı otonom sistemlerin algoritmik önyargılarla verdiği kararlar

12 M. Asada, “Towards Artificial Empathy”, *Int J of Soc Robotics* 7, (2015), 19–33. Ayrıca bkz. Ö.N. Yalçın - S. DiPaola, “Modeling Empathy: Building a Link Between Affective and Cognitive Processes”, *Artificial Intelligence Review* 53, (2020), 2983-3006.

13 A. Jobin vd., “The Global Landscape of AI Ethics Guidelines”, *Nat Mach Intell* 1, 2019, 389.

14 Bernd Carsten Stahl, “Ethical Issues of AI”, *Artificial Intelligence for a Better Future*, 2021, 35- 53.

15 Modern bürokratlar sıklıkla, sorumluluğu o kadar geniş bir şekilde dağıtan yerleşik prosedürlere sığınır ki sonuçta ortaya çıkan felaketlerden sorumlu olacak hiç kimse tespit edilemez. Bkz. Bostrom, 709.

16 Bkz. Paula Boddington, *Towards a Code of Ethics for Artificial Intelligence*, Heidelberg: Springer, 2017.

bu duruma bir örnektir.

Kısacası teknoloji şirketlerinin etik kriterlerini hangi değerlere dayandırdığı, doğru ya da yanlışın ne olduğuna dair tespitleri ve bu tespitlere dayalı olarak oluşturdukları kurallar tartışmaya ihtiyaç duymaktadır. YZ sistemlerinin araştırılması, geliştirilmesi ve uygulanmasındaki ilerlemeler, YZ etiği üzerine geniş ve kapsamlı bir söylemi doğurmaktadır. Bu nedenle, yenilikçi uygulamaların hızı göz önüne alındığında, herhangi bir teknolojik başıboşluğun yaşanmaması için YZ’de etiğin önemi göz ardı edilmemelidir.

Sonuç olarak geliştirilmekte olan belirli bir YZ donanımlı araçla ilgili olarak meydana gelen veya gelecekte oluşabilecek herhangi bir etik hatadan kimin sorumlu tutulacağı konusu tartışılması gereken ciddi bir konudur. Rosalind Picard’ın da iddia ettiği üzere, “bir makinenin özgürlüğü ne kadar büyükse, ahlaki standartlara da o kadar çok ihtiyaç duyulur.”¹⁷ Bu bağlamda YZ’de ahlaki bir ilkenin oluşturulması endüstrinin ve politika yapımcıların önemli sorunları arasında yer almalıdır. Olumsuz ya da kriminal bir durum baş gösterdiğinde, etik YZ politikalarının yasal zeminde nasıl bir karşılığının olacağı, mevcut yasal düzenlemelerin ihtiyaca cevap verip vermediği ya da ne şekilde yasal düzenlemelere ihtiyaç duyulduğu hususu YZ’nin etik olmayan kullanımını tespit etmek açısından önemlidir. Bu nedenle YZ’nin kötü aktörler tarafından etik dışı kullanımını önlemek için özellikle de politika yapımcılar tarafından bazı planlar yapılmalı ve tedbirler alınmalıdır.

YZ algoritmalarının ya da herhangi bir otonom sistemin ulusal ve uluslararası denetimlerde sergileyeceği şeffaflığın derecesi etik değerlerin YZ ile ne derecede bütünlük arz ettiğiyle yakından ilintilidir. Nitekim Bostrom’a göre yalnızca güçlü ve ölçeklenebilir değil aynı zamanda denetime karşı şeffaf olan YZ algoritmaları geliştirmek giderek daha önemli hâle gelecektir.¹⁸ YZ algoritmalarının manipülasyona ve dezenformasyona açık sorunlarının göz ardı edilmesi yerel ve küresel boyutta birtakım etik sorunlar doğurmaktadır. Düşünen makinelerin tasarımında ne tür bir programlama kullanıldığı çok önemli bir kriterdir. Bu noktada, makineler için etik kriterler geliştirmenin yanı sıra makine kullanan insanlar için de etik değerler ve ilkeler geliştirmenin önemi açıktır.

YZ’de etiği vazgeçilmez kılan bir diğer husus, YZ’nin şeffaf, sorumlu ve hesap verebilir bir forma sokulmasıdır. Bostrom’a göre YZ’nin tek istenen özelliği şeffaflık değildir. Sosyal işlevleri devralan YZ algoritmalarının, yönettikleri kişiler tarafından öngörülebilir olması da önemlidir.¹⁹ Bu ise YZ kullanımının yasalar, düzenlemeler ve normlar temelinde yeniden değerlendirilmesini gerektirmektedir.

Etik YZ stratejileri oluşturma çabaları, önyargılı veri veya algoritmaların kullanımına karşı insanlar için güvenli bir ortam sağlamayı mümkün kılmaktadır. Anderson’a göre otonom makinelerin kendi başlarına doğru etik kararlar verebileceklerine güvenmek isteriz ve bu da makineler için bir etik oluşturmamızı gerektirir.²⁰ Ancak yine de etik bir YZ uygulamasının istenmeyen ve öngörülemeyen olumsuz durumlara neden olabileceği ihtimali bir makineye güvenmeyi zorlaştırmaktadır.

3. ETİK YAPAY ZEKÂYA DAİR BEKLENTİLER

Etik bir YZ sisteminden kuralları ve ilkeleri açıklayabilmesi ve uygun davranışa rehberlik edecek motivasyonu sağlayabilmesi beklenir. Ayrıca etik YZ kapsayıcı, açıklanabilir, olumlu bir amaca sahip ve verileri sorumlu bir şekilde kullanabilir olmalıdır. Bu bağlamda YZ etiği, YZ teknolojisinin geliştirilmesini ve sorumlu kullanımını bilgilendirmeyi amaçlayan bir ahlaki ilkeler ve teknikler sistemidir.²¹ İnsan haklarına saygı göstermek, insani değerlerle uyum sağlamak ve mümkün olduğunca insanı güçlendirirerek bütünsel refahı teşvik etmek gayesiyle tasarlandıkları sürece YZ sistemleri toplumda iyiliğin sürdürülmesi için muazzam bir güç olabilir. Değerler hususu, YZ’de etiği kaçınılmaz kılan ve farkında olsak da olmasak da

17 Rosalind Picard, *Affective Computing* (Cambridge, MA and London: MIT Press: 1997), 19.

18 Nick Bostrom ve Eliezer Yudkowsky, *Cambridge Handbook of Artificial Intelligence*, “The Ethics of Artificial Intelligence”, eds. William Ramsey and Keith Frankish, Cambridge University Press, 2011, s. 919.

19 Bostrom, “The Ethics of Artificial Intelligence”, 920

20 Michael Anderson, *Machine Ethics*, (Cambridge: Cambridge University Press, 2011), 1.

21 Lawton, “AI Ethics (AI Code of Ethics)”, 1.

seçimlerimizin dayanağı ile yakından ilintili olup çevremizi ve doğal kaynaklarımızı korumak için tasarlanmalıdır. Bu değerler, politika yapıcılara olduğu kadar mühendislere, tasarımcılara ve geliştiricilere de rehberlik etmelidir. YZ sistemlerindeki ilerlemeler yalnızca küçük gruplara, tek bir ulusa veya tek bir şirkete fayda sağlamaktan ziyade tüm insanların hizmetinde olmalıdır.²²

Bu gayeye bağlı olarak son yıllarda, bir dizi etik yönerge yayınlandı. Bu yönergeler, yeni YZ teknolojilerinin “yıkıcı” potansiyellerinden yararlanmayı amaçlayan normatif ilkeler ve önerileri içermektedir.²³ Avrupa Komisyonu, 8 Nisan 2019’da “Güvenilir Yapay Zekâ Sistemleri İçin Etik İlkeler Rehberi”ni yayınlarak güvenilir YZ sisteminin; yasal, etik ve teknik bakımdan sağlam ve güçlü olması gerektiği kararına vardı. Benzer şekilde UNESCO’ya üye ülkeler, YZ etiği ile ilgili ilk küresel anlaşmayı 24 Kasım 2021’de kabul etti. Anlaşma, insan haklarını ve insan onurunu sadece korumakla kalmayacak aynı zamanda teşvik edecek ve dijital dünyada, hukukun üstünlüğüne güçlü bir saygının inşa edilmesini sağlayan etik bir yol gösterici pusula ve küresel bir normatif ana kaya olacaktır.

Bahsi edilen yönergelerin yanı sıra son yıllarda AB, topluma büyük ölçüde zarar verebilecek sorumsuz kullanımı önlemek için YZ’yi çevreleyen yasalar çıkararak onu güvenli, şeffaf, etik ve tarafsız hâle getirmek için bir komisyon oluşturdu.²⁴ Bu tür sorumlu YZ çalışmalarına devam etmek, teknolojinin nasıl kullanıldığına şeffaflık ve açıklık getirmek açısından oldukça önemlidir. YZ’nin verdiği kararların kullanıcılar üzerinde geri dönüşü olmayan ve yaşamı değiştiren bir etkisi olabileceği göz önüne alındığında YZ’nin daha demokratik ve sorumlu hâle getirilmesi “etik YZ” oluşturmak için önemli bir adım olacaktır.

Bir başka nokta da düşünen makinelerin insan topluluğuna yarı-özerk failer olarak girmesi halinde bu failerin, toplumun sosyal ve ahlaki normlarını takip etmeleri yönündeki beklentidir. Normların bu tür yarı özerk sistemlere yerleştirilmesi, konuşlandırılacakları topluluğun net bir tanımını gerektirir. Ayrıca, belirli bir topluluk içinde bile farklı teknik düzenlemeler farklı normlar gerektirecektir. İlk adım, sistemlerin içinde konuşlandırılacağı belirli topluluğun normlarını ve özellikle gerçekleştirmek üzere tasarlandıkları görev türleriyle ilgili normları belirlemektir.²⁵ Görüldüğü üzere çok kültürlülük nedeniyle neyin iyi olduğu veya neden iyi olarak kabul edilmesi gerektiği konusunda yani etik uzlaşmaların nasıl sağlanacağı hususunda çoğu zaman pek bir fikir birliği sağlanamayabilir. Çoğulcu bir dünyada kimin normlarının hâkim olacağına ilişkin çatışmalarla karşı karşıyayız ve bu aşılması çok zor bir sorun olmaya devam ediyor. Günümüz dijital dünyasında böyle bir “etik fikir birliği”²⁶ küresel düzeyde bulunmamaktadır. Oysa dijital teknolojiler ve veri akışları zaten yoğun bir şekilde küresel olduğu için fikirbirliğine acilen ihtiyacımız söz konusudur.

Bu tür anlaşmazlıkların yanı sıra, etik YZ’de yardımseverlik, güvenlik, başarı ve öz yönetim gibi paylaşılan etik değerleri tanımlamak için çok sayıda girişimde bulunulmuştur. YZ’ye dair etik beklentiler arasında en dikkat çeken hususlar kullanıcı dostu, şeffaf, açıklanabilir, adil ve tarafsız olmasının yanı sıra insanların daha iyi hayatlar yaşamalarına olanak sağlamaya yönelik sundukları refah vaatleridir.

YZ sistemleri için ilk etik kod, *Runaround*’ta²⁷ “Üç Robot Yasası”nı (The Three Laws of Robotics) sunan ünlü bilim kurgu yazarı Isaac Asimov tarafından tanıtıldı. Otonom YZ’nin potansiyel tehlikelerini gelişmelerinden çok önce öngören Asimov, bu riskleri sınırlamanın bir yolu olarak bu üç yasayı daha sonra *Robots and*

22 Bkz: IEEE, *Ethically Aligned Design- A Vision for Prioritizing Human Well-being with Autonomous and Intelligent Systems*, (USA: IEEE, 2019), 10.

23 Bkz. Thilo Hagendorf, *The Ethics of AI Ethics: An Evaluation of Guidelines*, *Minds and Machines* 30: (2020), 99. Yarı sistematik bir değerlendirme olarak tasarlanan bu makale, örtüşmelerin yanı sıra ihmalleri de vurgulayarak 22 kılavuzu analiz edip karşılaştırmaktadır.

24 Sray Agarwal- Shashin Mishra, *Responsible AI Implementing Ethical and Unbiased Algorithms*, (Switzerland: Springer, 2021), 175.

25 IEEE, *Ethically Aligned Design- A Vision for Prioritizing Human Well-being with Autonomous and Intelligent Systems*, (USA: IEEE, 2019), 6.

26 İnsan hakları savunucuları genellikle etik çerçevelerin kullanımını evrensellikten yoksun olmakla eleştirir. Bk. Sakiko Fukuda-Parr ve Elizabeth Gibbons, “Emerging Consensus on ‘Ethical AI’: Human Rights Critique of Stakeholder Guidelines”, *Global Policy* Vol. 12 (2021), 32 – 44.

27 Bkz. I. Asimov, *Runaround: A Short Story*, (New York: Street and Smith, 1942).

Empire'da²⁸ “Robotiğin Sıfırncı Yasası” olarak adlandırılan dördüncü bir yasayla destekledi. Bu dört yasa şöyle formüle edilir: (1.) Bir robot bir insana zarar veremez veya hareketsiz kalarak bir insanın zarar görmesine izin veremez. (2.) Bir robot, birinci yasayla çelişmediği sürece, insanlar tarafından verilen emirlere uymak zorundadır. (3.) Bir robot, birinci veya ikinci yasa ile çelişmediği sürece kendi varlığını korumalıdır. (4.) Bir robot insanlığa zarar veremez veya hareketsiz kalarak insanlığın zarar görmesine izin veremez.

Asimov hikayelerinde bulunan robotik yasasının bir alternatifi olan “sorumlu robotiğin üç alternatif yasası”nı ana hatlarıyla belirten ilk bilimsel çalışma 2009 yılında Murphy ve Woods tarafından yapılmıştır. Buna göre sorumlu robotiğin alternatif üç kanunu şu şekildedir: (1.) Bir insan, en yüksek yasal ve profesyonel güvenlik ve etik standartlarını karşılayan insan-robot çalışma sistemi olmadan bir robotu konuşlandıramaz. (2.) Bir robot, bir insana rollerine uygun şekilde yanıt vermelidir. (3.) Bir robot, birinci ve ikinci yasalara uygun olarak kontrolün diğer ajanlara sorunsuz transferini sağladığı sürece, kendi varlığını korumak için yeterli konumlanmış özerkliğe sahip olmalıdır.²⁹ Gerek Asimov’un öngörülerıyla şekillenen yasalar gerekse Murphy ve Woods’un alternatif yasaları, etik kaygılardan beslenerek ortaya çıkan bu endişeleri tanımlamak ve olası risklerin önüne geçmek amacıyla kaleme alınmaları bakımından etik YZ’nin oluşumunda önemli bir yer teşkil etmektedir.

Buradan da anlaşılacağı üzere, etik YZ ile ilgili beklentiler çoğunlukla etik standartlarımızı benimseyecek insan dostu YZ sistemleri oluşturmak gibi niteliklerle donatılmış özerk yapıların olması gerektiği konusunda ortak bir paydada birleşiyor. Ayrıca çevreye duyarlı, güvenilir ve etik bir YZ inşa etmeye dair beklentiler arasında sosyal sorumluluğu gözetmesi, insanları ve ortak iyiliği faydalı bir şekilde koruması ve hizmet etme amacına cevap vermesi yer almaktadır.

Sonuç olarak belirtilmelidir ki toplumsal sorumluluğu, insanları ve ortak iyiyi yararlı bir şekilde korumaya ve hizmet etmeye yönelik bir amaca cevap vermesi beklenen YZ’nin; çevreye duyarlı, güvenilir ve etik bir form üzerine inşa edilmesi YZ’den beklenen hedefler arasındadır. Avrupa Birliği’nin Yapay Zekâ Konusunda Üst Düzey Uzman Grubu’nun savunduğu gibi, “YZ kendi başına bir amaç değil insani gelişmeyi artırmak, böylece bireysel ve sosyal refahı ve ortak iyiliği geliştirmek, ilerleme ve yenilik getirmek için umut verici bir araçtır.”³⁰

4. YAPAY ZEKÂYA DAİR ETİK KAYGILAR

YZ’nin yaşamımızdaki rolünün artması, insan toplumlarını yeni teknolojilerden kaynaklanan tehditlere maruz kalma konusunda endişeye sevk etmektedir. YZ etiğinin, YZ sistemleri ve bunlar üzerinde çalışan insanlar tarafından nasıl algılandığını daha iyi anlamak açısından “SHERPA”³¹ önemlidir. Bu projede amaçlanan literatür incelemelerinin ötesine geçerek insanların YZ’nin etik sorunlarından bahsederken akıllarında ne olduğunu ampirik olarak bulmaktır. Bunun yanı sıra, YZ söz konusu olduğunda, hepimizin çok dikkatli bir şekilde düşünmesi gereken temel etik zorluklardan bazılarını vaka çalışmaları ve Delphi çalışmasıyla tespit eden Stahl, *Ethical Issues of AI* (Yapay Zekâda Etik Sorunlar) adlı çalışmasında YZ’ye dair tüm etik sorunları toplam 39 kategoride sıralanmaktadır:

“(1.) İnovasyon maliyeti, (2.) Fiziksel bütünlüğe zarar, (3.) Kamu hizmetlerine erişim eksikliği, (4.) Güven eksikliği, (5.) Yapay zekânın “Uyanışı”, (6.) Güvenlik sorunları, (7.) Kaliteli veri eksikliği, (8.) İşlerin ortadan kalkması, (9.) Güç asimetrisi, (10.) Sağlık üzerinde olumsuz etki, (11.) Bütünlük sorunları, (12.) Verilerin doğruluğu eksikliği, (13.) Gizlilik eksikliği, (14.) Şeffaflık eksikliği, (15.) Askeri kullanım potansiyeli, (16.) Bilgilendirilmiş onay eksikliği, (17.) Önyargı ve ayrımcılık, (18.) Adaletsizlik, (19.) Eşitsiz güç ilişkileri, (20.) Kişisel verilerin kötüye kullanılması, (21.) Adalet sistemi üzerinde olumsuz etki, (22.) Demokrasi üzerinde olumsuz etki, (23.) Suç ve kötü niyetli kullanım potansiyeli, (24.) Özgürlük ve bireysel

28 Bkz. I. Asimov, *Robots and Empire: The Classic Robot Novel*, (New York: HarperCollins, 1986).

29 R. Murphy – D. Woods, Beyond Asimov: The three laws of responsible robotics, *IEEE Intelligent Systems*, 24(4), 14–20, 2009, 19.

30 European Commission, *High-Level Expert Group on Artificial Intelligence, Ethics Guidelines for Trustworthy AI*, Brussels: 2012. <https://www.aepd.es/sites/default/files/2019-12/ai-ethics-guidelines.pdf>. Erişim: 08 Haziran 2022.

31 Bkz SHERPA, <https://www.project-sherpa.eu/> Erişim: 10 Haziran 2022.

özerklik kaybı, (25.) Tartışmalı veri sahipliği, (26.) İnsan temasının azaltılması, (27.) Veri ve sistemlerin kontrol ve kullanım sorunları, (28.) Tahmine dayalı önerilerin doğruluğunun olmaması, (29.) Bireysel olmayan önerilerin doğruluğunun olmaması, (30.) Ekonomik gücün yoğunlaşması, (31.) Tedarik zincirinde temel insan haklarının ihlali, (32.) Son kullanıcıların temel insan haklarının ihlali, (33.) İstenmeyen, öngörülemeyen olumsuz etkiler, (34.) “Yanlış” sorunların önceliklendirilmesi, (35.) Hassas gruplar üzerinde olumsuz etki, (36.) Sorumluluk ve sorumluluk eksikliği, (37.) Çevre üzerinde olumsuz etki, (38.) Karar vermede insan faktörünün kaybı, (39.) Bilgiye erişim ve bilgi özgürlüğü eksikliği.”³²

Genel olarak ön plana çıkan etik zorluklar yukarıda yer verildiği şekilde sıralansa da birincil ve sıklıkla dile getirilen etik kaygı, gizlilik ve veri koruma sorununa ilişkin kaygıdır. Bu noktada gizlilik ve veri korumanın aynı şey olmadığını belirtmek gerekir.³³ Ancak YZ etiği amaçları doğrultusunda, temel gizlilik endişesi bilgi gizliliğidir ve veri koruma, bilgi gizliliğini korumanın bir yolu olarak anlaşılabilir.³⁴ YZ'nin büyüleyici ve bir o kadar da korkutucu dünyasında, mahremiyet ve gözetim hususu da etik meselelerle yakından ilgili bir konu olarak belirmektedir. Özellikle de yaygın gözetimin YZ ile zirveye tırmanması, insanların ve toplumların birçok sahada gözetlenmesine neden olmaktadır. Şirketler ve hükümetler tarafından YZ sistemleri aracılığıyla insanların ne yaptığının sürekli olarak takip edilmesi mahremiyet alanlarının aşınmasına sebebiyet vermektedir. YZ'nin etkin olduğu yakın gelecekte beklenen aşırı gözetim, mahremiyetin sonu olarak yorumlanırken bu durum aynı zamanda etik anlamda da endişe uyandırmaktadır.

İnsanların izinleri olmaksızın “Büyük Birader” tarafından devamlı izlenmesi, özellikle de yüz tanıma teknolojilerinin gelişimiyle stadyumlarda, parklarda ve kamusal alanlarda geniş kalabalıklar içinde kimliklerinin tespiti daha da kolay hâle gelmektedir. Bu durum, herkesin yaşamı üzerinde yalnızca birkaç şirketin ya da hükümetin bilgi ve kontrole sahip olabileceği varsayımı dikkate alındığında daha da rahatsız edici bir hal almaktadır.

Sanayi devriminden sonra daha da otomatik hâle gelen toplumun maruz kaldığı sorunlardan bir diğeri de sahte medya ve dezenformasyonun yükselişiyle YZ'nin kötü amaçlara alet edebilmesidir. Nitekim yenilikçi uygulamaların her geçen gün daha da güçlenmeye devam etmesi YZ'nin birey ve toplum yararından ziyade bir takım grupların çıkarları için araç olarak kullanılmasını kaçınılmaz kılacaktır. Bu ise propaganda, dezenformasyon, kötü niyetli müdahale ve şantaj gibi ciddi saldırılar şeklinde kendini gösterecektir. Ezcümle bunlar, YZ bazlı yeni teknolojilerin güvenlik açığına işaret etmektedir.

Özellikle belirtmelidir ki YZ'ye sahip olamayacağı nitelikler vererek onu somut kılmamaya dikkat edilmelidir. Erdemleri insanlar yerine YZ'ye uygulamaktan kaçınmak önemlidir. “Güvenilir yapay zekâ” ve “sorumlu yapay zekâ” bu yönelimin iki örneğidir. Güven nesnelere değil, kişilerarası bir ilişkidir. Güven duygusu YZ'yi tasarlayan, programlayan, geliştiren ve kullanan insanlara mahsus bir yetidir dolayısıyla YZ teknolojik formundan ötürü güven konusu yapılmamalıdır. YZ için Avrupa etik yönergeleri üzerinde çalışan komisyonun uzman grubunun bir üyesi olan Alman filozof Profesör Thomas Metzinger'in iddia ettiği gibi, güvenilir YZ hikâyesi bir pazarlama anlatısıdır. Potansiyel yaptırımlarla ilgili sorumluluk, YZ'ye ait olmayıp temsilcisinin niyetine göre değerlendirilir. YZ Kantçı anlamda otonom hâle gelmedikçe yani bir tür iradeyi ifade edemedikçe sorumlu tutulamaz. Ancak onu tasarlayan, programlayan, geliştiren ve kullanan insanlar sorumlu tutulmalıdır.³⁵

YZ ile ilgili çokça atıfta bulunulan diğer endişe konusu ise “önyargıdır.”³⁶ Makine öğrenimi sistemlerinin kasıtlı veya kasıtsız olarak, hâlihazırda var olan önyargıların yeniden üretilmesine neden olabilmesi önemli bir sorundur. Söz gelimi, algoritmik sistemler aracılığıyla yapılan işe alımlardaki cinsiyet yanlılığı-

32 Stahl, “Ethical Issues of AI”, 35–53.

33 G. Buttarelli, “Choose Humanity: Putting Dignity Back into Digital”, In: *Speech At 40th International Conference of Data Protection and Privacy Commissioners*, Brussels, (2018) https://www.privacyconference2018.org/system/files/201810/Choose%20Humanity%20speech_0.pdf. Erişim 10 Eylül 2020.

34 Stahl, “Ethical Issues of AI”, 35–53.

35 Emmanuel R. Goffi, *Responsible AI Implementing Ethical and Unbiased Algorithms*, ed. Sray Agarwal ve Shashin Mishra (London: Springer, 2021), ix-x.

36 CDEI, *Interim Report: Review into Bias in Algorithmic Decision-Making*, Centre for Data Ethics and Innovation, London, 2019. <https://www.gov.uk/government/publications/interim-reports-from-the-centre-for-data-ethics-and-innovation/interim-report-review-into-bias-in-algorithmic-decision-making> Erişim: 08 Ağustos 2022.

nın veya denetimli serbestlik süreçlerinde makine öğrenimi yoluyla ırksal önyargıların devam ettirildiğine dair çokça bulgu söz konusudur.³⁷ Anlaşılacağı üzere YZ kaynaklı olarak toplumun karşı karşıya kaldığı yasal ve etik sorunlar arasında mahremiyet ve gözetim, önyargı veya ayrımcılık yer alır. Dolayısıyla potansiyel olarak felsefi zorluk, insan yargısının rolüdür. Çünkü verilere gömülü önyargılarla baş etmek nerdeyse imkânsızdır.

Yeni dijital dönüşüm çağında, emek yoğun işgücünün YZ destekli sistemlerle değiştirilmesi, çalışanların yerinden edilmesiyle ilgili endişeleri artırıyor. Günümüz teknoloji çağında otomatlar mağaza görevlilerinin, ilgili yasal belgeleri bulan otomatik keşif programları avukatların ve yasal yardımcılarının, otomatik sanal asistanlar ise müşteri hizmetleri temsilcilerinin yerini alıyor. Yapay zekâ; hastalıkları teşhis etme, dilleri tercüme etme, müşteri hizmetleri sağlama gibi birçok “insan” işinde başarılı oluyor ve hızla gelişiyor. YZ'nin tüm bu başarıları, onun insan işçilerin yerini alacağına dair makul korkuları artırıyor.³⁸ Dijital araçlar bize, biz de onlara karşı daha önce hiç bu kadar duyarlı olmamıştık. YZ, işin nasıl yapıldığını ve kimin yaptığını kökten değiştirecek olsa da teknolojinin daha büyük etkisi, insan yeteneklerinin yerini alması değil tamamlaması ve artırması olacaktır.

YZ donanımlı sistemlere ilişkin korkuların yersiz olduğunu savunanlara göre insanların yaptığı iş yok olmaz, tam tersine iş şeklinin değiştiğini ve yeni beceriler gerektirdiğini görüyoruz. İnsanlara ve teknolojiye saygı duyarak YZ'yi yönetebilmemiz için akıllı makinelerle birlikte çalışma konusunda daha iyi olmamız gerekiyor. Bu anlamda insanlar ve makineler birbirlerinin güçlü yanlarını artırabilir. Hatta YZ'yi iş süreçlerine dahil etmek angaryayı ortadan kaldırır. Bu tür işbirlikçi zekâ yoluyla insanlar ve YZ, birbirlerinin tamamlayıcı güçlerini aktif olarak geliştirir. Bir taraftan insan özelliklerinden olan liderlik, takım çalışması, yeni şeyler ortaya koyabilme yeteneği ve sosyal beceriler, diğer taraftan ise YZ'nin öne çıkan nitelikleri arasında yer alan hız, ölçeklenebilirlik ve niceliksel yetenekler gibi her iki alanın güçlü yanlarının birbirini tamamlaması, çok daha etkili sonuçlar ortaya çıkaracaktır.

YZ'nin insan toplumları için ortaya koyduğu başlıca etik zorluklardan biri de insanın kendisini ikinci planda bir varlık olarak görmesine neden olmasıdır. Hızlı, güçlü ve verimli hâle gelen YZ sistemleri karşısında insanlar kendilerini bazı açılardan aciz hissedebilmektedir. İşlerimizi gerçekleştirme şeklimizin otomatikleşmeye devam etmesi insan olarak katkımızın ne olacağını gündeme getirmektedir. YZ'nin bu yönleri, insan olmanın gerçekte ne anlama geldiği konusunda bizleri anlam arayışına sürüklemektedir.

Bununla birlikte özel hayatımızdaki teknolojik müdahaleler karşısında uyanık kaldığımız sürece, YZ'nin gelişimi bireyin özerkliği ve özgürlüğünü zayıflatamayacaktır. Son olarak belirtilmelidir ki bazı iddiaların aksine makineler, insanlık için varoluşsal bir tehdit oluşturmaz. Çünkü özerkliği tamamen teknolojik olan YZ yalnızca bilgi almaktan karar vermeye giden maddi nedensellik zincirlerine tekabül etmektedir. Öte yandan, makinelerin hiçbir ahlaki özerkliği yoktur dolayısıyla karar verme sürecinde kafamızı karıştırsalar ve bizi yanıltmalar bile, kendi iradeleri yoktur ve onlara verdiğimiz komutlara tabidirler.³⁹

YZ'ye sahip sistemlerde başı çeken endişelerden biri de robot ordularıdır. Özerk bir komuta yapısına sahip robot orduların oluşturulmasının kaçınılmaz olduğunu belirten uzmanlar, bu robotların mutlak güç ve cezasızlık duygularının ortaya çıkmasına yol açacağına inanmaktadır. Bir çatışma halinde insanlar otonom yapıların aksine karşı tarafın hareket şeklini, ahlaki tutumlarını, duygularını görmezden gelmezler. Yani

37 Bkz. FA. Raso vd. *Artificial Intelligence & Human Rights: Opportunities & Risks*, Berkman Klein Center Research Publication No. 6, 2018.

38 “Elbette, birçok şirket süreçleri otomatikleştirmek için yapay zekâyı kullandı ancak onu esas olarak çalışanları yerinden etmek için kullananlar yalnızca kısa vadeli üretkenlik kazanımları görecektir. 1.500 şirketi kapsayan araştırmada, firmaların en önemli performans artışlarını insanlar ve makineler birlikte çalıştığında elde ettiği sonucuna varıldı. Bu tür işbirlikçi zekâ yoluyla, insanlar ve YZ, birbirlerinin tamamlayıcı güçlerini aktif olarak geliştirir: Birincisinin liderlik, ekip çalışması, yaratıcılık ve sosyal becerileri ve ikincisinin hızı, ölçeklenebilirliği ve nicel yetenekleri. İnsanlara doğal gelen (örneğin şaka yapmak) makineler için yanıltıcı olabilir ve makineler için basit olan (gigabaytlarca veriyi analiz etmek) insanlar için neredeyse imkânsız kalır. İş, her iki tür yetenek de gerektirir.” Bkz., “Collaborative Intelligence: Humans and AI are Joining Forces”, *Harvard Business Review*, <https://hbr.org/2018/07/collaborative-intelligence-humans-and-ai-are-joining-forces> Erişim: 20 Temmuz 2022.

39 Bkz. R. A. Calvo vd., “Supporting Human Autonomy in AI Systems: A Framework for Ethical Enquiry,” *Ethics of Digital Well-Being: A Multidisciplinary Approach*, ed. C. Burr- L. Floridi (Cham: Springer, 2020), 31–54.

profesyonel askerler arasında şefkat ve duyarlılık zamanla azalsa bile başkalarının acılarını doğrudan gözlemlemeleri, askeri personel üzerinde hala caydırıcı bir etkiye sahiptir. Silah uzmanı ve gazeteci Vasili Sychev'in de belirttiği üzere, robot orduları komuta edecek insanlar olmadan tam özerklikle düşmanlıklarını yürütebilecek ototnom silahlar için endişelenmeye başlamamızın zamanı geldi.⁴⁰ Sychev'in bu uyarısını dikkate alacak olursak otonom makinelerin, doğru etik kararları kendi başlarına verebilmelerine güvenmek isteriz. Bu ise bizlerin makineler için ideal bir etik oluşturmasıyla mümkündür.⁴¹

Diğer bir endişe kaynağı ise sosyal robotlar konusuyla bağlantılı olarak antropomorfizmdir. Antropomorfizm, genel olarak, belirgin şekilde insana benzer duyguların, zihinsel durumların ve davranışsal özelliklerin cansız nesnelere, hayvanlara ve genel olarak doğal fenomenlere ve doğaüstü varlıklara atfedilmesi olarak tanımlanır.⁴² Bu görüşe göre, YZ'nin insan özelliklerine benzediğinin inkârı, insanın benzersiz ve özel olduğuna dair hümanist bir anlayışı, yani felsefi ve bilimsel nedenlerle geniş çapta eleştirilen insan merkezli görüş türünü ima eder.⁴³ Bu argüman, yanlışlıkla ontolojik ve ahlaki düşünceleri karıştırıyor gibi görünmektedir. YZ'de insana benzer özelliklerin varlığını sorgulamak, ahlaki insan istisnacılığını savunmayı değil ontolojik bir perspektiften insanların ve YZ'nin farklı olduğu gösterir. Bu tür bir sorgulama, hümanist bir düşüncüyü benimsemeyi, diğer zekâ "türlerinin" mümkün olduğunu inkâr etmeyi ya da insan seviyesindeki zekânın zekâ için en iyi standart olduğunu düşünmeyi gerektirmez. Kavramsal düzeyde, insan özelliklerini YZ'ye yansıtmak, insanları model veya paradigma olarak konumlandırmakla sonuçlanabilir ve bu, potansiyel olarak farklı zekâ biçimlerine kapalı, insan merkezli bir ontoloji biçimine geri dönmeyi sağlar.

Bu bakımdan, insanbiçimleştirme eğiliminin birtakım etik sonuçları vardır. İlk olarak, YZ'leri insan benzeri olarak algılamak, onları hem ahlaki failer olarak hem de eylemlerini özerk bir karar verme sürecinin sonucu olarak görmeyi gerektirir.⁴⁴ İkincisi, insansı robotlar hem YZ'ye ilişkin aşırı korkuların (örneğin, YZ'nin insanları modası geçmiş kılacağı) hem de eleştirel olmayan iyimserliğin (YZ'nin gerçekte insanlar gibi davranabilme ve zor görevleri insanlardan daha iyi yapabilme derecesine ilişkin) kaynağı olabilir. Son olarak, yukarıda değindiğimiz türden bir antropomorfizm, ahlaki ve ontolojik sınırları bulanıklaştırarak etik karışıklık doğurabilir.⁴⁵

40 Vasily Sychev, "The Threat of Killer Robots", *Unesco Courier*, <https://en.unesco.org/courier/2018-3/threat-killer-robots> Erişim: 13 Haziran 2022.

41 Anderson, *Machine Ethics*, 1.

42 Bkz. G. Airenti, "The Cognitive Basis of Anthropomorphism: From Relatedness to Empathy", *International Journal of Social Robotics* 7(1): 2015, 117–127. Ayrıca Bkz. N. Epley vd., "On Seeing Human: A Three-Factor Theory of Anthropomorphism", *Psychological Review* 114(4) 2007, 864-886.

43 Bkz. Rosi Braidotti, *The Posthuman* (Cambridge: Polity, 2013).

44 A. Waytz vd. "Who Sees Human? The Stability and Importance of Individual Differences in Anthropomorphism", *Perspectives on Psychological Science* 5(3): (2010), 219-232.

45 Arleen Salles vd. "Anthropomorphism in AI", *AJOB Neuroscience*, 11:2 (2020), 88-95.

SONUÇ

Değişen dijital dünyada, kendi kendini idame ettiren YZ bazlı makineler ve bilgisayar uygulamaları her geçen gün daha da karmaşık hâle gelmekte ve yaşamın birçok yönünü dönüştürmektedir. Teknoloji felsefesinin en dinamik konularından biri haline gelen YZ etiği geleneksel ahlak anlayışlarımızı yeniden tanımlama ve sorgulama gerekliliğini doğurmaktadır. Etik YZ araştırmalarına ilişkin son literatürde gündeme getirilen hususlar arasında güvenlik, şeffaflık ve mahremiyet gibi geniş çapta tartışılan etik sorunlarla birlikte, otonom araçların çalışma sektöründe baskın hale gelmesinin doğuracağı istihdam sorunu ve birçok alanda akıllı makinelerin insanın yerini alacağı yönünde büyüyen kaygılar yer almaktadır.

Esasında her teknolojide olduğu gibi YZ de tarafsız bir yapı arz eder. Ancak YZ'nin birey ve toplum yaşamına yapıcı ya da yıkıcı etkisi onu kullanırken yahut tasarlarken benimsediğimiz etik tercihlerle bağlantılıdır. YZ donanımlı sistemlerin insanlığın, diğer canlı türlerin ve gezegenin ortak iyiliğini temel alarak inşa edilmesi, etik sorular ve kaygılar konusunda sergilediğimiz hassasiyetle yakından ilintilidir. Etik değerlerden yoksun ya da bu değerlerin titizlikle işlenmediği bir YZ'nin bize vadedeceği yaşam her ne kadar hayatı kolaylaştıran bir etkiye sahip gibi görünse de beraberinde getireceği adaletsizliklerle çok da mutlu bir yaşam sunmayacaktır. Dolayısıyla etik YZ oluşturmanın gerekliliği ve zorluğunun bilincinde olarak insandan daha akıllı YZ'ler tasarlarken, eşitliği ve adaleti bu sistemlerin kalbinde tutmak önceliklerimiz arasında yer almalıdır. İleri teknolojinin insanlar ve doğa için tehlikeli boyutlara gelmemesi ve iyi bir gelecek tasavvuru için kontrollü YZ düzenlemeleri elzemdir. Bu bağlamda yakın gelecekte, YZ'nin post-hümanların yeteneklerini geride bırakarak farklı bir türe evrilmesi olasılığının etik sonuçlarının neler olabileceği şimdiden sorgulanmalıdır. Çünkü biz hazır olsak da olmasak da YZ geliyor.

KAYNAKÇA

- Agarwal, Sray- Mishra, Shashin. *Responsible AI Implementing Ethical and Unbiased Algorithms*. Switzerland: Springer, 2021.
- Airenti, G. “The Cognitive Basis of Anthropomorphism: From Relatedness to Empathy”. *International Journal of Social Robotics* 7(1): 2015, 117-127.
- Anderson, Michael. *Machine Ethics*. Cambridge: Cambridge University Press, 2011.
- Asada, M. “Towards Artificial Empathy”. *Int J of Soc Robotics* 7 (2015), 19-33.
- Asimov, Isaac. *Robots and Empire: The Classic Robot Novel*. New York: HarperCollins, 1986.
- Asimov, Isaac. *Runaround: A Short Story*. New York: Street and Smith. 1942.
- Bostrom, Nick- Yudkowsky, Eliezer. “The Ethics of Artificial Intelligence”. *Cambridge Handbook of Artificial Intelligence*, eds. William Ramsey- Keith Frankish: Cambridge University Press, 2011.
- Braidotti, Rosi. *The Posthuman*. Cambridge: Polity, 2013.
- Calvo, R. A. vd. “Supporting Human Autonomy in AI Systems: A Framework for Ethical Enquiry”. *Ethics of Digital Well-Being: A Multidisciplinary Approach*. ed. C. Burr- L. Floridi. Cham: Springer, 2020, 31-54.
- CDEI. *Interim Report: Review into Bias in Algorithmic Decision-Making*. Centre for Data Ethics and Innovation. London, 2019. <https://www.gov.uk/government/publications/interim-reports-from-the-centre-for-data-ethics-and-innovation/interim-report-review-into-bias-in-algorithmic-decision-making> Erişim 08 Ağustos 2022.
- Crotoft, Rebecca. The Killer Robots Are Here: Legal and Policy Implications. *36 Cardozo L. Rev:* 1837 (2015).
- Epley, N. vd. “On Seeing Human: A Three-Factor Theory of Anthropomorphism”. *Psychological Review* 114(4) 2007, 864-886.
- Epstein, Richard G. “The Case of The Killer robot (part 2)”. *Acm Sigcas Computers and Society* 24.4 (1994), 12-32.
- Ethical AI. *Glossary*, <https://c3.ai/glossary/artificial-intelligence/ethical-ai/> Erişim: 10 Mayıs 2022.
- European Commission. *High-Level Expert Group on Artificial Intelligence, Ethics Guidelines for Trustworthy AI*. Brussels: 2019. <https://www.aepd.es/sites/default/files/2019-12/ai-ethics-guidelines.pdf> Erişim: 08 Haziran 2022.
- FA., Raso. vd. *Artificial Intelligence & Human Rights: Opportunities & Risks*. Berkman Klein Center Research Publication: No. 6, 2018.
- Goffi, Emmanuel R. *Responsible AI Implementing Ethical and Unbiased Algorithms*. ed. Sray Agarwal-Shashin Mishra. London: Springer, 2021.
- Grant, George. *Technology and Justice*. Concord, Ontario: House of Anansi, 1986.
- Hagendorf, Thilo. “The Ethics of AI Ethics: An Evaluation of Guidelines”. *Minds and Machines* 30: (2020) 99- 120.
- Harvard Business Review, “Collaborative Intelligence: Humans and AI are Joining Forces”, <https://hbr.org/2018/07/collaborative-intelligence-humans-and-ai-are-joining-forces> Erişim: 20 Temmuz 2022.
- IEEE. *Ethically Aligned Design- A Vision for Prioritizing Human Well-being with Autonomous and Intelligent Systems*, USA: IEEE. 2019.
- Jobin, A. vd. “The Global Landscape of AI Ethics Guidelines”. *Nat Mach Intell* 1, (2019), 389-399.
- Lawton, George. “AI Ethics (AI code of ethics)”, *TechTarget*, <https://www.techtarget.com/whatis/definition/AI-code-of-ethics>. Erişim: 12 Mayıs 2022.
- Murphy, R.- Woods, D. “Beyond Asimov: The Three Laws of Responsible Robotics”. *IEEE Intelligent Systems* 24(4), (2009), 14-20.
- Nyholm, Sven. “The Ethics of Crashes with Self-driving Cars: A roadmap, I”. *Philosophy Compass* Vol. 13 (2018), 2-10.
- Parr, Sakiko Fukuda- Gibbons, Elizabeth. “Emerging Consensus on ‘Ethical AI’: Human Rights Critique of Stakeholder Guidelines”. *Global Policy* Vol. 12 (2021), 32 – 44.
- Picard, Rosalind. *Affective Computing*. Cambridge, MA and London: MIT Press: 1997.
- R. Richardson vd. “Dirty Data, Bad Predictions: How Civil Rights Violations Impact Police Data, Predictive Policing Systems, and Justice”. *N Y Univ Law Rev Online* 192. (2019) <https://ssrn.com/abstract=3333423>. Erişim 10 Eylül 2022.
- Salles, Arleen vd. “Anthropomorphism in AI”. *AJOB Neuroscience*, 11:2 (2020), 88-95.
- SHERPA, <https://www.project-sherpa.eu/> Erişim: 10 Haziran 2022.
- Sparrow, Robert. “Killer Robots”. *Journal of Applied Philosophy*, 24(1), (2007). 62-77.
- Stahl, Bernd Carsten. “Ethical Issues of AI”, *Artificial Intelligence for a Better Future*, (2021), 35-53.
- Sullins, John P. “When Is a Robot a Moral Agent?”. *International Review of Information Ethics* Vol. 6 (12/2006), 23-30.
- Sychev, Vasily. *The Threat of Killer Robots*, <https://en.unesco.org/courier/2018-3/threat-killer-robots> Erişim: 13 Haziran 2022.
- Topakkaya, Arslan- Eyibaş, Yağmur. Yapay Zekâ ve Etik İlişkisi. *Felsefe Dünyası*, Sayı: 70, (2019), 81-99.
- Waytz, A. vd. Who Sees Human? The Stability and Importance of Individual Differences in Anthropomorphism. *Perspectives on Psychological Science* 5(3): (2010), 219-232.

Yalçın, Ö.N.- DiPaola, S. Modeling Empathy: Building a Link Between Affective and Cognitive Processes. *Artificial Intelligence Review* 53, (2020), 2983-3006.

Yeşilkaya, Nazan. “Yeni Bir Varlık Türü Olarak Biyoteknolojik Varlıklar ve Adalet Sorunu”. *Şırnak Üniversitesi İlahiyat Fakültesi Dergisi* 26 (2021), 233-256.

EXTENDED ABSTRACT

In recent years, we have been considering transhumans made of flesh and steel, and cyborgs, in more comprehensive terms, post-humans. Leaving behind the view that humans were machines that could only be produced with techné, that is, art, we move towards the idea that humans will mix with machines. We are rapidly approaching a period when the design, distribution and control of new technologies and technological processes are transferred to machines. AI applications that cause significant changes at the global level by integrating many aspects of our lives; It transforms many sectors from health to transportation, from production to security, from education to entertainment. While AI algorithms or thinking machines take on the serious obligations of modern life and become an indispensable stage of social life, they also bring with them some ethical problems.

The integration of new technologies into daily life brings about the emergence of problems related to moral integrity. As the use of AI becomes widespread, complex ethical problems are coming to the agenda, which are highlighted by the hesitations and objections made. For example, self-driving cars with AI or autonomous weapons coded to kill without human intervention, “killer robots” are the subject of ongoing discussions in AI applications today as well as yesterday. Most of the discussions focused on ethics in AI cover problematic issues that need to be addressed from a moral point of view. The most prominent of these issues is the ethical debate about whether machines with AI can harm humans or other beings.

In these days, which are moving towards producing humanoid robots that dominate ethical values, another agenda item is about the “moral status” of thinking machines. While AI is gaining importance in making decisions in various aspects of our lives, it is necessary to examine whether these non-humanoid structures are “ethical” and whether they benefit people by adhering to the normative expectations of justice. Because machines and computers are becoming more complex by taking the initiative in terms of their social roles. The advent of machines with AI that can match or replace human capabilities can be interpreted as a major challenge to humans’ traditional sense of self as beings with the highest moral status in the world. Noting that AI has the potential to redefine our traditional moral concepts, ethical approaches and moral theories, the possibility of machines working autonomously without ethical assurance can be considered as a situation that will disturb most of us.

At this point, it should be noted that the more complex the technological developments, the more complex the ethical problems that will arise. Considering that great power brings great responsibility, whether machines with AI, which is touted as great power, will harm people or other assets, and the issue of responsibility are among important ethical discussions. For example, AI applications that stand out with developments such as “lethal autonomous weapon systems (LAWS)”, “autonomous cars” and “predictive policing” that can kill without human intervention are at the center of complex ethical issues. For this reason, AI can only work with people in trust and justice, by building an ethical AI that includes transparency, responsibility and accountability.

In fact, as in every technology, AI also presents an impartial structure. However, the constructive or destructive effect of AI on individual and social life is related to the ethical choices we adopt when using or designing it. Building AI-equipped systems based on the common good of humanity, other species, and the planet is closely related to our sensitivity to ethical questions and concerns. Although the life promised to us by an AI devoid of ethical values or in which these values are not meticulously processed may seem to have an effect that makes life easier, it will not offer a very happy life with the injustices it will bring along. Therefore, keeping equality and justice at the heart of these systems should be among our priorities while designing AIs that are smarter than humans, being aware of the necessity and difficulty of creating ethical AI. Controlled AI arrangements are essential for advanced technology not to become dangerous for humans and nature, and for a good future envisioning. In this context, the ethical consequences of the possibility of AI evolving into a different genre in the near future, leaving the capabilities of post-humans behind, should be questioned now. Because whether we’re ready or not, AI is coming.

Due to the factors mentioned above, it is important to highlight the inevitability of ethics in AI in order to evaluate the thematic issues and ethical concerns that require philosophical and ethical analysis of the areas of responsibility of non-human intelligences, and to avoid any technological ramblings by considering the pace of innovation. In this study, which includes thematic problems related to artificial intelligence supported systems, which are launched as transformative technology, it is aimed to clarify the responsibility areas of artificial intelligence and what the ethical concerns and problems about artificial intelligence are. In the article, which draws attention to the fact that artificial intelligence has the potential to redefine our traditional moral concepts, ethical approaches and moral theories, the need for ethical principles in artificial intelligence design is demonstrated with examples from current research in the field of technology. Emphasizing that we are rapidly approaching a period in which the design, distribution and control of new technologies and technological processes are transferred to machines, the aim of this study is to avoid any technological straying that may occur depending on the speed of innovation. The article calls on both policy makers and legislators to meticulously process ethical values and principles into AI under the guidance of philosophy in order to resolve the ethical problems that arise in AI and to alleviate concerns.