



Avrupa'da Dijital Etik, İnsan Hakları Bağlamında Yapay Zekâ ve Algoritmik Ayrımcılık

Nihan AKINCILAR KÖSEOĞLU¹, Belgin ÇETİN²

Öz

Makale, Yapay Zekâ (YZ) kaynaklı insan hakları ihlallerini ortaya koymayı amaçlamaktadır. YZ kullanımının etik çerçeve içinde gerçekleşmediği durumlarda insan hakları ihlallerinin ortaya çıktığını/çıkacağını iddia eden bu makale, önce YZ etiği ve dijital etik kavramlarının üzerinde durmaktadır. Makalenin kapsamı gereği, YZ etiğini meydana getiren Avrupa'da YZ ve insan hakları konulu yasal düzenlemeler temel alınmakta; ayrıca hakime/avukata ve bireylere yardımcı olan YZ uygulamalarından kısaca bahsedilerek, Avrupa'daki robot hakim uygulamalarına odaklanılmaktadır. Dolayısıyla makalenin sınırlılığı kapsamında, dijital etik ve YZ etiğinden bahsedildikten sonra, insan hakları bağlamında yapay zekâ kullanımı, başta Avrupa Birliği (AB) olmak üzere, OECD ve Avrupa Konseyi'nin hazırladığı yasal düzenlemeler, yapay zekâdan kaynaklanan insan hakları ihlalleri, yapay zekâ uygulamalarının ayrımcılık yasağı üzerindeki etkileri ve yeni bir ayrımcılık türü olarak algoritmik ayrımcılık konularına değinilmektedir. Türkiye'deki akademik çalışmalara bakıldığında, algoritmik ayrımcılık konulu henüz yeterli çalışma bulunmadığından, makalenin sunacağı akademik katkı yenilikçi olacaktır. Bu konuda -özellikle son birkaç yılda azımsanmayacak sayıda- Türkçe yayınlanmış diğer çalışmalar incelendiğinde, sanal mahkemeler, hakime/avukata ya da bireylere yardımcı olan yapay zekâ ve robot hakim yapay zekâ uygulamalarının tüm dünya genelindeki örneklerinin ele alındığı görülmektedir. Bu çalışmanın özgün değeri, YZ kaynaklı insan hakları ihlallerini önlemek için Avrupa'da YZ etiğinin oluşturulma sürecine odaklanmasındadır. Sonuç olarak, Avrupa'da YZ etik kurallarının oluşturulması, yaygınlaştırılması ve uygulanması süreci tamamlanmadıkça YZ kaynaklı insan hakları ihlallerinin devam edeceği kanısına varılmaktadır.

Anahtar Kelimeler: Dijital Etik, Yapay Zekâ, İnsan Hakları, Robot Hakimler, Algoritmik Ayrımcılık, Avrupa

Jel Kodu: N44, O52, N74

Digital Ethics, Artificial Intelligence in the Context of Human Rights and Algorithmic Discrimination in Europe

Abstract

This article aims to reveal human rights violations caused by Artificial Intelligence (AI). Claiming that human rights violations occur/will occur when the use of AI is not carried out within an ethical framework, this article first focuses on the concepts of AI ethics and digital ethics. Due to the scope of the article, the legislations on AI and human rights in Europe that create AI ethics are taken as basis; Additionally, AI applications that assist judges/lawyers and individuals are briefly mentioned, focusing on robot judge applications in Europe. Therefore, within the scope of the limitation of the article, after mentioning digital ethics and AI ethics, the use of artificial intelligence in the context of human rights, legislations prepared by the European Union, OECD and the Council of Europe, human rights violations caused by artificial intelligence, the effects of artificial intelligence applications on the prohibition of discrimination. and algorithmic discrimination as a new type of discrimination are touched upon. Considering the academic studies in Turkey, the academic contribution of the article will be innovative, as there are not yet sufficient studies on algorithmic discrimination. When other studies published in Turkish on this subject - especially in the last few years - are examined, it can be seen that examples of virtual courts, artificial intelligence that assists judges/lawyers or individuals, and robot judge artificial intelligence applications all over the world are discussed. The unique value of this study is that it focuses on the process of establishing AI ethics in Europe to prevent AI-related human rights violations. As a result, it is concluded that AI-related human rights violations will continue unless the process of establishing, disseminating and implementing AI ethical rules in Europe is completed.

Keywords: Digital Ethics, Artificial Intelligence, Human Rights, Robot Judges, Algorithmic Discrimination, Europe

Jel Codes: N44, O52, N74

¹ **Sorumlu Yazar/Corresponding Author:** Doç. Dr., Fenerbahçe Üniversitesi, İktisadi, İdari ve Sosyal Bilimler Fakültesi, Siyaset Bilimi ve Uluslararası İlişkiler Bölümü, İstanbul, Türkiye. **E-posta:** nihan.koseoglu@fbu.edu.tr **Orcid no:** 0000-0002-8565-506X

² Kütüphane Direktörü, Fenerbahçe Üniversitesi, Kütüphane ve Dökümantasyon, İstanbul, Türkiye. **E-posta:** belgin.cetin@fbu.edu.tr **Orcid no:** 0000-0001-7696-3012

Extended Abstract

Introduction

Today's world is rapidly digitalizing through 5G, cloud technologies, artificial intelligence, algorithms, machine learning, data analytics, edge computing and the internet of things. This digitalization of every field requires greater emphasis on digital ethics. Digital ethics can be defined as integrating digital technology and human values in a way that digital technology advances human values rather than harming them. The integration of ethics and technology is of interest to a broader community beyond developers, ethicists, and philosophers. With the increasing incorporation of technology into our daily practices, significant uncertainties arise regarding the ethical use of technology. Knowledge about technology is abstract, unlimited and endless, and the quest to encode ethics into technologies is a necessity.

Digitalization and information technologies are widely used in many countries around the world. The increasing use of artificial intelligence (AI) in recent years is being used in all areas of daily life. The use of AI in the field of law manifests itself with examples of AI that assist judges/lawyers or robot judge applications that replace judges in order to alleviate the excessive workload of judges and lawyers. While the use of AI in making court decisions in these two ways, on the one hand, alleviates the excessive workload of judges and lawyers and reduces costs for individuals, on the other hand, it raises questions about transparency, responsibility for decisions or appeals, and causes human rights violations that manifest themselves in the form of discrimination due to algorithmic errors. Therefore, international organizations are trying to take the necessary measures to prevent possible human rights violations in this regard.

Method

Within the scope of the limitation of the article, after mentioning digital ethics and AI ethics, the use of artificial intelligence in the context of human rights, legislations prepared by the European Union, OECD and the Council of Europe, human rights violations caused by artificial intelligence, the effects of artificial intelligence applications on the prohibition of discrimination. and algorithmic discrimination as a new type of discrimination are touched upon. Considering the academic studies in Turkey, the academic contribution of the article will be innovative, as there are not yet sufficient studies on algorithmic discrimination. When other studies published in Turkish on this subject-especially in the last few years - are examined, it can be seen that examples of virtual courts, artificial intelligence that assists judges/lawyers or individuals, and robot judge artificial intelligence applications all over the world are discussed. The unique value of this study is that it focuses on the process of establishing AI ethics in Europe to prevent AI-related human rights violations.

Result and Discussion

Today, as digital innovations develop, an information society emerges that requires digital ethics and therefore respect for human rights. In this context, AI applications can lead to human rights violations and discrimination. The main reason for this is often based on biased algorithms. The development of digital ethics is necessary at this point. The use of facial recognition technology or recruiting software violates human rights. AI techniques that access customer information violate the protection of personal information. On the other hand, if AI accesses people's posts through the algorithms used in social media programs, it violates the person's freedom of expression.

In order to prevent human rights violations and algorithmic discrimination resulting from AI applications, an international legal standard should first be created and AI applications that comply with international human rights standards should be preferred. The possibility of biased decisions arising from algorithmic problems that may arise in the decision-making process of the robot judge and the possibility of not satisfying the society about the decisions made are the most fundamental problems of AI applications in the judiciary.

At this point, the EU AI Act Proposal of 2021 is very important. This proposal must become law, be accepted and implemented by all EU member and candidate countries, and shed light on the international AI legal standard. Considering that 47% of employment, including judges at the ECHR, can be automated by 2030, AI can be used in every field. As a result, it is concluded that AI-related human rights violations will continue unless the process of establishing, disseminating and implementing AI ethical rules in Europe is completed.

GİRİŞ

Dijitalleşme ve bilgi teknolojileri, dünya üzerindeki birçok ülkede yaygın olarak kullanılmaktadır. Son yıllarda artan YZ kullanımına, zamanla günlük hayatın her alanında başvurulmaktadır. Hukuk alanında YZ'nin kullanılması, hâkim ve avukatların aşırı iş yükünü hafifletmek için hakime/avukata yardımcı olan YZ örnekleriyle ya da hakim yerine geçen robot hakim uygulamalarıyla kendini göstermektedir. YZ'nin bu iki şekilde mahkeme kararlarının alınmasında kullanılması bir yandan hâkim ve avukatların aşırı iş yükünü hafifletip bireyler için maliyeti düşürürken, diğer yandan şeffaflık, kararların sorumluluğu ya da temyiz konularında soru işaretleri yaratıp algoritmik hatalar nedeniyle de ayrımcılık şeklinde kendini gösteren insan hakları ihlallerine neden olmaktadır. Dolayısıyla uluslararası örgütler bu konuda olası insan hakları ihlallerini önlemek amacıyla gerekli önlemleri almaya çalışmaktadır.

Bugünün dünyası 5G, bulut teknolojiler, yapay zekâ, algoritmalar, makine öğrenimi, veri analitiği, uç bilgi işlem ve nesnelerin interneti aracılığıyla hızla dijitalleşmektedir. Her alanın bu şekilde dijitalleşmesi, dijital etik konusuna daha fazla vurgu yapılmasını gerektirmektedir. Dijital etik, dijital teknoloji ile insani değerleri, dijital teknolojinin insani değerlere zarar vermek yerine onları ilerletecek şekilde bütünleştirmesi olarak tanımlanabilir (Rogerson, 2020). Etik ve teknolojinin entegrasyonu, geliştiricilerin, etik uzmanlarının ve filozofların ötesinde daha geniş bir topluluğun ilgisini çekmektedir. Teknolojinin günlük uygulamalarımıza giderek daha fazla dâhil olmasıyla birlikte, teknolojinin etik kullanımına ilişkin önemli belirsizlikler ortaya çıkmaktadır (Verma & Garg, 2023). Teknolojiye ilişkin bilgi soyut, sınırsız ve sonsuzdur ve etiği teknolojilere kodlama arayışı bir zorunluluktur (Green, 2021:209).

Makale, Yapay Zekâ (YZ) kaynaklı insan hakları ihlallerini ortaya koymayı amaçlamaktadır. YZ kullanımının etik çerçevede içinde gerçekleşmediği durumlarda insan hakları ihlallerinin ortaya çıktığını/çıkacağını iddia eden bu makale, önce YZ etiği ve dijital etik kavramlarının üzerinde duracaktır. Makalenin kapsamı gereği, YZ etiğini meydana getiren Avrupa'da YZ ve insan hakları konulu yasal düzenlemeler temel alınacak; ayrıca hâkime/avukata ve bireylere yardımcı olan YZ uygulamalarından kısaca bahsedilerek, Avrupa'daki robot hâkim uygulamalarına odaklanılacaktır. Dolayısıyla makalenin sınırlılığı kapsamında, dijital etik ve YZ etiğinden bahsedildikten sonra, insan hakları bağlamında yapay zekâ kullanımı, başta Avrupa Birliği olmak üzere, OECD ve Avrupa Konseyi'nin hazırladığı yasal düzenlemeler, yapay zekâdan kaynaklanan insan hakları ihlalleri, yapay zekâ uygulamalarının ayrımcılık yasağı üzerindeki etkileri ve yeni bir ayrımcılık türü olarak algoritmik ayrımcılık konularına değinilecektir.

Türkiye'deki akademik çalışmalara bakıldığında, algoritmik ayrımcılık konulu henüz yeterli çalışma bulunmadığından, makalenin sunacağı akademik katkı yenilikçi olacaktır. Bu konuda -özellikle son birkaç yılda azımsanmayacak sayıda Türkçe yayınlanmış diğer çalışmalar incelendiğinde, sanal mahkemeler, hâkime/avukata ya da bireylere yardımcı olan yapay zekâ ve robot hâkim yapay zekâ uygulamalarının tüm dünya genelindeki örneklerinin ele alındığı görülmektedir. Bilgin (2022) ve Sümer (2021), robot hâkimlerin mahkeme kararı almasından; Karabel ve Aydemir (2023), Medeni Usul Hukukunda yapay zekânın kullanımından; Singil (2022), yapay zekânın neden olduğu insan hakları ihlalleri örneklerinden; Yılmaz (2020), Avrupa Etik Şartı'ndan ve Yılmaz vd. (2021), Avrupa Birliği ve Amerika Birleşik Devletleri'ndeki YZ ile ilgili düzenlemelerden karşılaştırmalı olarak bahsetmektedir. Bu çalışmanın özgün değeri ise, YZ kaynaklı insan hakları ihlallerini önlemek için Avrupa'da YZ etiğinin oluşturulma sürecine odaklanmasındadır. Sonuç olarak, Avrupa'da YZ etik

kurallarının oluşturulması, yaygınlaştırılması ve uygulanması süreci tamamlanmadıkça YZ kaynaklı insan hakları ihlallerinin devam edeceği kanısına varılmaktadır.

1. DİJİTAL ETİK

Dijital dönüşüm, sağlık hizmetlerinden akıllı şehirlere ve küresel ısınmaya kadar özel ve kamusal yaşamı ve çevremizi iyileştirmek için büyük fırsatlar sunmaktadır. Ancak bu tür fırsatlar önemli etik zorlukları da beraberinde getirmektedir. Özellikle, giderek daha fazla verinin (büyük veri) yaygın kullanımı ve seçimleri şekillendirmek, kararlar almak ve analiz etmek için makine öğrenimi, yapay zekâ, robotik teknolojilerinin kullanımı; algoritmalara artan güven ve insan katılımının veya birçok otomatik süreç üzerindeki gözetimin kademeli olarak azalması, adalet, sorumluluk ve insan haklarına saygı hakkında acil sorunlar doğurmaktadır. Sanal ortamda etik, dijital yeniliklerin gelişmesi ve uygulanmasını teşvik ederken, aynı zamanda insan haklarına saygıyı ve açık, çoğulcu, hoşgörülü bilgi toplumunun şekillenmesini sağlamalıdır. Bu nedenle insan yaşamı üzerinde etkisi olan herhangi bir dijital proje için güçlü bir etik denge kurmak önemlidir. Bunu başarmak için dijital etik, son otuz yıldır bilgi ve iletişim teknolojilerinin getirdiği zorluklara odaklanan bilgisayar ve bilgi etiği tarafından sağlanan temel üzerine inşa edilmektedir (Floridi, 2013:19; Bynum, 2015). Böylece dijital etik, bilgisayar ve bilgi etiğinde zaman içinde onaylanmış yaklaşımları, etik sorgulamaların soyutlama düzeylerini (Levels of Abstraction (LoA)) bilgi merkezli (information-centric (LoAI)), dijital merkezli olana (digital-centric (LoAD)) değiştirerek gelişmektedir (Floridi vd., 2019:2). Bu alandaki ilk araştırmalar, bilgisayarların yaygınlaşmasının yarattığı etik sorunları hem tasarımcıların hem de kullanıcıların mesleki sorumlulukları açısından ele alan insan merkezli bir soyutlama düzeyini benimsemiştir (Parker, 1968:270). Soyutlama düzeyi daha sonra 1980'lerin ortalarında bilgisayar merkezli somut temele (computer-centric (LoAC)) kaymıştır ve 20. yüzyılın başında tekrar soyut temellere dayanan bilgi merkezli (LoAI) soyutlama olarak değiştirilmiştir (Moor, 1985:273; Floridi, 2006:9). Bu değişiklikler önemli kavramsal çıkarımları olduğu için hızlı, yaygın teknolojik dönüşümlere yanıt vermiştir. Örneğin, bilgisayar merkezli soyutlama (LoAC), bilgisayarların sosyal dinamikleri şekillendirme ve çevre tasarımı üzerindeki etkisini anlamayı kolaylaştırarak, bilgisayarların evrensel ve baş edilebilir (kodlanabilir) araçlar olduğunu vurgulamıştır (Moor, 1985:271). Daha sonra, soyutlama düzeyi (LoA) odak noktasını teknolojik araçlardan (donanım, bilgisayarlar, cep telefonları vb.), bu araçlar aracılığıyla oluşturulabilen, kaydedilebilen, işlenebilen ve paylaşılabilen içeriğe (bilgiye) kaydırmıştır. Bilgi merkezli soyutlama (LoAI), bilginin farklı ahlaki boyutlarını (ahlaki eylemlerin kaynağı, sonucu veya hedefi olarak bilgi gibi) vurgulayarak makro-etik bir yaklaşımın tasarlanmasına yol açmıştır. Ayrıca bilginin paylaşım, depolama, korunma, kullanım ve olası imha döngüsünün tamamını ele almıştır (Floridi, 2006:8).

Bugün artık etik stratejisi, konunun sadece belirli bir teknoloji veya donanım olmadığını (çevrimiçi platformlar, bulut bilişim, Nesnelerin İnterneti, YZ vb. dahil), aksine dijital teknoloji tarafından yaratılan ve manipüle edilen tüm ekosistem olduğunu anlamıştır. Dijital etik kavramı ile odak noktası artık, etik sorunlara neden olan donanımın kendisi değil, yeni etik sorunlara yol açan donanımın yazılımla, verilerle, araçlarla davranışları ve ilgili ortamlarda nasıl davrandığıdır. Bu nedenle, “robo-etik”, ‘yapay zekâ etiği’ veya “makine etiği” gibi kavramlar oluşmaya başlamıştır. Dijital etik, tüm dijital gerçeklik yelpazesinin farklı ahlaki boyutlarını odak noktasına getirmektedir. Örneğin, algoritma etiği, özellikle makine öğrenimi uygulamaları söz konusu olduğunda, algoritmaların artan karmaşıklığı ve özerkliğinin ortaya koyduğu sorunları ele almaktadır (Floridi vd., 2019:4). Öngörülemeyen ve istenmeyen sonuçlarla ilgili olarak hem tasarımcıların hem de veri

bilimcilerin ahlaki sorumluluğu ve hesap verebilirliği çok önemli zorluklardır. Bunlar, algoritmaların etik tasarımı ve denetimi, potansiyel istenmeyen sonuçlarla mücadele (örneğin, ayrımcılık veya insan hakları ihlalleri) gibi konulardır (Floridi vd., 2019:7).

Dijital etik en kısa tanımı ile ahlaki açıdan iyi çözümleri formüle eden, destekleyen ve ilgili uygulamalar ve altyapılar ile en standart veri ve algoritmaları inceleyen ve değerlendiren etik dalıdır (Floridi vd., 2019:8). Diğer bir bakış açısı ile bilgi toplumunu olumlu yönde etkilemek, insanların kendi arasında ve çevresiyle olan etkileşimini geliştirmek, insan onurunu korumak ve dünyada insan gelişimini teşvik etmek için hayati önem taşıyan, yeni ve hızla gelişen bir araştırma alanıdır. Dijital etiğin sürdürülebilirliği için birçok uzmanlık alanı ile çok disiplinli iş birliğine ve etik düşüncenin tüm dünyada gelişen teknolojik yeniliklere uygulanabilmesi için gerekli değerlere sürekli ve yılmadan odaklanması önemlidir (Floridi vd., 2019:9).

1.1. Uygulamalı Etik Alanı Olarak Dijital Etik

Etik ve teknoloji arasındaki etkileşim üzerine yapılan araştırmalar, teknolojinin insan yaşamı için yeni olanaklar sağladığını ve aynı zamanda yeni etik sorunlar yarattığını göstermektedir. Bir anlamda “uygulamalı etik” teknolojik gelişmenin ürünüdür (Khalili & Ghasemzadeh, 2021:11). 1980'lerde ve 1990'larda dijital etik konusunun varlığı veya gerekliliği hem felsefe bilminde hem de uygulamalı etikte ve bilgisayar bilminde tartışılmaya başlanmıştır. Dijital etikte ve uygulamalı etikte ortaya çıkan karakteristik özelliklerden biri, bir problemin nasıl araştırmaya değer bir problem haline geldiğidir. Geleneksel felsefede ölçüt, genellikle gelenekte mevcut bir tartışma olması ve bu tartışmada felsefi olarak ilginç, çözülmemiş bir olay olmasıdır. Modern felsefe ise dijital etik kavramında olduğu gibi insanları tasarım sürecinde duyarlı hale getirmeye ('tasarım yoluyla etik') ve insanların etik olarak doğru olan şeyin ne olduğunu gerçekten bilmediği sorunları seçmeye ve pratik problemlere çözüm üretmeye çalışır (Müller, 2022:15).

Dijital dönüşüm, toplumun her düzeyinde köklü değişimlere neden olmuştur. Ancak, kapasite artışı, minyatürleştirme (her zamankinden daha küçük mekanik, optik ve elektronik ürünler ve cihazlar üretme), mikro donanımların yaygınlaşmasıyla birlikte çeşitli kaynaklardan büyük miktarda veri toplama ve analizi için algoritmik sistemlerin kullanılmaya başlanması ile etik sorunlar katlanarak büyümektedir (Floridi, 2013:14). Yirminci yüzyılın ikinci yarısından bu yana bilgisayar ve bilgi etiği tarafından atılan temeller üzerine inşa edilen dijital etik alanı (Moor, 1985:266; Floridi, 2013:290; Floridi vd., 2019:3), her ikisinin de ortaya koyduğu etik zorluklarla mücadele etmektedir (Becker vd., 2022:2629). Yazılım geliştiricilerin, tasarımcıların ve yönetimsel karar vericilerin, dijital ürünleri oluştururken etik değerleri göz önünde bulundurmaları ve normatif değerlendirmeler yapmaları her zamankinden daha fazla beklenmektedir (Zuber vd., 2022:4). Rogerson (2017:1), teknoloji geliştirme aşamasında; etkinlik, verimlilik ve etiğin öneminden bahseder ve bu öğelerin dijital ürünün (tasarım, yazılım vs.) ihlale sebep olmayacak şekilde başlangıcından itibaren doğru uygulanması gerektiğini vurgulamaktadır.

Dijital dönüşüm içinde dijital etik kavramının sahaya indirgenmesi, yaygınlaştırılması, dijital ayrımcılığın önüne geçilmesi ve sürdürülebilirlik için çerçeve politikalar, prosedürler ve standartların oluşturulması son derece önemlidir. Etik modelleri ve ilkeleri uygulamalara entegre etmek amacıyla çoğunlukla Avrupalı bölgesel örgütler tarafından geliştirilmiş ve yürürlükte olan önemli yasal düzenlemeler mevcuttur. Makalenin ikinci başlığında bu yasal düzenlemeler detaylı bir şekilde incelenecektir.

1.2. Yapay Zekâ Etiği

YZ tabanlı teknoloji; yüz tanıma, tıbbi teşhis, kendi kendini süren arabalar gibi otomotiv sektöründe ve ekonomik büyüme, sosyal kalkınma, insan refahı, güvenliğin iyileştirilmesi gibi alanlarda büyük faydalar vaat etmektedir. Bununla birlikte, YZ tabanlı teknolojinin kullanımı düşük düzeyde açıklanabilirlik, veri yanlışlıkları, veri güvenliği, veri gizliliği gibi etik sorunlar ve kullanıcılar, geliştiriciler ve toplumlar için önemli riskler oluşturmaktadır. YZ ile ilişkili etik ve ahlaki zorlukların nasıl ele alınacağı önemli bir konudur (Siau & Wang, 2020:74). YZ sistemlerinin yerel veya geçici etik standartları yansıtması pek olası değildir ve farklı sosyal normlara uyacak şekilde uygulamalara, ürün tasarımına entegrasyonu zordur (Johnson, 2020:46). YZ ile ilgili etik ve ahlaki sorunları anlamak ve ele almak hala emekleme aşamasındadır. Ancak, insanlığın geleceği YZ etiğinin doğru gelişimine bağlı olabilir (Siau & Wang, 2020:84).

YZ'nin bir akademik araştırma alanı olarak kurulmasının başlangıcı 1950'lere kadar gitmektedir. Etik tartışması da bir o kadar eskidir. Ancak, YZ sistemlerinin yeteneklerindeki ve uygulamalarındaki etkileyici ilerlemeler, yalnızca son yıllarda YZ'nin toplum için meydana getirdiği fırsatları ve riskleri daha keskin hale getirdiğinden YZ'nin toplum üzerinde yarattığı etkilere ilişkin politikalara talep artmaktadır (Floridi, 2018:2). YZ teknolojileri toplumda daha yaygın hale geldikçe, YZ kurallarını düzenlemek ve kontrol ilkelerini uygulamak önem kazanmaktadır. YZ'yi tasarlayan, geliştiren veya uygulayan her şirket, devlet kurumu veya akademik kurumun bunu etik çerçeveye uygun olarak yapma yükümlülüğü vardır (Floridi, 2018:3). Aksi takdirde birçok farklı insan hakları ihlallerine neden olmaktadır ve olacaktır (Floridi, 2018:5).

2. AVRUPA KITASINDAKİ BÖLGESEL ÖRGÜTLERDE İNSAN HAKLARI BAĞLAMINDA YAPAY ZEKÂ

Günümüzde birçok alanda YZ kullanımı arttıkça insan hakları alanında olumsuz gelişmeler yaşanmaktadır. YZ'nin insan hakları ihlallerine yol açmasını önlemek adına yakın tarihlerde Avrupa kıtasındaki bölgesel örgütler özellikle, Avrupa Konseyi, OECD ve Avrupa Birliği- YZ ve dijital etik konularında ilke ve değerler belirlemektedir.

2.1. Avrupa Konseyi'nde İnsan Hakları Temelli Yapay Zekâ Düzenlemeleri

Avrupa Konseyi ilk olarak Aralık 2018'de, "Yapay Zekânın Yargı Sisteminde Kullanılmasına Dair Avrupa Etik Şartı"nı (CEPEJ European Ethical Charter on the Use of Artificial Intelligence (AI) in Judicial Systems and Their Environment) kabul etmiştir (Council of Europe, 2018). Bu metin, YZ'nin yargı sistemlerinde kullanımına ilişkin etik ilkeleri ortaya koyan ilk Avrupa metni olma özelliğindedir. Şart, üye ülkelerin politika yapıcılara, yasa koyucularına ve yargı çalışanlarına yol gösterebilecek bir çerçeve sunmaktadır. Şart'ın AIHS'ne de atıfta bulunarak altını çizdiği "temel haklara saygı ilkesi" ve "ayrımcılık yapmama ilkesi" YZ'nin yargı sistemlerinde kullanımının odak noktasına konulmuştur (Yılmaz, 2020).

Avrupa Konseyi'nin Şubat 2019'da Helsinki'de düzenlediği Üst-Düzey Konferans'ta, "Oyun Değiştiricinin Yönetilmesi- Yapay Zekâ Geliştirmenin İnsan Hakları, Demokrasi ve Hukukun Üstünlüğü Üzerindeki Etkileri" (Governing the Game Changer – Impacts of Artificial Intelligence Development on Human Rights, Democracy and the Rule of Law) başlığı altında Avrupa Konseyi'nin insan haklarını koruma konusunda evrensel rolünün olduğu, YZ'nin insan hakları, demokrasi ve hukukun üstünlüğüne yönelik risklerini irdelemek gerektiği, ve oyunun mu bizi yöneteceğinin yoksa bizim mi oyunu yöneteceğimizin belirlenmesi gerektiği üzerinde durulmuştur (Council of Europe, 2019).

2.2. Ekonomik İşbirliği ve Kalkınma Teşkilatı'nda (OECD) İnsan Hakları Temelli Yapay Zekâ Düzenlemeleri

Diğer yandan, 22 Mayıs 2019 tarihinde OECD “Yapay Zekâ Konseyi Tavsiyesi”ni (OECD Council Recommendation on Artificial Intelligence) kabul etmiştir. Bu tavsiyeye göre, güvenilir yapay zekânın sorumlu yönetimi için beş tamamlayıcı ilke bulunmaktadır ve OECD üye ülkeleri bu ilkeleri teşvik etmeli ve uygulamalıdır (OECD Legal Instruments, 2019). Bu ilkeler, “kapsayıcı büyüme, sürdürülebilir kalkınma ve refah”, “insan merkezli değerler ve adalet”, “şeffaflık ve açıklanabilirlik”, “sağlamlık, güvenlik ve emniyet”, ve “hesap verebilirlik”tir. Haziran 2019’da gerçekleşen G20 zirvesinde de kabul edilen bu ilkeler ışığında (G20 Ministerial Statement on Trade and Digital Economy-Annex: G20 AI Principles), yapay zekâ aktörlerinin tüm yapay zekâ döngüsü boyunca hukukun üstünlüğüne, hürriyet, haysiyet ve özerklik, mahremiyet ve veri koruma, ayrımcılık yapmama ve eşitlik, çeşitlilik, adalet, sosyal adalet ve uluslararası kabul görmüş işçi hakları gibi- insan haklarına ve demokratik değerlere saygı göstermesi gerektiği vurgulanmıştır (OECD, 2019).

2.3. Avrupa Birliği'nde İnsan Hakları Temelli Yapay Zekâ Düzenlemeleri

AB, yapay zekânın etik unsurlarının belirlenmesi konusunda ilk adımı Mart 2018 tarihinde atmış, “Yapay Zekâ Uzman Grubu” kurulmasına karar vermiştir. Sonrasında, Nisan 2018’de Komisyon yapay zekâyla ilgili ilk strateji belgesini yayınlamıştır (Artificial Intelligence for Europe) (European Commission, 2018). Haziran 2018’de “Avrupa Yapay Zekâ İttifakı”, Avrupa Komisyonu tarafından yapay zekâ forumu olarak kurulmuştur (European AI Alliance) (European Commission, 2022). Bu forumda, AB yapay zekâ stratejisi çerçevesinde sivil toplum, akademisyenler, sendikalar ve çeşitli uzmanlar bir araya gelerek güvenilir yapay zekâ stratejilerini tartışmakta ve geleceğe yönelik tavsiyeler oluşturmaktadırlar (Yılmaz vd., 2021:453).

Nisan 2019’da, “Güvenilir Yapay Zekâ için Etik İlkeler Kılavuzu” (Ethics Guidelines for Trustworthy AI) yayınlanmıştır (European Commission, 2019). Bu kılavuza göre, güvenilir YZ’nin üç bileşeni vardır; tüm yasalara uygun olmalı, tüm etik ilke ve değerlere uygun olmalı ve hem teknik hem de sosyal açılarından güçlü olmalıdır (European Commission, 2019). Her bileşen tek başına önemli olsa da birlikte uyum içinde işlemeli, eğer çatışma çıkarsa da toplumun çabasıyla sorun çözülmelidir. Komisyon, yapay zekâ sistemleri için gerekli yedi temel ilkeyi belirlemiştir. Bunlar; (1) insan etkinliği ve gözetim, (2) teknik sağlamlık ve emniyet, (3) gizlilik ve veri yönetimi, (4) şeffaflık, (5) çeşitlilik, ayrımcılık yasağı ve adil yargılanma, (6) çevresel ve toplumsal iyilik ve (7) hesap verebilirliktir (European Commission, 2019). Bu ilkeler aynı zamanda yapay zekâ ve dijital etiğin en temel sorunlarına işaret etmektedir.

Makalenin konusu gereği, “Güvenilir Yapay Zekâ için Etik İlkeler Kılavuzu”nda Komisyon’un ayrımcılık yasağının içeriğini genişletme çabası görülmektedir. Yani yapay zekâ uygulamalarında, “tüm insanların tüm ahlaki değer ve itibarına eşit saygı duyulmak zorunda” olduğunun altı çizilmiştir. Komisyon bu maddenin ayrımcılık yasağı maddesinin uygulama alanının üzerine çıktığını, yani genişlettiğini iddia etmektedir. Yapay zekâ bağlamında eşitlik, yanlı/ön yargılı çıktılar yaratmayan uygulamalar gerektirir. Bu durumda, YZ işçiler, kadınlar, engelli bireyler, etnik azınlıklar, çocuklar, tüketiciler ya da dışlanma riski altında olan diğer dezavantajlı kişi veya gruplara yeterli saygı göstermelidir.

YZ sistemlerinin eğitim ve operasyon için kullandığı data setleri, kasıtsız tarihsel ön yargı, eksiklik ve kötü yönetim modelleri yüzünden direkt ya da dolaylı ön yargı ve ayrımcılığa neden olabilmekte. Bunu önlemek için, tanımlanabilir ve ayrımcı ön yargı, mümkünse veri toplama aşamasında ortadan kaldırılmalıdır (European Commission, 2019).

Ayrıca Kılavuz'da, YZ aracılığıyla yapılan herhangi bir tür vatandaş puanlamanın (citizen scoring) temel insan haklarının ve özellikle ayrımcılık yasağının ihlaline neden olduğu belirtilmektedir. Okul sistemleri, elektronik eğitim ve ehliyetlerde vatandaş puanlama sistemi uygulansa da bu sürecin şeffaf bir şekilde tüm vatandaşlara açık olarak yürütülmesi gerektiği belirtiliyor. Ancak unutulmamalıdır ki sürecin şeffaflığı ayrımcılık uygulamalarının önüne geçemiyor (European Commission, 2019).

Şubat 2020'de Komisyon, tanımlanan etik sorunların nasıl ele alınacağına ilişkin detaylı yönlendirmeler içeren “Yapay Zekâ Üzerine Beyaz Kitap- Mükemmellik ve Güven Üzerine Avrupa Yaklaşımı” (White Paper on Artificial Intelligence-A European Approach to Excellence and Trust) başlıklı raporunu yayınlamıştır (European Commission, 2020).

Beyaz Kitap öncelikle, dijital teknolojilerin güvenilir olması gerekliliğini ön koşul olarak belirlemiştir (Yılmaz vd., 2021:454). Beyaz Kitap, YZ'nin opak karar alma, toplumsal cinsiyete dayalı veya diğer tür ayrımcılıklar gibi potansiyel risklerinin olduğunu belirtmektedir (European Commission, 2020). Beyaz Kitap'ın amacı, YZ'nin güvenilir ve güvenli bir şekilde gelişimi için tüm AB vatandaşlarının tüm değer ve haklarına tam saygılı politika opsiyonları sunmaktır. Komisyon ayrıca YZ uygulamalarının temel haklara saygılı olma zorunluluğunun altını çizmekte ve bu değerleri dünyaya yaymak için çaba sarf edeceğini bildirmektedir.

Ayrıca Beyaz Kitap'ta YZ kullanımının AB'nin üzerine kurulduğu değerleri³ etkileyebileceğinden ve özel hayatın, kişisel verilerin ve tüketicinin korunması alanında ifade özgürlüğü, toplanma özgürlüğü, insan haysiyeti, ayrımcılık yasağı, adil yargılanma gibi temel hakların ihlal edebileceğinden bahsedilmektedir (European Commission, 2020). Bu riskler, yapay zekâ sistemlerinin genel tasarımındaki kusurlardan veya verilerin olası ön yargıları düzeltmeden, mesela sadece erkeklerden elde edilen verilerin, kullanılmasından kaynaklanabilir.

Beyaz Kitap, YZ daha önce yalnızca insanlar tarafından yapılabilen birçok işlevi yerine getirebildiği için artık bireyler ve tüzel kişilerin gittikçe daha fazla YZ karar ve eylemlerine maruz kalacağını iddia etmektedir. Ayrıca YZ'nin “devlet yetkilileri veya diğer kuruluşlar tarafından toplu gözetim için ve işverenler tarafından çalışanlarının nasıl davrandığını gözlemlemek için” ya da “çevrimiçi araçlar tarafından kullanıcıları için bilgilere öncelik vermek ve içerik denetimi gerçekleştirmek için” kullanılma riskini de vurgulamaktadır (European Commission, 2020). Son olarak Komisyon, YZ karar ve eylemlerinin AB yasalarına uygun şekilde gerçekleştirilip gerçekleştirilmediğine dair doğrulama mekanizmalarından yoksun olduğunu, bu nedenle bu tür kararlardan olumsuz etkilenen birey ve tüzel kişilerin “adalete etkin erişim konusunda zorluklarla karşılaşabileceğini” değerlendirmektedir (European Commission, 2020). Nitekim, “yargılama sürecinde kullanılan bir algoritmanın sonucu, doğruluğundan şüphe duysa dahi hakimin aksi yönde karar vermesini engelleyici psikolojik bir güce” sahip olacaktır (Yılmaz vd., 2021:455-6).

Son olarak, Avrupa Birliği Komisyonu 21 Nisan 2021 tarihinde “Yapay Zekâya İlişkin Uyumlu Kurallar Koyulması ve Bazı Avrupa Birliği Düzenlemelerinin Değiştirilmesi” (Artificial Intelligence Act-Fostering a European approach to Artificial Intelligence) yasasını teklif etmiştir (European Commission, 2021). Teklif edilen Yapay Zekâ Yasası, YZ sistemlerinin özel kullanımı ve ilgili riskler hakkındaki yasal çerçeveye odaklanmakta, ve tüm AB ülkelerinde ve YZ sistemlerini AB pazarına sokan üçüncü ülkelerde kurulan ve kullanılan tüm YZ sistemleri için geçerli olan yatay AB yasama

³ İnsan Haysiyeti, Özgürlük, Demokrasi, Eşitlik, Hukukun Üstünlüğü, İnsan Hakları; bkz. madde 2 AB Antlaşması, (Official Journal of the European Union, 2016).

aracı olmayı hedeflemektedir (European Commission, 2021). Opaklık, karmaşıklık, verilere bağımlılık ve özerk davranış özellikleri nedeniyle YZ kullanımı birçok temel hak ve kullanıcı güvenliğini olumsuz yönde etkilemektedir. Bu nedenle önerilen YZ Yasası, somut risk düzeyine göre yasal müdahaleyi mümkün kılan risk temelli bir yaklaşımı izlemektedir. Bu yaklaşıma göre, YZ sistemlerini (i) kabul edilemez risk, (ii) yüksek risk, (iii) sınırlı risk ve (iv) düşük risk düzeylerine göre sınıflandırmaktadır. Bu durumda, YZ uygulamaları yalnızca belirli risk düzeylerini ele alacak şekilde düzenlenecektir (European Commission, 2021). Yasa aynı zamanda, üye devlet temsilcileri ve Avrupa Komisyonu temsilcilerinden oluşan bir Avrupa Yapay Zekâ Kurulu kurmayı önermekte, ve bu kurulun yeni kuralların uyumlu bir şekilde uygulanmasını ve ulusal denetim birimleri ve Komisyon arasındaki işbirliğini sağlaması gerektiğini ileri sürmektedir. Yasaya uymayan üye devletlere ihlal seviyesine göre 30 milyon Euro'ya kadar veya dünya çapındaki toplam yıllık cironun %6'sı kadar idari para cezası verilebileceğini söylemektedir (European Commission, 2021).

3. AVRUPA'DA YAPAY ZEKÂ UYGULAMALARININ İNSAN HAKLARI ÜZERİNDEKİ ETKİLERİ VE ROBOT HAKİM UYGULAMALARI

Günümüzde YZ askeri, yargı, sağlık ve ticaret alanları başta olmak üzere pek çok alanda kullanılmaktadır, bu yüzden uluslararası antlaşmaların sağladığı insan hakları olumlu veya olumsuz şekilde etkilenmektedir (Singil, 2022:10). Burada önemli olan nokta, YZ uygulamalarının hangi bireyler ya da gruplar üzerinde hangi insan hakları ihlallerine neden olduğunu ortaya çıkarabilmektir. Dolayısıyla, bu tür insan hakları ihlallerinin engellenmesi için Avrupa Konseyi, OECD ve AB tarafından atılan adımların etkin bir şekilde uygulanması ve daha da önemlisi üye ülkelerde yasalaşması gerekmektedir.

YZ kapsamında toplumsal cinsiyet eşitliğini sağlamak için UNESCO ortaklığıyla 2017'de üretilmiş Pakistan'ın özellikle kırsal bölgelerindeki kadınlarla “üreme sağlığı, hijyen ve güvenlik” hakkında konuşabilen YZ sohbet robotu “RAAJI” gibi olumlu örnekler olmakla beraber (Singil, 2022:11), makalenin konusu gereği bu bölümde daha çok YZ kaynaklı insan hakları ihlallerine odaklanılacaktır. Bu noktada, özellikle algoritmalar oluşturulurken insan haklarına dayalı YZ sistemleri tercih edilmeli, çünkü ön yargılı ya da ayrımcı veriler içerdiği takdirde bireylerin ya da birçok dezavantajlı grubun din, dil, siyasi görüş, etnik köken gibi birçok konuda ayrımcılığa uğrayacağı açıktır.

Avrupa'da robot hakim uygulamaları ve hakim için karar tahminleri yapmak üzere geliştirilen YZ örnekleri Almanya, Hollanda, Fransa, Estonya, İngiltere ve Amerika'da görülmektedir (Abanoz, 2020:197). Öncelikle Amerika'da karar tahmininde bulunan YZ sistemleri denenmeye başlamıştır. Bu sistemler, “davaya bakmakta olan hâkimin, davayı hangi hukukî gerekçeye dayanarak hangi olasılığa göre karara bağlayacağını” çok sayıda karar inceleyerek tüme varımsal olarak tahminde bulunurlar (Bilgin, 2022:410). ABD Yüksek Mahkemesi'nin 1816-2015 yılları arasında verdiği yaklaşık 240 bin yargıç oyu ve 28 bin dava sonucu YZ ile tahmin edilmeye çalışılmış; yargıç oylarının %71,9'u ve dava sonuçlarının da %70,2'si doğru bulunmuştur (Abanoz, 2020:197).

Diğer yandan Almanya'da Alman Federal Anayasa Mahkemesi kararları incelenerek “Argumentum” adlı bir YZ projesi geliştirilmiş, bu proje ile yargılama sürecinde görev alanların argüman yapıları tahmin edilmeye çalışılmıştır (Sümer, 2021:1563). Ayrıca, yakın gelecekte yargı sürecinde YZ uygulamalarına geçilmesi planlandığı için “Almanya'da tüm mahkemelerin 1/1/2026 tarihine kadar elektronik dava dosyalarına geçmesi gerekliliği vurgulanmıştır” (Bilgin, 2022:414).

Hollanda'da dolandırıcılık suçlarıyla mücadele için bir YZ uygulaması geliştirilmiş, bu uygulamayla geçmişte bu suçu işleyen kişilerden “risk profilleri oluşturarak “benzer” vatandaş profillerini taramaktadır” (Sümer, 2021:1564). Ancak bu uygulamanın algoritması, daha çok düşük gelirli ve azınlık bireyleri seçtiğinden ayrımcılık yaratması nedeniyle eleştirilmiştir. Ayrıca yine Hollanda'da 2016 yılında kurulan çevrimiçi mahkeme, “tam otomatik bir uyuşmazlık çözüm hizmeti sunmaktadır” (Bilgin, 2022:414). Yargı maliyetlerini de oldukça düşüren bu sistem, uyuşmazlık çözümü konusunda bireyleri memnun etmektedir.

İngiltere'de çevrimiçi “Trafik Ceza Mahkemesi”, İngiltere ve Galler'de verilen trafik cezalarına itiraz etmek için kurulmuş bir nevi temyiz merciidir (Sümer, 2021:1564). Fransa'da iki temyiz mahkemesinin işten çıkarma, boşanma ve komşuluk konularından doğan uyuşmazlıklarla ilgili kararları 2017 yılında YZ ile test edilmiştir, ancak Adalet Bakanlığı hakimlerin karar verme sürecine katkı sunmadığını düşünerek uygulamayı kaldırmıştır (Bilgin, 2022:413). Estonya'da ise YZ kullanılarak “uyuşmazlık değeri 7.000-Euro'dan az olan özel hukuk temelli sözleşmeden kaynaklı uyuşmazlıkları”n çözülmesi hedeflenmiştir (Bilgin, 2022:413). Hâkim kararına yardım amaçlı oluşturulan en son YZ uygulaması 15 Şubat 2023'te duyurulan, sözleşme analizi ve durum tespiti yapabilen Harvey'dir (Karabel & Aydemir, 2023:558). Bu uygulamanın ne kadar başarılı olduğu yakın zamanda anlaşılacaktır.

3.1. Avrupa İnsan Hakları Mahkemesi'nde (AİHM) Robot Hakimler/Yapay Zekâ Kararları

AİHM'ye her sene ortalama 45 bin dava başvurusu yapılmakta ve yargıçların her geçen gün artan iş yükü nedeniyle de yeni çözümler üretilmeye çalışılmaktadır (Euronews, 2023). YZ'nın insan hakları alanında denenmesi de tam da bu nedenden dolayı ortaya çıkmıştır.

Ulenaers (2020:6)'a göre, 2016 yılında işkence yasağı, adil yargılanma hakkı ve özel ve aile yaşamına saygı hakkına ilişkin olarak yaklaşık 600 AİHM kararı yapay zekâ sistemi tarafından analiz edilmiş ve bu hakların ihlal edilip edilmediğine dair verilen yapay zekâ kararları ile AİHM yargıçlarının verdiği kararlar %79 uyum göstermiştir. Robot hâkim, “davada geçen ifadelere, kullanılma sıklığına ve yoğunluğuna bakarak” AİHS'nde yer alan herhangi bir insan hakkının ihlal edilip edilmediğine karar vermiştir (Abanoz, 2020:197). 2020 yılında ise Avrupa İnsan Hakları Sözleşmesi'nde (AİHS) kabul edilen dokuz hakka yönelik yaklaşık 1950 adet AİHM kararı YZ ile analiz edildiğinde uyum oranı %75 çıkmıştır (Abanoz, 2020:197). Bu yüzdelerle bakıldığında AİHM kararlarında YZ, hızlı bir şekilde karara varmak için yargıçlara yararlı bir yardımcı araç olabilir. Özellikle emsal teşkil eden AİHM kararları YZ sistemine tanımlanır, kümelendirilir ve öncelik sırasına koyulursa, algoritma yeni verilerle beslenirse ve yargıçlar son zamanlardaki davalarda ortaya çıkan güncel eğilimlerden haberdar edilirse YZ insan hakları davalarında kullanılabilir.

Yine de bir insan hakkı ihlali olup olmadığına karar vermek gibi çok önemli bir görevi olan AİHM'nin YZ'yı kullanıp kullanamayacağı ya da hangi sınırlarda kullanılabileceği konusu henüz kolayca kabul edilebilecek ve uygulanabilecek gibi gözükmemektedir.

3.2. Yapay Zekâ Uygulamalarının Ayrımcılık Yasağı Üzerindeki Etkileri: Algoritmik Ayrımcılık

AİHS'nin 14. maddesi ve 12 No'lu Ek Protokolü'nde düzenlenen ayrımcılık yasağı, hukuken tanınmış tüm hakların, “cinsiyet, ırk, renk, dil, din, siyasi veya diğer kanaatler, ulusal ve sosyal köken, ulusal bir azınlığa mensup olma, servet,

doğum veya herhangi bir diğer statü bakımından hiçbir ayrımcılık yapılmadan” sağlanmasını öngörür (Avrupa Konseyi, 1950).

AB Temel Haklar Şartı madde 21’de ise ayrımcılık yasağı, “cinsiyet, ırk, renk, etnik veya toplumsal köken, genetik özellikler, dil, din veya inanç, siyasi veya başka konulardaki görüşler, ulusal bir azınlığın üyesi olmak, mal, doğum, engellilik/özürlülük, yaş, cinsel eğilimler gibi hususlara dayanılarak ayrımcılık yapılamaz” (Official Journal of the European Union, 2012) şeklinde geçer.

Son yıllarda insan hakları ve Avrupa çalışmaları literatüründe “Algoritmik Ayrımcılık” tartışılmaya başlanmış, ayrıca algoritmaların, makine öğrenmesinin ve otomatik karar vermenin neden olduğu ayrımcılığa karşı korunma hakkı geliştirilmiştir.

YZ uygulamalarının ayrımcılığa yol açmasının nedeni genellikle ön yargılı algoritmalara dayanır. Ön yargılı algoritmaların kullanılması, özellikle dezavantajlı gruplar üzerinde ayrımcılık riski yaratmaktadır. Bu tür riskleri önlemek için veri etiğinin geliştirilmiş olması gerekmektedir.

Schwartz ve Ulbricht (2022), ABD adalet sisteminde sanıkların 2 yıl içinde tekrar suç işleme riski hakkında yargı personelini bilgilendirmek için 1998 yılından beri kullanılan “Alternatif Yaptırımlar için Düzeltici Suçlu Yönetim Profili Oluşturma” (COMPAS) yazılımını incelemiş ve COMPAS puanlarının siyahlar için ön yargılı olduğunu iddia etmişlerdir. Boyd vd. (2014:54) işverenin algoritmik karar verme sürecinde -LinkedIn gibi bir mecra aracılığıyla mesela- kişisel ağını kullanarak işe alım yaptığını ve böylece ayrımcılığa neden olduğunu iddia eder.

Xenidis ve Senden (2020) de 2016 yılında Microsoft tarafından üretilen Tay.AI adlı sohbet robotunun (chatbot) ırkçı ve cinsiyetçi söylemleri nedeniyle 24 saat içinde geri çekildiğini söylemektedir. Bir yandan klişeleştirme ve ön yargı, grupların toplumdaki eşit temsili etkilerken; diğer yandan geçmişteki -zamanla kurumsallaşmış ve somutlaştırılmış- ayrımcılıklar yapısal eşitsizlik biçimlerine yansır (Xenidis & Senden, 2020:6). Böylece iki çeşit yeniden üretilen eşitsizlik mekanizması algoritmik ayrımcılığın en temel kaynaklarıdır.

Diğer yandan, MIT (Massachusetts Institute of Technology) projesi olan “Kodlanmış Bakış” (The Coded Gaze) ortaya çıkarmıştır ki siyah kadınların yetersiz temsil edildiği eğitim verileri, siyah/beyaz erkek yüzlerini ve beyaz kadın yüzlerini tanıma konusunda çok daha iyi performans gösterirken siyah kadınların yüzlerini tanıma konusunda daha az performans gösteren bir yüz tanıma yazılımına neden olmaktadır (Xenidis & Senden, 2020:8).

Amazon’un kullandığı işe alma yazılımının kadınlara karşı ön yargılı olduğu ve erkek başvuranları işe alımda tercih ettiği ortaya çıkmıştır (Zuiderveen Borgesius, 2018:25). Aynı YZ uygulamasının siyah kadınlara karşı ön yargı geliştirmesi nedeniyle geri çekildiği bilinmektedir (Kriebitz & Lütge, 2020:15). Diğer yandan, 2013 yılında Google’a Afro-Amerikan isimler yazıldığında siyah bir bireyin tutuklama kaydını gösteren reklamlar çıkarmaktaydı, bu da Google’ın YZ sisteminin sahip olduğu ırkçı ön yargıyı ortaya koymaktadır (Zuiderveen Borgesius, 2018:26). Yine 2016 yılında Google’a “üç siyah ergen” yazıldığında sabıka fotoğrafı çıkarken “üç beyaz ergen” yazıldığında mutlu beyaz çocukların fotoğraflarını göstermekteydi (Zuiderveen Borgesius, 2018:28). Ayrıca yine Google’ın, cinsiyetini kadın ya da erkek olarak bildiren kişiler karşılaştırıldığında erkeklere daha yüksek maaşlı işlerin reklamlarını gösterdiği ve kadınların bu nedenle ayrımcılığa uğradığı bilinmektedir (Zuiderveen Borgesius, 2018:26).

Çalışkan vd.'ne göre (2017), Google Çeviri'ye "He is a doctor. She is a nurse." Yazıp Türkçe'ye çevirmek istendiğinde "O bir doktor. O bir hemşire." şeklinde çevirir çünkü Türkçe'de "he" ve "she" kelimelerinin karşılığı yoktur, yani "o" kelimesi cinsiyet ayrımı yapmamaktadır. Ancak bu Türkçe cümleler tekrar İngilizce'ye çevrilmek istendiğinde Google Çeviri yine "He is a doctor. She is a nurse." yazar. Yani, hemşireye kadın, doktora erkek cinsiyetini yüklemiştir.

Tüm bu örnekler toplum içindeki dezavantajlı grupların YZ tarafından da görmezden geldiğini ortaya koymaktadır. Mesela dünyaca ünlü teknoloji markası Huawei hakkında 2021 yılında Uygur Türkleri'ni algılayan YZ tekniklerini geliştirdiği haberleri çıkmıştır (Kelion, 2021). Huawei hızlıca bu haberleri yalanlasa da YZ'nin hangi amaçlar için kullanılabilmesi ihtimalini göstermesi açısından önemlidir.

SONUÇ

Günümüzde dijital yenilikler geliştikçe, dijital etik gereksinimini ve dolayısıyla insan haklarına saygıyı da beraberinde gerektiren bir bilgi toplumu ortaya çıkmaktadır. Bu bağlamda YZ uygulamaları insan hakları ihlallerine ve ayrımcılığa yol açabilmektedir. Bunun temel nedeni genellikle ön yargılı algoritmalarla dayanır. Dijital etiğin geliştirilmesi tam da bu noktada gereklidir. Yüz tanıma teknolojisinin ya da işe alma yazılımlarının kullanımı insan haklarını, müşteri bilgilerine ulaşan YZ teknikleri de kişisel bilgilerin korunmasını ihlal etmektedir. Diğer yandan YZ sosyal medya programlarında kullanılan algoritmalarla kişilerin yaptığı paylaşımlara ulaşırsa, kişinin ifade özgürlüğünü ihlal etmiş olmaktadır.

YZ uygulamalarından kaynaklanan insan hakları ihlallerini ve algoritmik ayrımcılığı engellemek için öncelikle uluslararası yasal standart yaratılmalı ve uluslararası insan hakları standartlarına uygun YZ uygulamaları tercih edilmelidir. Robot hâkimin karar verme sürecinde ortaya çıkabilecek algoritmik sorunlardan kaynaklanan ön yargılı karar verme ihtimali ve toplumu verilen kararlar konusunda tatmin edememe olasılığı, yargıda YZ uygulamalarının en temel sorunlarıdır.

Bu noktada, 2021 tarihli AB YZ Yasası Teklifi çok önemlidir. Bu teklifin yasalaşması, tüm AB üye ve aday ülkeleri tarafından kabul edilip uygulanması ve uluslararası YZ yasal standardına ışık tutması gerekmektedir. 2030'a kadar AİHM'deki yargıçlar dâhil istihdamın %47'sinin otomatikleştirilebileceği düşünüldüğünde YZ her alanda kullanılacaktır.

Bu çalışmanın özgün değeri, YZ kaynaklı insan hakları ihlallerini önlemek için Avrupa'da YZ etiğinin oluşturulma sürecine odaklanmasındadır. Sonuç olarak, Avrupa'da YZ etik kurallarının oluşturulması, yaygınlaştırılması ve uygulanması süreci tamamlanmadıkça YZ kaynaklı insan hakları ihlallerinin devam edeceği kanısına varılmaktadır.

ETİK BEYAN VE AÇIKLAMALAR

Etik Kurul Onay Bilgileri Beyanı

Bu çalışma etik onay gerektiren herhangi bir insan veya hayvan araştırması içermemektedir.

Yazar Katkı Oranı Beyanı

Yazarlar tüm çalışmalarını birlikte yürütmüştür.

Çıkar Çatışması Beyanı

Çalışma ile ilgili herhangi bir kurum veya kişi ile çıkar çatışması bulunmamaktadır.

KAYNAKÇA

- Abanoz, B. (2020). Hukuki Bilgiye Dijital Erişim. *Marmara Üniversitesi Hukuk Fakültesi Hukuk Araştırmaları Dergisi*, 26(1), 190-203. DOI: 10.33433/maruhad.678020
- Avrupa Konseyi. (1950). *Avrupa İnsan Hakları Sözleşmesi*. Strasbourg: Avrupa İnsan Hakları Mahkemesi. Erişim adresi: https://www.echr.coe.int/documents/convention_tur.pdf, Erişim tarihi: 15.11.2022.
- Kelion, L. (2021, 13 Ocak). Huawei, Çin'de Sokakta Yürüyen Uygurların Tespitini Sağlayan Teknolojinin Patenti İçin Başvuru Yaptı. *BBC News Türkçe*. Erişim adresi: <https://www.bbc.com/turkce/haberler-dunya-55647839>, Erişim tarihi: 15.03.2021.
- Becker, S. J., André, T.N., Lucas, S., Heinitz, R. M., Klevesath, M. & Charton, J. E. (2022). A Code of Digital Ethics: Laying the Foundation for Digital Ethics in a Science and Technology Company. *AI & SOCIETY*, 38(11), 2629-2639.
- Bilgin, H. (2022). Yapay Zekânın Mahkeme Kararlarında Kullanımına Uluslararası Bir Bakış ve Robot Hâkimler Hakkında Düşünceler. *İnönü Üniversitesi Hukuk Fakültesi Dergisi*, 13(2), 405-419.
- Boyd, D., Levy, K. & Marwick, A. (2014). *The Networked Nature of Algorithmic Discrimination*. Open Technology Institute, New America, 53-57.
- Bynum, T. (2015). *Computer and Information Ethics*. The Stanford Encyclopedia of Philosophy. Erişim adresi: <http://plato.stanford.edu/archives/win2015/entries/ethics-computer/>, Erişim tarihi: 15.12.2022.
- Council of Europe. (2018). *CEPEJ European Ethical Charter on the Use of Artificial Intelligence (AI) in Judicial Systems and Their Environment*. Erişim adresi: <https://www.coe.int/en/web/cepej/cepej-european-ethical-charter-on-the-use-of-artificial-intelligence-ai-in-judicial-systems-and-their-environment>, Erişim tarihi: 15.11.2022.
- Council of Europe. (2019). *High-Level Conference: Governing the Game Changer—Impacts of Artificial Intelligence Development on Human Rights, Democracy and the Rule of Law*. Helsinki, Finland. Erişim adresi: <https://rm.coe.int/conference-report-28march-final-1-/168093bc52>, Erişim tarihi: 15.11.2022.
- Çalışkan, A., Bryson, J. J. & Narayanan, A. (2017). Semantics Derived Automatically from Language Corpora Contain Human-Like Biases. *Science*, 356(6334), 183-186.
- Euronews. (2023). *AİHM'e Geçen Yıl Yapılan Başvuruların Yüzde 26,9'u Türkiye'den*. Erişim adresi: <https://tr.euronews.com/2023/01/26/aihme-gecen-yil-yapilan-basvurularin-yuzde-269u-turkiyeden>, Erişim tarihi: 26.01.2023.
- European Commission. (2018). *Artificial Intelligence for Europe*. COM(2018) 237 Final. Brussels. Erişim adresi: <https://eur-lex.europa.eu/legal-content/EN/TXT/PDF/?uri=CELEX:52018DC0237&from=EN>, Erişim tarihi: 15.11.2022.
- European Commission. (2019). *Ethics Guidelines For Trustworthy AI High-Level Expert Group on Artificial Intelligence*. B-1049 Brussels. Erişim adresi: <https://digital-strategy.ec.europa.eu/en/library/ethics-guidelines-trustworthy-ai>, Erişim tarihi: 15.11.2022.
- European Commission. (2020). *White Paper On Artificial Intelligence-A European Approach to Excellence and Trust*. COM(2020) 65 Final. Brussels. Erişim adresi: https://ec.europa.eu/info/sites/default/files/commission-white-paper-artificial-intelligence-feb2020_en.pdf, Erişim tarihi: 15.11.2022.
- European Commission. (2021). *Proposal for a Regulation of The European Parliament and of the Council Laying Down Harmonised Rules On Artificial Intelligence (Artificial Intelligence Act) and Amending Certain Union Legislative Acts*. {SEC(2021) 167 Final} - {SWD(2021) 84 Final} - {SWD(2021) 85 Final}. COM(2021) 206 Final. 2021/0106 (COD). Brussels. Erişim adresi: https://eur-lex.europa.eu/resource.html?uri=cellar:e0649735-a372-11eb-9585-01aa75ed71a1.0001.02/DOC_1&format=PDF, Erişim tarihi: 15.11.2022.
- European Commission. (2022). *The European AI Alliance*. Erişim adresi: <https://digital-strategy.ec.europa.eu/en/policies/european-ai-alliance>, Erişim tarihi: 15.11.2022.
- Floridi, L. (2006). Information Ethics, Its Nature and Scope. *SIGCAS Computer Society*, 36(3), 21–36.
- Floridi, L. (2013). *The Ethics of Information*. Oxford: Oxford University Press.

- Floridi, L. (2018). Soft Ethics, the Governance of the Digital and the General Data Protection Regulation. *Philosophical Transactions of the Royal Society A: Mathematical, Physical and Engineering Sciences*, 376(2133), 1-11.
- Floridi, L., Cath, C. & Taddeo, M. (2019). *Digital Ethics: Its Nature and Scope*. The 2018 Yearbook of the Digital Ethics Lab., 9-17.
- Green, B. (2021). The Contestation of Tech Ethics: A Sociotechnical Approach to Technology Ethics in Practice. *Journal of Social Computing*, 2(3), 209-225. doi: 10.23919/JSC.2021.0018.
- Johnson, A. T. (2020). Ethics in the Era of Artificial Intelligence. *IEEE Pulse*, 11(3), 44-47.
- Karabel, E.G. & Aydemir, D. (2023). Medeni Usul Hukukunda Yargılamanın Hızlandırılması ve Adalete Erişim Hakkı Bakımından Çevrimiçi Yargılama ve Yapay Zekanın Kullanımı. *Marmara Üniversitesi Hukuk Fakültesi Hukuk Araştırmaları Dergisi*, 29(1), 530-573. DOI:10.33433/maruhad.1263609
- Khalili, L. & Ghasemzadeh, A. A. (2021). Bibliometric Analysis of Ethics and Technology Research Based on Google Scholar from 1983 to 2020. *International Journal of Ethics and Society*, 3(3), 10-18. Erişim adresi: <http://ijethics.com/article-1-147-en.html>, Erişim tarihi: 01.12.2022.
- Kriebitz, A. & Lütge C. (2020). Artificial Intelligence and Human Rights: A Business Ethical Assessment. *Business and Human Rights Journal*, 1–21. doi:10.1017/bhj.2019.28
- Moor, J. H. (1985). What Is Computer Ethics? *Metaphilosophy*, 16(4), 266–275. <https://doi.org/10.1111/j.1467-9973.1985.tb00173.x>
- Müller, V. C. (2022). *History of Digital Ethics*. In: Carissa Véliz (Ed.), *Oxford Handbook of Digital Ethics* (pp. 3-9). Oxford: Oxford University Press.
- OECD. (2019). *G20 Ministerial Statement on Trade and Digital Economy-Annex: G20 AI Principles*. Erişim adresi: https://g20-digital.go.jp/asset/pdf/g20_2019_japan_digital_statement.pdf, Erişim tarihi: 01.12.2022.
- OECD Legal Instruments. (2019). *Recommendation of the Council on Artificial Intelligence*. Erişim adresi: <https://legalinstruments.oecd.org/en/instruments/OECD-LEGAL-0449#:~:text=The%20OECD%20Council%20adopted%20the,on%2022%2D23%20May%202019.&text=The%20OECD%20Recommendation%20on%20AI,governments%20in%20their%20implementation%20efforts>, Erişim tarihi: 01.12.2022.
- Official Journal of the European Union. (2012). *Charter of Fundamental Rights of the European Union*. C 326/392. Erişim adresi: <https://eur-lex.europa.eu/legal-content/EN/TXT/PDF/?uri=CELEX:12012P/TXT&from=EN>, Erişim tarihi: 01.12.2022.
- Official Journal of the European Union. (2016). *Consolidated Versions of the Treaty on European Union and the Treaty on the Functioning of the European Union*. Erişim adresi: <https://eur-lex.europa.eu/legal-content/EN/TXT/HTML/?uri=CELEX:12016ME/TXT&from=EN#d1e32-13-1>, Erişim tarihi: 14.11.2022.
- Parker, D. B. (1968). Rules of Ethics in Information Processing. *Communications of the ACM*, 11(3), 198–201.
- Rogerson, S. (2017). *Coding Ethics into Technology*. Hack & Craft News. Erişim adresi: <http://hcnnews.com/coding-ethics-technology>. Erişim tarihi: 24.04.2021.
- Rogerson, S. (2020) *Re-Imagining the Digital Age through Digital Ethics*. Erişim adresi: <https://dora.dmu.ac.uk/bitstream/handle/2086/20626/Ethics%20and%20the%20Internet%20Position%20Paper%20Simon%20Rogerson.pdf?sequence=1>. Erişim tarihi: 24.04.2021.
- Schwartz, R. & Ulbricht, L. (2022). Why Organization Matters in “Algorithmic Discrimination”. *Köln Z Soziol (Suppl 1)*, 74, 307-330.
- Siau, K. & Wang, W. (2020). Artificial Intelligence (AI) Ethics: Ethics of AI and Ethical AI. *Journal of Database Management (JDM)* 31(2), 74-87.
- Singil N. (2022). Yapay Zekâ ve İnsan Hakları. *Public and Private International Law Bulletin*, 42(1), 1-38.
- Sümer, S. Y. (2021). Ceza Yargılamasının Geleceği: Robot Hakim. *Dokuz Eylül Üniversitesi Hukuk Fakültesi Dergisi*, 23, 1543-1591.

- Ulenaers, J. (2020). The Impact of Artificial Intelligence on the Right to a Fair Trial: Towards a Robot Judge? *Asian Journal of Law and Economics*, 11(2), 1-38.
- Verma, S. & Garg, N. (2023). The Trend and Future of Techno-Ethics: A Bibliometric Analysis of Three Decades. *Library Hi Tech*. Erişim adresi: <https://www.emerald.com/insight/content/doi/10.1108/LHT-10-2022-0477/full/html>.
- Xenidis, R. & Senden, L. (2020). *EU Non-Discrimination Law in the Era of Artificial Intelligence: Mapping the Challenges of Algorithmic Discrimination*. In: U. Bernitz, X. Groussot, J. Paju & S. A. De Vries (Eds.), *General Principles of EU Law and the EU Digital Order* (pp. 151-182). Alphen aan den Rijn: Kluwer Law International.
- Yılmaz, G. (2020). Yapay Zekânın Yargı Sistemlerinde Kullanılmasına İlişkin Avrupa Etik Şartı. *Marmara Avrupa Araştırmaları Dergisi*, 28(1), 27-55.
- Yılmaz, İ., Sözer, C. & Elver, E. (2021). Yapay Zekâ ile İlgili Güncel Düzenlemeler: Avrupa Birliği ve Amerika Birleşik Devletleri'nde Alman Aksiyonlar Işığında Bir Değerlendirme. *Adalet Dergisi*, 66, 445-469.
- Zuber, N., Kacianka, S. & Gogoll, J. (2022). Big Data Ethics, Machine Ethics or Information Ethics? Navigating the Maze of Applied Ethics in IT. *arXiv:2203.13494 [cs.SE]*.
- Zuiderveen Borgesius, F. (2018). *Discrimination, Artificial Intelligence, and Algorithmic Decision-Making*. Council of Europe, Directorate General of Democracy. Erişim adresi: <https://rm.coe.int/discrimination-artificial-intelligence-and-algorithmic-decision-making/1680925d73>, Erişim tarihi: 14.11.2022.