MÜHF-HAD Marmara Üniversitesi Hukuk Fakültesi Hukuk Araştırmaları Dergisi • Cilt 29, Sayı 2, Özel Sayı, Aralık 2023, ISSN: 2529-0142, ss. 1003-1040 • DOI:10.33433/maruhad.347497

ARAŞTIRMA MAKALESİ
RESEARCH ARTICLE

ChatGPT ve Üretici Yapay Zekâ Modellerinde Mahremiyet ve Güvenliğin Hukuki Boyutu

Legal Aspects of Privacy and Security in ChatGPT and Generative Artificial Intelligence Models

Başak Ozan ÖZPARLAK*, Müge ÇETİN**

ÖZ

Bu çalışma, ChatGPT gibi üretici yapay zekâ modellerindeki mahremiyet ve güvenliğe dair etkileri hukuki açıdan ele alarak; bu konuda ortaya çıkmaya başlayan sorunları ve olası çözüm yollarını tartışmaktadır. ChatGPT, geniş dil modeli (large language model) yapısına sahip olan ve pekiştirmeli öğrenme (reinforcement learning) esası ile sürekli gelişen bir yapay zekâ tabanlı sohbet botudur. Bu sohbet botu, kullanıcılarla etkileşimde bulunarak farklı görevleri yerine getirebilmektedir. Çalışmamızda; ChatGPT ve genel olarak üretici yapay zekâ modellerinin çalışma ilkeleri ve kullanım potansiyeli açıklanarak, bu teknolojinin halihazırdaki ve gelecekteki mahremiyet ve güvenlik riskleri ortaya konulmaktadır. Bu sorunların yürürlükteki ve tasarı aşamasındaki yasal düzenlemeler ışığında tartışılması ve olası çözümlerin belirlenmesi önemlidir. Çalışmamızın ana argümanı, üretici yapay zekâ modellerinin, mahremiyet ve güvenlik risklerinin önceki teknolojilerden daha faklı olduğu, bu nedenle bu risklerin en aza indirgenmesini sağlamanın sadece teknolojik önlemler ile değil etkili yasal düzenlemelerle mümkün olabileceğidir.

Anahtar Kelimeler: Üretici Yapay Zekâ, Yapay Zekâ ve Hukuk, Mahremiyet, Siber Güvenlik Hukuku, Veri Koruma Hukuku, ChatGPT.

ABSTRACT

This article discusses privacy and security concerns in generative artificial intelligence models, such as ChatGPT, from a legal perspective and addresses the emerging problems and possible solutions in this regard. ChatGPT is an artificial intelligence-based chatbot that has a large language model structure and is improving constantly based on reinforcement learning. This chatbot can perform different tasks by interacting with users. In this study, the working principles and usage potential of ChatGPT and generative artificial intelligence models are explained, and the current and future privacy and security risks of this technology are revealed. It is important to discuss these problems in the light of current and draft legal regulations and to identify possible responses. The main argument of our study is that the privacy and security risks of generative artificial intelligence models are different from previous technologies, and therefore, minimizing these risks is possible not only through technological measures but also through effective legal regulations.

Keywords: Generative Artificial Intelligence, Artificial Intelligence and Law, Privacy, Cyber Security Law, Data Protection Law, ChatGPT.

* Dr. Öğretim Üyesi, Özyeğin Üniversitesi Hukuk Fakültesi, Bilişim Hukuku Ana Bilim Dalı, E-posta: basak.ozparlak@ozyegin.edu.tr, ORCID: https://orcid.org/0000-0002-6861-4290.

** Dr. Öğretim Üyesi, Özyeğin Üniversitesi Hukuk Fakültesi, Bilişim Hukuku Ana Bilim Dalı, E-posta: muge.cetin@ozyegin.edu.tr, ORCID: https://orcid.org/0000-0002-9825-3262.

Sorumlu Yazar/Correspondence Author: Başak Ozan ÖZPARLAK

E-posta/E-mail: basak.ozparlak@ozyegin.edu.tr

Geliş Tarihi/Received: 21.08.2023 **Kabul Tarihi/Accepted:** 13.11.2023

"Bu makineye verebildiğim güce ben kendim de hayret ediyorum. Bir yıl önce, böyle bir sonucun mümkün olabileceğine kesinlikle inanmazdım"[1]. Charles Babbage

GİRİŞ

30 Kasım 2022 bilim ve insanlık tarihinde önemli bir kırılma noktası olarak kabul edilebilir. Bu tarihte, OpenAI[2] şirketi tarafından geliştirilen ve insanlar ile diyalog yolu ile etkileşim sağlayan geniş dil modeline dayalı ChatGPT *(Generative Pre-Trained Transformer*[3]), dünya kamuoyuna duyurulmuştur[4]. ChatGPT, kullanıcıların tüm sorgularını dil yapısını anlayarak[5] ayrıntılı cevap vermek ve pek çok farklı görevi yerine getirmek için tasarlanmıştır. Bununla birlikte ChatGPT; insan dilini öğrenmek ve çok yönlü yapay zekâ asistanları olarak insanlarla etkileşimde bulunmak amacıyla kullanılabileceği gibi[6], bugüne dek sadece insanlar tarafından yapılabilmiş olan şiir ve hikaye yazmak gibi yaratıcılık gerektiren eylemleri ya da kod üretmek gibi görevleri yerine getirebilmektedir. OpenAI tarafından ChatGPT ile ilgili teknik raporda ifade edildiği üzere; bu model henüz kusursuz

1 Babbage'ın bu ifadesi Menabrea'nın Analitik Makine üzerine yazdığı makale tercümesinde yer verdiği Babbage'ın mektuplarında geçmektedir. Luigi Federico Menabrea, 'Sketch of the Analytical Engine Invented by Charles Babbage, Esq.' in Richard Taylor (ed), *The Transactions of Foreign Academies of Science and Learned Societies* (Taylor and Francis 1843) 667.

2 Open AI şirketi, yapay genel zekânın şeffaflık olmadan üretilmesinden ortaya çıkabilecek risklere karşı sonradan yönetim kurulundaki görevinden ayrılan Elon Musk ve şirketin mevcut CEO'su Sam Altman'ın da içinde bulunduğu bir grup teknoloji girişimcisi tarafından 2015 yılında kurulmuştur. Şirketin kurucularından Sam Altman, günümüzdeki temel hedeflerinin yapay genel zekânın (artificial general intelligence – AGI) tüm insanlığa faydalı olmasını sağlamak olduğunu belirtmektedir. Sam Altman, 'Planning For AGI and Beyond' (24 Feb 2023) <https://openai.com/blog/planning-for-agi-and-beyond> accessed 05 July 2023. Microsoft şirketinin ilki 2019 yılında olmak üzere OpenAI şirketine yaptığı yüksek bütçeli yatırımlar ve bulut servisi Azure'u OpenAI şirketinin bulut servisi olarak belirlemesi ile iki şirket arasındaki yapay zekâ alanındaki işbirliği giderek artmıştır. Microsoft, 'Microsoft and OpenAI Extend Partnership' (Microsoft 23 Jan 2023), <https://blogs.microsoft.com/blog/2023/01/23/microsoftandopenaiextendpartnership/> accessed 22 Jul 2023. Microsoft, 2023 yılında OpenAI şirketine yaklaşık 10 milyon Amerikan doları yatırım yapmıştır. Cade Metz and Karen Weise, 'Microsoft to Invest $10 Billion in OpenAI, the Creator of ChatGPT' <https://www.nytimes.com/2023/01/23/business/microsoft-chatgpt-artificial-intelligence.html>, accessed 22 Jul 2023.

3 "Generative Pre-Trained Transformer" ifadesinin Türkçe karşılığı "üretici, önceden eğitilmiş transformer" olarak verilebilir. Terimde geçen transformer, özellikle doğal dil işlemenin temeli olan derin öğrenme ağlarında dikkat mekanizması oluşturulmasında kullanılan bir ağ mimarisidir. Ayyüce Kızrak, 'Nesne Algılama Yaklaşımlarına Transformer Devrimi' (Medium, 10 April 2021) <https://ayyucekizrak.medium.com/nesne-alg%C4%B1lama-yakla%C5%9F%C4%B1mlar%C4%B1na-transformer-devrimi-baf583a29a23> accessed 22 Jul 2023. GPT kavramı, bir tür yapay zekâ modelini ifade etmektedir. GPT modellerine dayanan yapay zekâ sistemleri; belirli bir görevi gerçekleştirmek için büyük veri kümeleri üzerinde eğitilmiş, verilere dayalı tahminler yapan ve beyin benzeri bir nöral algoritmanın karşılığı olan bir derin öğrenme modelidir. Angie Lee, 'What is a Pre-Trained AI Model?' (NVIDIA, 8 Dec 2022) <https://blogs.nvidia.com/blog/2022/12/08/what-is-a-pretrained-ai-model/> accessed 22 Jul 2023.

4 OpenAI, 'Introducing ChatGPT' (OpenAI 30 Nov 2022) <https://openai.com/blog/chatgpt> accessed 22 Jul 2023. OpenAI, şu anki versiyonundan önce GPT1, GPT2 , GPT3 modellerini geliştirmiştir.

5 Bilgisayarların insan dilindeki ifadelerin anlamlarının (semantiğin) farkına varıp varmadıklarının önemi yakın zamana dek tartışma konusu olmuşsa da; *Kaku* tarafından ifade edildiği üzere, makinelerin dilin yapısına (sentaks) iyice hakim olduğu çağa doğru ilerlerken, bir ifadenin anlamının bilgisayar tarafından anlaşılıp anlaşılmaması önemsizleşecektir: *"Sentaksa iyice hakim olan bir robot, nereden bakılırsa bakılsın söylenen şeyi anlar, başka bir deyişle sentaksa kusursuz şekilde hakim olmak, anlamaktır."* Michio Kaku, *Olanaksızlığın Fiziği* Engin Tarhan (çev) (ODTÜ Yayıncılık 2014) 137.

6 T.C. Cumhurbaşkanlığı Dijital Dönüşüm Ofisi, 'Chatbot Uygulamaları ve ChatGPT Örneği' (15 March 2023) 39.

olmayıp, *"halüsinasyon"* olarak tabir edilen, gerçeğe aykırı sonuçlar üretebilmektedir[7]. Bu nedenle özellikle sağlık gibi riskli alanlarda kullanımında insanlar tarafından sonuçlarının teyit edilmesi veya yüksek risk içerecek bazı alanlarda hiç kullanılmaması düşünülmektedir[8]. ChatGPT'nin Kasım 2022'de yayınlanan GPT 3.5 dil modelinin gelişkin bir ardılı olan GPT4 modeli ise, 14 Mart 2023 tarihinde kullanıma açılmıştır. OpenAI tarafından yayınlanan teknik raporda, bu çok modelli yapay zekâ sisteminin baro sınavı gibi pek çok testte insan performansına eşdeğerde, bazen de insanlardan daha iyi sonuçlar ortaya çıkardığı ifade edilmektedir[9].

ChatGPT gibi üretici yapay zekâ modelleri, çok geniş veri setleri ile eğitilmekte; aynı zamanda pek çok farklı alanda kullanılabilen temel modellerinin (*foundation model*) yüksek kapasiteli çeşidi olarak nitelendirilen ileri model *(frontier model)* olarak da adlandırılmaktadır[10]. Bu yeni nesil yapay zekâ modelleri sayısal bilimler yönünden olduğu kadar sosyal bilimler açısından doğurabileceği sonuçları ile de tartışılmaktadır. Hukukun inceleme alanı bakımından ise, bu teknolojinin doğurabileceği risklerin temelinde yer alan mahremiyet ve güvenliğe dair etkileri açısından ele alınması gerekir. Bu bakış açısı, bu teknolojinin diğer alanlardaki etkilerini de anlamak ve yapılacak yasal düzenlemelerde dikkate alınması gereken hususlara işaret edebilmek için faydalı olabilir.

Çalışmamızın en temel argümanı; üretici yapay zekâ modelleri pek çok farklı alanda herkes tarafından kullanılabilecek son derece güçlü sistemler olduğundan, mahremiyet ve güvenlik risklerinin kendinden önceki yapay zeka modellerinden çok daha fazla olduğu, bu nedenle bu risklerin en aza indirgenmesini sağlamanın sadece teknolojik önlemler ile değil; etkili yasal düzenlemelerle mümkün olabileceğidir. Üretici yapay zekâ gibi büyük veri setleri kullanılarak eğitilen ve yaygın bir kullanım alanına sahip olan yapay zekâ modellerinin mahremiyet ve güvenlik açısından bir risk oluşturduğu yadsınamaz bir gerçektir. Bunun yanı sıra, bu risklerin; üretici yapay zekânın verinin beyin-makine arayüzleri gibi diğer teknolojilerin geliştirilmesinde kullanılması ile artış göstereceği öngörülmektedir[11]. Nitekim, Avrupa Birliği (AB) Yapay Zekâ Tüzüğü Taslağı'nda[12] 16 numaralı gerekçeye 14 Haziran 2023 tarihinde AB Parlamentosu'nun önerisi ile eklenen açıklamada belirtildiği üzere; fiziksel (beyin makine arayüzü teknolojisinin geliştirilmesi veya kullanımının yol açabileceği fiziksel zararlar gibi) veya psikolojik (insan bilincinin manipülasyonu ile ortaya çıkabilecek depresyon ve benzeri ruhsal sağlığı bozucu) zararların meydana gelmesinin olası olduğu

7 OpenAI, 'ChatGPT4 Technical Report' (Arxiv 27 Mar 2023) <https://doi.org/10.48550/arXiv.2303.08774> 10, accessed 22 Jul 2023.

8 ibid 10.

9 ibid 1, 6. Bahsi geçen baro sınavı *Uniform Bar Exam* baro sınavının simülasyonu olsa da; ChatGPT'nin test edilmesi için geliştirilen bu simülasyon sınav, gerçek baro sınavlarında geçmişte çıkan gerçek sorulardan yola çıkılarak hazırlanmıştır.

10 Markus Anderljung, Joslyn Barnhart, Anton Korinek and Jade Leung 'Frontier AI Regulation: Managing Emerging Risks to Public Safety' (11 Jul 2023) <https://arxiv.org/pdf/2307.03718.pdf> 6, accessed 22 Jul 2023.

11 Nöro teknoloji ve yapay zekâ giderek iç içe geçmiş durumdadır. Etik ve insan haklarına dair etkiler bu iki alanın hızla yakınlaşması ile daha da artmaktadır. UNESCO, 'Ethics of Neurotechnology' (Unesco) <https://www.unesco.org/en/ethics-neurotech> accessed 22 Jul 2023.

12 European Commission, Proposal for 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, 21.4.2021, COM(2021) 206 final, 2021/0106 COD <https://eur-lex.europa.eu/legal-content/EN/TXT/PDF/?uri=CELEX:52021PC0206&from=EN> accessed 22 Jul 2023.

durumlarda, insan davranışını etkilemek ve/veya bozmak amacıyla yapay zekâ ile desteklenen nöro teknolojileri bünyesinde bulunduran yapay zekâ modellerinin piyasaya çıkarılması, hizmete sunulması veya kullanılması yasaklanmalıdır[13]. Bu sınırlamanın kapsamına, gerçek bir kişinin davranışını manipüle edecek şekilde; beyin-bilgisayar arayüzleri aracılığıyla toplanan nöro verileri izlemek, kullanmak veya etkilemek için kullanılan yapay zekâ modelleri tarafından desteklenen nöro teknolojiler girmektedir.

Çalışmamızın ilk bölümünde; ChatGPT özelinde genel olarak üretici yapay zekâ modellerinin çalışma ilkeleri ve kullanım alanları açıklanmaktadır. İkinci bölümde, bu teknolojinin mahremiyet ve güvenlik açısından doğurduğu riskler, günümüzde ortaya çıkan örnekler üzerinden ele alınmaktadır. Son bölümde ise, konuya dair çözüm önerilerimiz, yürürlükteki ve tasarı aşamasındaki yasal düzenlemeler ışığında tartışılmakta ve üretici yapay zekâ modellerinin, mahremiyet ve güvenlik risklerine ilişkin yasal düzenlemeler için öneriler getirilmektedir.

I. ÜRETİCİ YAPAY ZEKÂ MODELLERİ

A. GENEL OLARAK

1959 yılında *Arf*; tasarlanması sırasında akla gelmemiş olan problemleri çözebilen bir makine yapılıp yapılmayacağı sorusuna cevap aramıştı[14]. *Arf*ın yarım yüzyıldan fazla bir zaman önce cevap aradığı bu soru; bilgisayarların, kendilerine insanlar tarafından sağlanan girdilerden yeni sonuçlar ortaya çıkarmalarının, bir başka ifadeyle, geniş kullanım alanlarına sahip olmalarının veya "üretici" olmalarının nasıl mümkün olabileceğinin yanıtlarını arıyordu. 2023 yılının sonlarına yaklaştığımız bu günlerde; teknolojik ilerleme ile ulaşılmaya çalışan hedefin genel amaçlarla kullanılabilen ve gerektiğinde "üretici" olabilen bir yapay zekâ yaratmak olduğunu ifade edilmelidir.

Dünyanın ilk bilgisayar programcısı[15] olarak kabul edilen *Lovelace*, 19. yüzyılda *Babbage* tarafından tasarlanan ve *Babbage* ile üzerinde çalıştıkları Analitik Makine *(Analytical Engine)*[16] ile ilgili bir makaleyi tercüme etmişti. *Lovelace*'ın makaleye eklediği çevirmen notlarının; Analitik Makine'nin potansiyelindeki öngörülerinin isabetli oluşu ve günümüzde genel amaçlı kullanılan bilgisayar programlarının ana fikrinin temelleri olduğu ifade edilmektedir[17]. Şöyle ki; *Lovelace*, Analitik Makine'nin her ne kadar sadece hesaplama yapmak için yaratılmış olsa da, insan tarafından belirlenen amaçlarla çok farklı görevler için kullanılabilecek bir genel makinenin ilk adımı olduğunu

13 Ibid, 38 Numaralı Değişiklik, 16 numaralı gerekçe.

14 Cahit Arf, 'Makine Düşünebilir mi ve Nasıl Düşünebilir?' Tarık Tuna Gözütok (ed), *Cahit Arf ve Atatürk Üniversitesindeki Halk Konferansları, 1958-1960* (1959) (Atatürk Üniversitesi Yayınları 1959), s. 95.

15 Bilgisayar programı görevi tanımlayan bir algoritmaya ve işlenen bilgilerin sayısal gösterimine dayanan veri setlerinden oluşur. Ethem Alpaydın, *Yapay Öğrenme: Yeni Yapay Zekâ,* Aylin Ağar (çev) (Tellekt 2020) 27.

16 Analitik Makine 19. Yüzyılda mucit *Charles Babbage* tarafından geliştirilen ve büyük sayılarla karmaşık hesaplamaları yapmayı otomatikleştiren bir makineydi. *Babbage* bu icadından önce ise Fark Makinesi isimli bir cihaz geliştirmişti. Ancak bu cihaz sadece toplama yapabiliyordu Walter Isaacson, *Geleceği Keşfedenler* Duygu Dalgakıran (çev) (Domingo 2014) 25.

17 ibid 25.

ifade eder[18]. Bunun yanı sıra, *Lovelace*, toplumda teknolojik gelişmeler ile ilgili ve bugüne dek pek az değişikliğe uğrayan bir bakış açısına da dikkat çeker. *Lovelace*'e göre; her tür yenilik ortaya çıktığı ilk anda toplumda "abartmak" ve "küçümsemek" olmak üzere iki farklı tür tepki doğurmaktadır[19].

Lovelace'in dikkat çektiği toplumun genelinde tezahür eden bu bakış açısı, üretici yapay zekâ modelleri ile ilgili hukuki değerlendirme yaparken bazı güçlüklere sebebiyet verebilir. Şöyle ki, başta temel haklar olmak üzere insan hayatı üzerinde bu modellerin olumsuz etki yaratabilecek bir riski olmayacağını öngörmek veya bu teknolojinin durdurulamayacak denli güçlenmiş olduğunu varsaymak bizi eylemsizliğe yöneltebilir. Bu iki uç öngörünün doğru bir eylemliliğe evrilmesi için, öncelikle yapılması gereken, üretici yapay zekâ modellerini anlamak için çaba sarf etmek ve bu teknolojinin doğurabileceği riskleri doğru bir biçimde ortaya koymaktır. Ancak bu şekilde olması gereken hukuk bakımından yapılması gereken yasal düzenlemeler doğru bir bakış açısı ile değerlendirilebilir[20].

2017 yılında gerçekleşen bir etkinlikte, Belçikalı Avukat *Dobbelaere-Welvaert* gelişen teknolojilerin etkisi ile 2025 yılında avukatların %80'inin mesleklerini yapmıyor olacağını dile getirdiğinde; bir görüş bu öngörünün gerçekçi olmadığı ifade ederken; bir diğer görüş bu olasılığın çok daha erken bir tarihte gerçekleşebileceği ifade etmişti[21]. Aslında *Dobbelaere-Welvaert*'ın bu öngörüsü, *Kurzweil*'in 2004 yılında ileri sürdüğü öngörü ile de oldukça uyumluydu. Zira, *Kurzweil,* 2029 yılının sonuna dek makine zekâsının Turing Testi'ni geçebileceğini ileri sürmüştü[22]. Yine, modern bilgisayar biliminin öncülerinden olarak addedilen *Turing*[23], makine zekâsının insan zekâsından ayırt edilmesinin 2000'li yılların başlarında güçleşmeye başlayacağını savunmuştur[24]. *Turing* 1950 yılındaki çalışmasında[25]

18 ibid 25.

19 Ada Lovelace, 'Notes of the Translator' for 'Sketch of the Analytical Engine Invented by Charles Babbage' (Taylor and Francis 1842) 722, <https://repository.ou.edu/uuid/6235e086-c11a-56f6-b50d-1b1f5aaa3f5e#page/1/mode/2up> accessed 22 Jun 2023.

20 Bu değerlendirmenin yapılabilmesi için farkında olunmasının faydalı olduğunu düşündüğümüz bir başka husus ise; *Lovelace*'ın yukarıda andığımız notlarında belirtilmiştir. Belli bir amaç için kullanılmak üzere geliştirilen bir teknik, zaman içinde, insan gücü ve düşüncesine katkı veren ve önceden planlanmamış farklı sonuçlar da doğurabilir. Gerçekten, bu çalışmamızın konusunu oluşturan ChatGPT gibi üretici yapay zekâ modelleri bugün farklı konularda sorduğumuz sorulara çeşitli (doğru-yanlış) yanıtlar vermekte iken; yakın gelecekte belki de "bilgili" bir sohbet botunun doğuracağından çok daha farklı etki ve sonuçlar ortaya çıkararak, insanlık ve dünya tarihini dönüştürecektir. Bu dönüştürücü etki bize göre sadece günümüz teknolojisinin gelişmesi veya farklı tekniklerin ortaya çıkması ile değil, aynı zamanda ve hatta bundan daha fazla; hukuki, sosyal, politik ve ekonomik etkiler ile filizlenecektir.

21 Helsinki'de düzenlenen ve tüm dünyadan yaklaşık 600 hukukçunun katılımı ile gerçekleşen dünyanın ilk uluslararası hukuk tasarımı zirvesi niteliğindeki Legal Design Helsinki Summit etkinliğinde; hukukçuluğun geleceği ve henüz yürürlüğe girmemiş olan AB Genel Veri Koruma Tüzüğü (*General Data Protection Regulation – GDPR*)'ndeki müstakbel yasal yükümlülükler tartışılmıştı. Başak Ozan Özparlak, 'Yeni Çağın Hukukunu Teknoloji ve Tasarım Şekillendirecek' HBT Dergi Herkese Bilim Teknoloji, 12.1.2018.

22 *Kurzweil*, 2005 yılında bilgi tabanlı teknolojilerin birkaç on yıl içinde (2015 veya 2025 civarında) tüm insani bilgi ve becerileri kapsayabileceğini ve bunun sonucunda bu teknolojilerin insan beynine özgü örüntü tanıma, sorun çözme becerileri ile duygusal zekâyı da kapsar hale geleceği fikrini savunmuştu. Ray Kurzweil, *İnsanlık 2.0: Tekilliğe Doğru Biyolojisini Aşan İnsan*, Müge Şengel (çev) (Alfa Yayınları 2017) 388.

23 David Deutsch, *The Beginning of Infinity: Explanations That Transform the World* (Allen Lane 2011) 148.

24 Alan M. Turing, 'Computing Machinery and Intelligence' (1950) Mind, Vol. LIX, Sayı 236, 433–460, https://doi.org/10.1093/mind/LIX.236.433 accessed 22.7.2023, s. 442.

25 ibid 442.

kendi adı ile anılagelen, bir test yardımı ile makine zekâsının insan zekâsından ne kadar ayırt edilebilir olduğunu ölçer[26].

Tüm bu öngörüler için "erkenci" demek mümkün müdür? ChatGPT gibi üretici yapay zekâ modellerinin geliştirildiği 2023 yılında bu soruya olumsuz yanıt vermek giderek güçleşmektedir. Bununla birlikte, zekâ ile anlaşılması gerekenin ne olduğu sorusu da önem arz etmektedir. Nitekim, zekâ kavramı ile ifade olunanın ne olduğunun henüz tam ifade edilmeden bu konuda bir tespitte bulunmak güçtür.

Alpaydın, öğrenen algoritmaların bilgisayarların hızlanması ve öğrenme verisinin artmasıyla birlikte, gittikçe daha "zeki" olacağını ve içinde bulunduğumuz yüzyıl sona ermeden, bu tür öğrenen bir zekânın insan zekâsına erişebileceğini ifade eder[27]. Günümüzde bu örneklerin ChatGPT veya genel olarak üretici yapay zekâ modelleri tarafından yerine getirilmesi çok uzak görünmemektedir. Ancak, Turing Testi'nin bir makina tarafından geçilmiş olması tek başına "akıllı" bir makine ile karşı karşıya olduğumuzu söylemeye ve makinanın bu kapsamda kabul edilse dahi insana atfedilen hak ve sorumlulukları bu sistemlere atfetmeye yeterli olacak mıdır? Zekâ ve zihnin derinliklerini henüz kesin olarak ortaya koyamadığımızdan, bu soruya olumlu yanıt vermek için henüz erken olabilir. Nitekim, yapay zekâ konusu ele alınırken asıl önemli olan husus, bu modellerin akıllı olup olmaması değil; bir sorunu etkili bir şekilde çözüp çözmemesidir. Oysa, bir sorunu etkili bir şekilde çözebilme yeteneği mutlaka zekânın varlığına işaret etmeyebilir[28].

Kurzweil'e göre; Turing Testi'ni geçebilecek bir sistemin, öncelikle, doğal dilin insan düzeyinde anlama yeteneği gerektiren işlerden en zor üçü olan; film eleştirisi, basın toplantısı ve konuşma çevirisini gerçekleştirebilmesi gerekir[29]. Bu görevler, bugün esasen ChatGPT gibi geniş dil modelleri ile başarılabilecek işlerdir. *Kurzweil* makinelerin bu yetkinliklere kavuştuğu bir çağın sessizce, insanlar çok da farkında olmadan geleceğini öngörür[30]. Günümüzde, Birleşmiş Milletler toplantılarına sohbet botu entegre edilmiş robotlar dahil edilerek[31] toplum nezdinde görünür biçimiyle bu yeni çağ muştulanmaktaysa da; toplumların büyük çoğunluğu popüler haberler veya gösteriler dışında

26 Turing Testi olarak adlandırılan bu test, bir insan gözlemcinin sorduğu sorulara verilen yanıtlara baktığında, yanıtların bir programdan mı yoksa insandan mı geldiğini ayırt edip edemeyeceğini test etmeye dayalıdır. ibid 433.

27 Alpaydın (n 15) 13.

28 Luciano Floridi and Massimo Chiriatti, 'GPT-3: Its Nature, Scope, Limits, and Consequences' (2020) <https://dx.doi.org/10.2139/ssrn.3827044> accessed 22 Jul 2023.Yapay zekâ sistemleri için "zeki" veya "akıllı" çıkarımını yapmamız halinde bu kabulün sadece teknolojik veya felsefi değil, aynı zamanda hukuki sonuçları da ortaya çıkabilir. Bu sonuç ise; henüz şirketlerin, devletlerin ve bireylerin teknolojinin geliştirilmesi veya uygulanması ile ilgili sorumluluklarını ve verinin kontrolünün nasıl sağlanacağının henüz belirlenemediği günümüzde, yazılımlara "hukuki kişilik verilmeli mi?" gibi soruların yeniden ve belki de yine çok erken bir şekilde gündemimize gelmesine neden olacaktır. Bu durumun bir sonucu asıl sorumluların perdenin arkasında kalmasına neden olma potansiyelidir. Öte yandan üretici yapay zekâ modellerinin giderek gelişmesi karşısında kanımızca acilen gündeme gelmesi gereken bir hukuki nitelik tartışması bulunmaktadır. Bu da verinin hukuki niteliğidir. Zira bunu belirlemek yapay zekâ veya "akıllı" sistemlerin hukuki kişiliğini tartışmaktan çok daha önemli ve elzemdir Başak Ozan Özparlak, *Büyük Veri Çağında Çalışma İlişkilerinde Yapay Zekâ Sistemlerinin Kullanılması: Hukuki Bir Değerlendirme* (On İki Levha Yayınları 2021) 46 vd.

29 Kurzweil (n 22) 437.

30 ibid.

31 United Nations, 'Meet the Robots Who Are Making The World a Better Place' (UN News 06 Jul 2023) <https://news.un.org/en/story/2023/07/1138412> accessed 22 Jul 2023.

nasıl bir gerçeklik ile karşı karşıya olunduğunun henüz farkında değil gibi gözükmektedirler. Oysa, risklerini öngörmek kadar faydalarından yararlanabilmek; bu yeni teknolojinin kullanıcılar tarafından anlaşılabilir olması ile mümkündür.

B. GENİŞ DİL MODELİNE DAYANAN ÜRETİCİ YAPAY ZEKÂ

Kendininkine benzer, hatta onu aşacak bir zekânın peşinde olan insan; uzun zamandır zekânın anlamını aramaktadır. Bu nedenledir ki, zekânın işlemcisi olarak düşünülen insan beyni ile ilgili çalışmalar, bilgisayar bilimleri alanındaki çalışmalara paralel şekilde ve çoğu zaman iş birliği içerisinde gelişmektedir. İnsan olarak icatlar yapabilmemizin, farklı diller öğrenebilmemizin, farklı mekanlarda yaşamaya adaptasyon yeteneğimizin kaynağı nedir? *Homo Sapiens'i* diğer türlerden ayıran zihinsel bir özellik var mıdır? Bu sorular henüz kesin olarak cevaplandırılamamış olsa da insanlığın tıpkı uzay hakkında bilgi sahibi olma süreci gibi, kendi iç uzayımız olan zihnimizin derinliklerinde neler olduğuna dair merakımız ve bizi aşabilecek yapay zihin yaratma çabalarımız herhalde türümüz var oldukça devam edecektir.

Yapay zihin ve akıllı makineler yaratma fikri, salt bir merak veya hayalin ötesine geçerek bilimsel araştırma konusu olmaya başladığı ilk andan itibaren, insanın dil (*language*) ile öğrenmesi ve etkileşimi ile de yakından ilintili olmuştur. 1959 yılında; "yapay zekâ" terimini ortaya atan *McCarthy*, ideal bir bilgi işlem sisteminin; girdileri doğal dilde alabilecek ve ardından sorgulara zaten bildiği gerçeklerden çıkarımlar yaparak yanıt verecek bir sistem olacağını öne sürmüştü[32]. *Hunt* da 1975 yılında, *"Makineler doğal bir dildeki basit ifadeleri anlayabilseydi, bilgisayarları kullanmanın ne kadar kolay olacağını bir düşünün."*[33] sözleri ile makinelerin doğal dili anlaması halinde insanlar ile etkileşiminin kolaylaşacağını ifade etmişti. Böyle bir sistemde karmaşık sorulara verilen cevaplar, makinelerin yapamayacağı varsayılan "anlama" eyleminin gerçekleştiği yolunda atılmış bir adım olarak değerlendirilebilir[34].

Dil, pek çok bilimin araştırma alanına girebilir. Çalışmamız üretici yapay zekâ modelleri üzerine olduğundan, bu konu ile doğrudan ilintili olacak şekilde bilgisayar bilimleri açısından dilin ne şekilde ifade edildiğini vurgulamak önemlidir. Bilgisayar bilimi açısından dil; etkili ve kesin bir iletişim aracı olarak ifade edilir[35]. Doğal dil işleme (*natural language processing, NLP*), en yalın tanımı ile; insan tarafından kullanılan dillerin (Türkçe, İngilizce, Fince gibi) bilgisayarlar tarafından kullanılmasıdır[36].

32 John McCarthy, 'Programs With Common Sense' in D. Blake and A. Uttelly (eds) Proceedings of the Symposium on the Mechanization of through Processes (H.M. Stationery Office,1959) 2.

33 Earl B. Hunt, *Artificial Intelligence* (Academic Press 1975) 16.

34 Makinelerin insan dilinden girdileri yanıtlarken; bu cümlelerin anlamını bulabilmeleri halinde yorumlama yapmalarının da mümkün olabileceğini söylenebilir. Ancak içinde bulunduğumuz aşamada bu durumun gerçekleşebileceği noktasında kesin bir yargıda bulunmak için henüz erken olabilir. Bu konudaki tartışmalar için bkz. Kaku (n 5). Yorumlama eylemi ile kastedilenin ne olduğunun belirlenmesi de bu değerlendirmede önem taşır. Örneğin psikoloji alanında Freud, yorumlama eylemini bir şeyin içinde saklanmış olanı bulmak olarak tanımlamıştı. Bu tanımdan hareket ederek cümlelerin anlamına ilişkin yukarıdaki değerlendirmeyi yapmak mümkün olabilir. Sigmund Freud, *A Complete Introductory Lectures on Psychoanalysis*, James Strachey (çev) (W.W. Norton 1966) 87.

35 Martin Erwig, *Once Upon An Algorithm: How Stories Explain Computing* (The MIT Press 2017) 141.

36 Ian Goodfellow, Yoshua Bengio and Aaron Courville, *Deep Learning* (The MIT Press, 2016) 448.

Doğal dil işlemenin kapsamına; insan metinlerini ve konuşmasını analiz ederek, üreterek, değiştirerek veya bunlara yanıt vererek doğal dil işlevlerini otomatikleştiren bilgisayar programları ve araçları girmektedir. Doğal dil işleme, dili girdi olarak kullanan, dili çıktı olarak üreten veya her ikisini birden yapan yapay zekânın bir alt kümesidir. Sohbet robotları (*chatbots*); makine çevirisi sistemleri ve sanal asistanlar, doğal dil işlemenin en önemli uygulama alanını oluşturan dil modellerine örnek olarak gösterilebilir. Bu modeller genellikle makine öğrenmesinin ileri türleri sayesinde geliştirilmektedir. Makine öğrenmesinde, istatistik temelli çeşitli gözlemler sonucunda çıkarımlar (*inferences*) yapabilen, öğrenebilen algoritmalar söz konusudur[37]. Makine öğrenmesi gözetimli *(supervised)* veya gözetimsiz *(unsupervised)* şekilde olabilir: Gözetimli öğrenmede, veri girdisi sunularak belirtilen çıktıya dair sonuçlar alınır. Gözetimsiz öğrenmede ise, daha önceden belirlenmiş bir çıktı yerine sadece veri girişi vardır ve bu girdilerin belirli bir düzen oluşturup oluşturmadığı ortaya çıkar[38].

Yeterince veri ve yeterli bilgisayar gücü[39] ile giderek gelişen ve bilgisayarların basit kavramlardan karmaşık kavramlar inşa edebilmesine olanak veren[40] derin öğrenme (*deep learning*) metodu ise; en az insan katkısı ile algoritmaların daha hızlı bir şekilde kendi kendine öğrenmesine olanak sağlar[41]. Derin öğrenme, makine öğrenmesinin özel bir türüdür[42]. Pekiştirmeli öğrenme (*reinforcement learning*) yönteminde ise, diğer öğrenme türlerinden farklı olarak yapılacakları söyleyen bir öğreten yerine, seyrek aralıklarla geri bildirim veren bir eleştirmen bulunmaktadır. Derin öğrenme yöntemi ile birleştirilerek kullanıldığında, pekiştirmeli öğrenmenin en bilinen uygulaması *Go* adlı strateji oyunu oynamak için geliştirilen ve 2016 yılında dünyanın en iyi *Go* oyuncularından birini 4-1'lik bir skorla yenen *AlphaGo* programıdır[43]. *Google Deep Mind* ekibi; *AlphaGo* programını gözetimli öğrenme ve pekiştirmeli öğrenme metotlarını birlikte kullanarak geliştirmiştir[44]. Böylece *AlphaGo,* sadece programlama aşamasındaki veriler ile değil buna ek olarak, programlanma aşamasından sonra da öğrenmeye devam ederek yüzlerce yıllık *Go* oyununda daha önce düşünülmemiş yeni hamleler geliştirebilmiştir[45].

Yukarıda bahsi geçen *AlphaGo*'da olduğu gibi, ChatGPT gibi üretici yapay zekâ modelleri de kendi kendini denetleyen kapsamlı bir veri topluluğu üzerinde önceden eğitilerek, ardından insan

37 Ethem Alpaydın, *Machine Learning: The New AI* (The MIT Press, 2016) 27.

38 ibid 27.

39 Yapay zekâ eğitmek için kullanılan bilgisayar gücünün son 40 yılda artış hızının her dokuz ayda iki katına çıktığına dair bir değerlendirme için bkz. Oğuzhan Gençoğlu, 'For Decades of AI Compute' (Laconic 27 Feb 2023) <https://www.laconic.fi/ai-compute/> accessed 22 Jul 2023.

40 Goodfellow et all. (n 36) 5.

41 Alpaydın (n 15) 91.

42 Goodfellow et all. n 36) 95.

43 Alpaydın (n 15) 110-111. AlphaGo, önce uzman bir veri tabanı ile eğitilmiş, sonrasında ise kendi kopyasına karşı oynayarak pekiştirmeli öğrenme yolu ile Go şampiyonlarını yenecek seviyeye gelmiştir. Alpaydın (n 15) 111.

44 AlphaGo eğitilirken insan uzman oyunlarından elde edilen veriden beslenerek uygulanan gözetimli öğrenme metodu programın kendi kendine oynadığı oyunlardan edinilen veriden beslenen pekiştirmeli öğrenme metodu ile birleştirilmiştir. David Silver, Aja Huang and Chris J. Maddison, et al, 'Mastering the Game of Go with Deep Neural Networks and Tree Search' (2016) 529 Nature 484–489 <https://doi.org/10.1038/nature16961> accessed 20 July 2023.

45 Martin Ebers, 'Regulating A and Robotics: Ethical and Legal Challenges, Algorithms and Law' Martin Ebers ve Susana Navas Navarro (eds) *Algorithms and Law* (Cambridge University Press, 2020) 11.

geri bildirimiyle pekiştirmeli öğrenme (*reinforcement learning with human feedback*) yöntemi aracılığıyla insan tercihleriyle uyumlaştırılır[46]. Böylece, ChatGPT gibi üretici yapay zekâ modelleri; programlama aşamasından sonra erişime açıldığı andan itibaren dünyanın her yerinden, farklı türde insan kullanıcılardan gelen veri ve geri bildirimler ile gelişme ve öğrenme sürecini sürekli olarak devam ettirmektedir.

ChatGPT, geniş dil modeline dayalı olarak insan geri bildirimiyle pekiştirmeli öğrenme yöntemi ile geliştirilen[47] ve doğal dili işlemek için tasarlanmış, bir tür yapay zekâ sohbet botu olarak tanımlanabilir[48]. ChatGPT'nin de içinde yer aldığı geniş dil modelleri (*large language models*, LLMs)[49] üretici yapay zekâ (*generative artificial intelligence*)[50] grubunda sınıflandırılır[51]. Üretici yapay zekâ; normalde insanlar tarafından üretilen metin, görüntü, ses veya video gibi içeriklere benzeyen "sentetik içerikleri"[52] yapay bir biçimde üretme amacıyla tasarlanır[53]. Üretici yapay zekâ

46 Hugo Touvron/Louis Martin, 'Llama 2: Open Foundation and Fine-Tuned Chat Models', (Meta AI, 18 July 2023) <https://ai.meta.com/research/publications/llama-2-open-foundation-and-fine-tuned-chat-models/> accessed 22 July 2023.

47 OpenAI, (n 4).

48 Council Of European Union, 'ChatGPT in the Public Sector – Overhyped or overlooked?'(2023) <https://www.consilium.europa.eu/media/63818/art-paper-chatgpt-in-the-public-sector-overhyped-or-overlooked-24-april-2023_ext.pdf> accessed 22 July 2023; Norwegian Consumer Council, 'Ghost in the Machine: Addressing the Consumer Harms of Generative AI' (Norwegian Consumer Council, June 2023), www.forbrukerradet.no/ai, accessed 26 June 2023, 7; Kristin E. Bush, 'Generative Artificial Intelligence and Privacy: A Primer' (Congressional Research Service, 23 May 2023), USA Congressional Research Service Report, >https://crsreports.congress.gov/product/pdf/R/R47569> accessed 28 June 2023, 2. Esasında algoritmalar kullanılarak insan etkileşimini taklit eden bir model bilgisayar bilimci *Joseph Weizenbaum* tarafından 1960'lı yıllarda yaratılmıştı. Eliza, bilgisayarla doğal dilde konuşmayı mümkün kılan bir programdır. Bkz. Joseph Weizenbaum, 'ELIZA-A Computer Program For the Study of Natural Language Communication Between Man and Machine' (1966) 9(1) Communications of the ACM.

49 Geniş dil modellerinin, farklı alanlarda belli amaçlar için uyarlanarak kullanılabilmelerinden dolayı bu modeller için günümüzde, "temel modeli" (*foundation model*) kavramının da kullanıldığı görülmektedir. ChatGPT gibi "geniş dil modelleri", özellikle dil odaklı sistemlere atıfta bulunmaktadır. "Foundation model" kavramı ise ChatGPT gibi modelleri de içine alacak şekilde, gelecekte yeni sistem türlerini barındırmak için genişleyebilecek daha geniş işlev tabanlı modelleri genel olarak ifade eden bir şemsiye kavramdır. The Center for Research on Foundation Models at the Stanford Institute for Human-Centered Artificial Intelligence, 'On the Opportunities and Risks of Foundation Models' (Arxiv, 16 August 2021) <https://doi.org/10.48550/arXiv.2108.07258>, accessed 22 July 2023.

50 Üretici yapay zekâ sistemlerine bu ismin verilmesinin nedeni, bu sistemlerin görsel ve işitsel sanatlar ve edebiyat dahil olmak üzere, daha önce insan yaratıcılığını gerektiren pek çok farklı alanda çıktı (*output*) üretebilmesidir. Bu sistemler veri gizliliğinin yanı sıra özellikle sanat, mimarlık, yazılım gibi alanlardaki fikri mülkiyet haklarına dair yeni hukuki tartışmalar ve uyuşmazlıklar doğurma potansiyeline de sahiptir. Bu anlamda, sanatın ne olduğu ve insan yaratıcılığının anlamına dair yeni teorik tartışmalar da gündemdedir. Örneğin, üretici yapay zekâ sistemlerinde olduğu gibi, bir resmin, fotoğrafın, tasarımın ortaya çıkışında insan, sadece algoritmaya komut veriyor ve yaratıcı bir etkinlikte bulunmuyorsa, bu durumda tamamen algoritmaların oluşturduğu bir eser söz konusu olacaktır. Jane Ginsburg and Luke Ali Budiardjo, 'Authors and Machines' (2019) 34 (2) Berkeley Technology Law Journal <http://dx.doi.org/10.2139/ssrn.3233885> accessed 20.7.2023. Böyle bir durum, eser sahipliği hakkının halen insana sahip olup olamayacağı tartışmasını beraberinde getirmektedir. Doktrinde böyle bir durumda, ilgili eserin kamuya ait olduğu, algoritmanın bu eserin yaratıcısı olduğu veya böyle bir üretimde eser sahibi hakları olmadığı gibi görüşler öne sürülmüştür. Yapay zekâ sistemleri ve fikri haklar ile ilgili ayrıntılı bilgi için bkz. Işıl Selen Denemeç, *To Feed or Not to Feed?:An Analysis of the Copyright Issues Surrounding the Use of Machine Learning Algorithms* (Lykeion 2021).

51 Samuel R. Bowman, 'Eight Things to Know about Large Language Models' (Arxiv 2 April 2023) <https://arxiv.org/pdf/2304.00612.pdf>, accessed 22 July 2023.

52 Sentetik içerik, tamamen veya kısmi biçimde yapay zekâ veya benzer diğer teknolojiler kullanılarak yaratılan içerikler olarak tanımlanabilir. Bart van der Sloot and Yvette Wagensveld (n 52), 4.

53 Norwegian Consumer Council (n 48) s.7; Bir yapay zekâ sistemi üretici olmak için tasarlanmak zorunda değildir. Yapay

algoritmaları programlama aşamasında ve sonrasında bu amacı gerçekleştirmek için büyük veri setleri üzerinde eğitilirler[54]. Bu veri setleri, üretici yapay zekâ modelinin farklılığına göre değişiklik gösterebilse de bu algoritmaların eğitim süreçlerinin ilkeleri büyük oranda benzerdir.

ChatGPT gibi geniş dil modelleri, eğitildiği veri setleri sayesinde farklı kelimeler arasındaki ilişkileri çıkarabilmekte ve bu ilişkiye dayalı olarak benzer metinler üretebilmektedir[55]. Geniş dil modellerinin geliştirilmesinde kullanılan veri setleri genellikle; e-kitaplar, forum ve haber siteleri, sosyal medya içerikleri gibi internetten elde edilen içeriklerdir[56]. Yapay zekâ tarafından üretilen veri ise genellikle belli olasılıklara dayalı bir biçimde üretilmiş olup, kullanıcılardan da elde edilen farklı veri girdilerine göre farklılaşabilmektedir[57].

Aşağıda örnekleri açıklanacak olan farklı türde üretici yapay zekâ örneklerinin çoğu, internet bağlantısına sahip ve herhangi bir özel teknik bilgi gerektirmeyen biçimde, herkes tarafından kullanımı kolay bir web arayüzü ile kullanılabilir. Ayrıca bu sistemler başka bir takım dijital hizmetlere ve sosyal medya hesaplarına entegre edilebilir[58].

İfade etmek gerekir ki; geniş dil modeline dayalı olarak geliştirilen tek yapay zekâ sistemi, OpenAI'ın ChatGPT modeli değildir. 18 Temmuz 2023'te Meta[59] AI şirketi, açık kaynak yazılıma dayalı ve ücretsiz bir üretici yapay zekâ sınıfına giren *Llama 2* modelini kullanıma sunmuştur[60]. Bu modelin OpenAI şirketi tarafından geliştirilen ChatGPT'den farkı ise açık kaynak yazılıma dayalı olmasıdır[61]. *Google* da ilk olarak *Bard* isimli modelini, 2023 yılında ise *Palm 2* geniş dil modelini tanıtmıştır[62].

C. ÜRETİCİ YAPAY ZEKÂ MODELLERİNİN FARKLI KULLANIM ALANLARI

Üretici yapay zekâ, sadece metin ve yazı alanında değil, aynı zamanda ses ve görüntüye dayalı içerik üretmede de kullanılabilir. Derin öğrenme tekniklerinin geliştirilmesi ile özellikle 2012 yılı ve sonrasında farklı alanlarda içerik oluşturulmasında üretici yapay zekâ modelleri kullanılmaya başlanmıştır. Bu modeller; müzik, resim ve şiir gibi yaratıcılık gerektiren alanlarda kullanım imkanı

zekâ sistemleri; görüntü verileri gibi verileri sınıflandırmak veya otonom araçlardaki kararları almak gibi başka hedeflere de sahip olabilir. Laurie A. Harris, 'Generative Artificial Intelligence: Overview, Issues, and Questions for Congress' (Congressional Research Service 9 June 2023) <https://crsreports.congress.gov/product/pdf/IF/IF12426> accessed 28 June 2023.

54 ibid 1.

55 Council Of European Union (n 48) 8; Bush (n 48) s. 3.

56 Norwegian Consumer Council (n 48) 8.

57 ibid 8.

58 ibid 7, 13.

59 Facebook 2021 yılında ismini Meta olarak değiştirmiştir.

60 Meta AI, 'Meta and Microsoft Introduce the Next Generation of Llama' (18 July 2023) https://ai.meta.com/blog/llama-2/>, accessed 22 July 2023.

61 Öte yandan her iki şirketin sadece üretici yapay zekâ alanındaki çalışmaları değil, yatırımcıları da ortaktır. Zira Microsoft şirketi OpenAI'ın yatırımcısı olup, Meta AI şirketi ile de Llama 2 modelinde iş birliği ortağıdır. Bkz (n 2)

62 Palm 2, 100'den fazla dilde çıktı üretebilmekte olup geliştirildiği veri setine bilimsel makalelerin de dahil olduğu belirtilmektedir. Google, 'Introducing PaLM 2', (Google 10 Mayıs 2023) <https://blog.google/ technology/ai/google-palm-2-ai-large-language-model/> accessed 22 July 2023.

sağlar. İlk üretici yapay zekâ modellerinin kullanılması sanatçılara yaratıcılık imkânı sağlaması ve ilham vermesi için olmuştur[63].

Üretici yapay zekânın en bilinen kullanım örneklerinden biri müzik eserlerinin yaratım sürecindedir. Algoritmik müzik üretimi; 1950'li yıllardan bu yana üzerinde çalışılmakla birlikte[64], derin öğrenme modellerinin gelişimi ile bu alandaki çalışmalar ilerleme kaydetmiştir[65]. Tasarlanan model, üretilmesi istenen müzik türünün örneklerinden oluşan bir koleksiyon (veri seti) üzerinden eğitilir. Eğitim verisi olan müzik eserleri, istenen türdeki tüm olası örneklerden alınmakta ve modelin bu eğitim verisinden yeni bir müzik eseri çıkarabilmesi için yeterli veriye sahip olması hedeflenmektedir[66]. Bu kapsamda bir araştırma ekibi tarafından üretici yapay zekâ kullanılarak bir müzik eseri üretimi için tasarlanan modelde, "*Walk in the Park*" isimli müzik eseri üretilmiştir[67].

Görsel alanda üretici yapay zekânın en bilinen örneklerinden biri ise, ChatGPT'nin yaratıcısı OpenAI şirketinin Ocak 2021'de tanıttığı ve doğal dilde verilen betimlemelerden yola çıkarak gerçekçi görseller oluşturan DALL-E ve bir yıl sonra yayınlanan gelişkin versiyonu DALL-E 2'dir[68]. Özellikle görsel alanda üretici yapay zekanın kullanılması ile, *deep fake* içerik adı verilen sahte içerikler de üretilmeye başlamıştır. *Deep fake* teriminin genel kabul görmüş bir tanımı yoktur[69]. Bir tanıma göre *deep fake*, bir kişinin yapmadığı ya da söylemediği bir şeyi yapıyor ya da söylüyormuş gibi göründüğü, üretilen sahte içeriğin gerçek kişiler veya yapay zekâ modelleri açısından tespitinin zor olduğu, gelişmiş teknik araçlarla oluşturulmuş bir içeriktir. Bu içerik, oldukça gerçekçi görünen videolar, fotoğraflar olabileceği gibi ses de olabilir[70]. *Deep fake*, 2017 yılında ortaya çıkmış ve o zamandan beri internette kullanıcıların kendi *deep fake* görüntü ve videolarını yaratmayı sağlayan programlar ortaya çıkmıştır[71]. Üretici yapay zekâ modellerinin giderek gelişmesi ile bu tür sahte

63 Ayrıntılı bilgi için bkz. Jan Smits and Tijn Borghuis (2022) 'Generative AI and Intellectual Property Rights' in Bart Custers, Eduard Fosch-Villaronga (eds) *Law and Artificial Intelligence*, *Information Technology and Law Series,* (T.M.C. Asser Press 2022) 324.

64 Algoritmik müzik üretimine ilişkin erken dönem bir çalışma için bkz. Lejaren A. Hiller and Leonard Isaacson, *Experimental Music Composition with an Electronic Computer* (McGraw – Hill Book Company 1959).

65 Luca Angioloni, Tijn Borghuis and Lorenzo Brusci (2020) 'Conlon: A pseudo-song generator based on a new pianoroll, wasserstein autoencoders, and optimal interpolations' (Research Gate, October 2020) 876 >https://www.researchgate.net/publication/348909705_Conlon_A_pseudo-song_generator_based_on_a_new_pianoroll_wasserstein_autoencoders_and_optimal_interpolations> accessed 22 July 2023.

66 Bonnie Buchanan and Danika Wright, 'The Impact of Machine Learning on UK Financial Services' (2021) 37(3) Oxford Review of Economic Policy, 326.

67 Araştırmaya ilişkin makale için bkz. Borghuis, Brusci, Brugi (n 65) 876; Ayrıca "Walk in the Park" isimli müzik eserinin telifi hakkında hukuki tartışmalara ilişkin olarak bkz. Smiths ve Borghuis (n 63) 327.

68 DALL-E'nin ardılı olan DALL-E 2 için bkz: OpenAI, 'DALL-E 2 Extending Creativity' (OpenAI 14 July 2023) <https://openai.com/blog/dall-e-2-extending-creativity> accessed 22 July 2023.

69 Bkz. Bart van der Sloot and Yvette Wagensveld (n 52), 1.

70 Bir içeriğin sahte olup olmadığının tespiti için şu özelliklerin bulunup bulunmadığının araştırması önerilmiştir: veri taşıyıcısının türü, sahte içeriğin oluşturulmasında kullanılan teknolojinin ne kadar gelişmiş olduğu, manipülasyonun derecesi, manipülasyonun aktarılan bilgi için ne kadar önemli olduğu, sahte içeriğin gerçek bir kimse ile ilgili olup olmaması, kişilerin içeriği ne ölçüde doğru kabul ettiği. Van der Sloot and Wagensveld (n 52) 4.

71 Tyrone Kirchengast, 'Deepfakes and image manipulation: criminalisation and control' (2020) 29 (3) Information& Communications Technology Law, 310.

içerik üretmek sistemi kullanmak için özel bir beceri gerektirmediğinden bu sentetik içeriklerin üretimi herkes tarafından kolayca gerçekleştirilebilir bir hâl almaktadır.

Üretici yapay zekâ finansal hizmetlerde de kullanılabilir[72]. Finansal hizmetler sektörü yapay zekâya en çok yatırım yapılan sektörlerden biridir[73]. Robot danışmanlar (*robo-advisors)* olarak da bilinen çevrimiçi otonom platformlar; yatırım tavsiyesine ihtiyaç duyan kişilere algoritmalar aracılığı ile finansal danışmanlık hizmeti sağlamaktadır[74]. Yatırımcılar bu şekilde yarı kişiselleştirilmiş portföy yönetimi hizmetlerine erişim sağlayabilmektedir[75]. Bireysel yatırımcılar robot danışmanlık hizmetleri sayesinde daha ucuz, hızlı ve günün her saatinde erişilebilir biçimde hizmet verebilmektedir[76]. Bununla birlikte, robot danışmanlar; yatırımcıların geliri, finansal bilgileri, yatırım deneyimleri, riske karşı tutumu gibi birçok hassas veriye de sahip olur[77].

Üretici yapay zekânın bir başka kullanım alanı da sağlık sektörüdür. Sağlık sektörüne ilişkin tıbbi görüntüleme ve teşhisten, ilaç üretimine dek birçok farklı alanda üretici yapay zekâ kullanımı yaygınlaşmaktadır[78]. Örneğin 2017 yılında *Woebot* adındaki bir üretici yapay zekâ sohbet botu geliştirilmiştir. Sohbet botunun amacı, insanlar ile etkileşimde bulunarak onların ruh halini iyileştirmeyi amaçlayan öneriler yapmaktır[79]. Yine ilaç sektöründe yeni ilaç geliştirmek için de üretici yapay zekâ kullanımı mevcuttur. Bu alanda kullanım ise, yukarıda sayılan alanlara göre hem çok daha yeni hem de ortaya çıkarabileceği riskler açısından hukuken daha tartışmalıdır[80].

72 Buchanan and Wright (n 66) 538; Güncel bir örnek olarak JP Morgan, 11 Mayıs 2023 tarihinde menkul kıymet ve finansal varlıklar hakkında yatırım danışmanlığı yapabileceği ifade edilen, ayrıca reklam ve pazarlama alanında da kullanılabileceği belirtilen IndexGPT modelini tescil ettirmek için başvuruda bulunmuştur: Will Daniel, 'Meet IndexGPT, the AI stock picker JP Morgan is developing that may put your financial advisor out of business' (Fortune, 26 May 2023) <https://fortune.com/2023/05/26/jpmorgan-indexgpt-a-i-stock-picker/> accessed 22 July 2023.

73 2019 yılında, İngiltere Merkez Bankası ve Finansal Piyasa otoritesi (FCA); bankalar da dahil olmak üzere finansal piyasa katılımcıları ile yaptığı ankette, piyasa aktörlerinin %66' sının makine öğrenme modelini bazı alanlarda kullanmak için çalışmalar gerçekleştirdiği sonuçlarını varmıştır: Bank of England (2019), 'Machine Learning in UK Financial Services', Working Paper (Bank of England, 16 October 2019) <https://www.bankofengland.co.uk/report/2019/machine-learning-in-uk-financial-services> accessed 22 July 2023.

74 Philipp Maume, 'Regulating Robo – Advisory', 55(1) Texas International Law Journal, 51; Dominique Payette, 'Regulating Robo-Advisers in Canada', 33(3) Banking & Finance Law Review, 423; Merve Ayşegül Kulular İbrahim, *Robo Danışmanların Hukuken Değerlendirilmesi* (Adalet Yayınevi 2023) 23.

75 Buchanan and Wright (n 66) 552.

76 ibid 552; Wolf Georg Ringe ve Christopher Ruof, 'A Regulatory Sandbox for Robo Advice, European Banking Institute' (2018) 26, 2.

77 Ringe and Ruof (n 76) 2. Ayrıca bkz. Payette (n 74) 423. Robo-danışmanlık hizmetlerinin hedef kitlesinin bireysel yatırımcılar olduğu göz önüne alındığında; yapay zekâ tarafından kullanılan ve geniş bir sınıflandırmaya dayalı olabilecek yatırım tavsiyeleri, yatırımcıların bireysel tercih, durum ve özel ihtiyaçlarını dikkate almakta başarısız olabilir. Yine robo-danışmanların bireysel yatırımcılara belirli varlık sınıflarını benzer bir şekilde tavsiye ettiği durumlarda bu durum finansal piyasada sistemik riski artırabilir. Ringe ve Ruof (n 76) 2.

78 Sağlıkta uygulamalara örnek olarak bkz. Polat Göktaş, Gül Karakaya, Ali Fuat Kalyoncu and Ebru Damadoğlu, Artificial Intelligence Chatbots in Allergy and Immunology Practice: Where Have We Been and Where Are We Going?' (2023) The Journal of Allergy and Clinical Immunology: In Practice, <https://www.sciencedirect.com/science/article/pii/S221.321.9823006414>, accessed 21 Agu 2023.

79 Woebot Health, (Woebolt Health 2023) <https://woebothealth.com/> accessed 27 July 2023; Attia Qammar et al., 'Chatbots to ChatGPT in a Cybersecurity Space: Evolution, Vulnerabilities, Attacks, Challenges, and Future Recommendations' (2021) 14(8) Journal of Latex Class Files, 3.

80 Hayden Field, 'The First Fully A.I.-Generated Drug Enters Clinical Trials in Human Patients' (CNBC 29 June 2023)

II. ÜRETİCİ YAPAY ZEKÂ MODELLERİNDE MAHREMİYET VE GÜVENLİĞİN HUKUKİ BOYUTU

A. KONUNUN ÖNEMİ

Bilgisayar bilimi, aynı zamanda "sorun çözme" bilimidir[81]. Gerçekten de bilgisayar bilimi, insan hayatındaki pek çok alanda mevcut veya muhtemel sorunların daha hızlı ve etkili biçimde çözümü için olanaklar yaratmaktadır. Hukukun da bilgisayar bilimi gibi hayatın pek çok alanına ve toplumun tümüne etkileri bulunan ve özel bir uzmanlık gerektiren bir bilim dalı olması yönü ile bilgisayar bilimi ile benzeştiği söylenebilir. Benzerliklerinin yanı sıra, hukuk ve bilgisayar bilimi özellikle son yıllarda devamlı bir etkileşim halindedir. Hukuk ve bilgisayar bilimi arasındaki bu etkileşim, bilgisayar bilimlerindeki son gelişmelerden biri olan üretici yapay zekâ modelleri ile daha çok yoğunlaşmaktadır. Bu etkileşimin odak noktası ise, mahremiyet ve güvenliğe dair riskler ve olası çözüm yollarıdır.

Verinin korunması temelde mahremiyet ve güvenliğe hizmet etmektedir. Bugün bireylerin mahremiyeti, özellikle hayatların dijital alana büyük oranda kayması ile bu ortamda paylaştıkları ve kendileri dışındaki kişi veya kurumlarca işlenen verinin korunabilmesine bağlıdır. Nitekim bu amaçla, 2021 – 2025 dönemi T.C. Ulusal Yapay Zekâ Stratejisi'nde mahremiyetin korunması yapay zekâ için göz önüne alınması gereken temel ilkeler arasında sayılmaktadır[82]. Dijital ortam bakımından güvenlik kavramı ise, zamanla yalnızca siber güvenlik ile sınırlandırılamayacak kadar bireylerin her alandaki güvenliğini etkiler hale gelmiştir. Güvenlik ile korunmaya çalışılan esasen verinin güvenliğidir. Bu nedenle mahremiyet ve güvenlik, özellikle verinin ekonomik bir değer haline gelmesi ile birbirini tamamlayan iki unsur haline gelmiştir. Öyle ki, veri mahremiyetini korumak veri güvenliğine de hizmet etmektedir. Çalışmamızda bu nedenle bu iki kavram birbirini tamamlayıcı şekilde ele alınacak olup, ilerleyen bölümlerde üretici yapay zekâ özelinde yapılacak tartışmalarda her iki kavramın birbiri ile kendiliğinden bağlantısı verilen örneklerden de anlaşılacaktır.

B. ÜRETİCİ YAPAY ZEKÂ MODELLERİNDE MAHREMİYET

1. Eğitim Verisi Açısından Mahremiyet

Yapay zekâ teknolojilerinin geliştirilmesinin temel şartı hem nitelik hem de nicelik yönünden büyük miktarda veriye[83] sahip olmaktır. ChatGPT gibi geniş dil modellerine dayanan üretici yapay zekâ

<https://www.cnbc.com/2023/06/29/ai-generated-drug-begins-clinical-trials-in-human-patients.html> accessed 22 July 2023.

81 Martin Erwig (n 35) 18.

82 TC Cumhurbaşkanlığı Dijital Dönüşüm Ofisi 'Türkiye Cumhuriyeti Ulusal Yapay Zekâ Stratejisi' (Cumhurbaşkanlığı Dijital Dönüşüm Ofisi, Ağustos 2021) >https://cbddo.gov.tr/SharedFolderServer/Genel/File/TR-UlusalYZStratejisi2021-2025.pdf> accessed 22 Jul 2023.

83 Büyük veri tanımı pek çok kaynakta farklı şekilde tanımlanmış olsa da Avrupa Konseyi'nin Büyük Veri Dünyasında Veri İşlenmesine Dair Bireylerin Korunmasına ilişkin olarak 2017 yılında yayınladığı rehberdeki tanımı bu konuda iyi bir kaynaktır: "*Büyük veri; çok büyük hacimdeki farklı verinin yüksek hızla toplanması, işlenmesi ve bu veri ile yeni*

örnekleri, bu gerçeğin katlanarak artmasına ve somut olarak gözler önüne serilmesine yol açmıştır. GPT3, 2020 yılında ilk yayınlandığında, eğitim verisinin içinde kişisel nitelikte veri de içeren internet içeriklerinin kazınması (*scrapping)* yöntemi ile elde edilen kitaplar, blog yazıları gibi geniş kaynaklardan oluştuğu ortaya çıkmıştı[84]. İnternetin kazınarak veri toplanması eylemi, trolle balık avcılığına benzetilebilir; ağa yakalanan balıkların av yasağı kapsamında olup olmadığı nasıl ki bu tür avcılık yönteminde belirlenemezse, internetin kazınması sırasında elde edilen verilerin, KVKK veya GDPR gibi veri koruma düzenlemeleri ile korunup korunmadığına bakılmamaktadır. Bu tür veri setlerinin içerisindeki kişisel veriler ile ilgili olarak veri öznesinden ilgili yasaya göre usulüne uygun olarak veri işlenmesine rıza alınıp alınmadığı dahi belirsizdir. Belirtilmelidir ki, internet kullanıcılarının tüm dünyada ürettiği günlük veri hacmi düşünüldüğünde, kitaplar bu geniş dil modellerinin beslendiği veri setinin sadece küçük bir parçasıdır. İşte bu nedenle, e-postalar ve akıllı klavyelerdekiler dahil olmak üzere interneti kazıyarak veri toplamak, geniş dil modeli geliştirenler şirketler için oldukça caziptir[85]. Bu nedenle, yapay zekâ modellerinin ham maddesi olan verinin mahremiyeti ve güvenliği en öncelikli meselemizden olmalıdır[86].

Şu anda, çoğu üretici yapay zekâ modeli, modelin geliştirilmesi ve eğitilmesinde kullanılan kişisel veriler bakımından ilgili kişilerin rızasını almamakta ve onların sunduğu verilerinin hangi işlemlerde kullanılacağına ilişkin bildirim sağlamamaktadır[87]. Bunun yanında internette kamuya açık bir biçimde olan veriler bakımından onay ve bildirim yükümlülüklerini sağlamak daha güç olacaktır[88].

Oysa, bilgisayar biliminin aynı zamanda bir sorun çözme bilimi olduğuna işaret edilmişti[89]. Bilgisayarlar ile sorun çözebilmek için ise algoritmalara, algoritmanın bulunmadığı durumlarda ise eksik bilgiyi telafi edebilecek veriye ihtiyaç bulunmaktadır[90]. Verinin günümüzde bir ekonomik değer haline gelmesinin nedeni, çağımızın itici gücü olan yapay zekâ modellerinin ham maddesi olmasıdır.

İnternet üzerindeki kamuya açık verilerin üretici yapay zekâ modellerinin eğitim verisi olarak kullanılması bazı etik tartışmaları da beraberinde getirmektedir. Birçok üretici yapay zekâ modelinin kamuya açık verileri, eğitim verisi olarak kullanmasının ve geliştirilen modellerin veriyi üreten tüm paydaş ve veri üreticilerine eşit bir biçimde yarar sağlama imkânı sunmaması potansiyeli bulunmaktadır. Nitekim böyle bir durum, bireyi ve bireyin verisini temel alan mevcut veri koruma

ve tahmine dayalı bilginin elde edilmesine yönelik ilerleyen teknolojik beceridir." Avrupa Konseyi, 'Guidelines On The Protection of Individuals With Regard to the Processing of Personal Data In A World Of Big Data'<https://rm.coe.int/CoERMPublicCommonSearchServices/DisplayDCTMContent?documentId=090.000.16806ebe7a> accessed 22 July 2023.

84 Matt Burgess, 'ChatGPT Has a Big Privacy Problem', (Wired, 4 Nisan 2023) <https://www.wired.com/story/italy-ban-chatgpt-privacy-gdpr/> accessed 22 July 2023.

85 El-Mahdi El-Mahamdi, Sadegh Farhadkhani, Rachid Guerraoui et al. 'On the Impossible Safety of Large Language Models', (Arxiv, 9 May 2023), <https://arxiv.org/abs/2209.15259> accessed 22 July 2023, 3-4.

86 Ozan Özparlak (n 28) 17.

87 Bush (n 48) 6.

88 ibid 6.

89 Erwig (n 81).

90 Alpaydın, (n 15) 27.

düzenlemeleri ile çözülemeyecektir[91]. Zira, dijital eşitsizlik sadece veri koruma kuralları ile çözülmesi mümkün olmayabilecek, rekabet kuralları ile de ilgili bir politik ve ekonomik durumdur. Böyle bir durumda; yeni geliştirilen teknolojiyi kontrol edenler, yetki alanları yaratma, kullanım koşullarını belirleme ve yaratılan modele kimlerin erişebileceğine karar verme konusunda önemli bir potansiyele sahip olacaktır[92].

Bu kapsamda ilk akla gelen risklerden biri; büyük teknoloji şirketlerinin tüm dünyada sundukları dijital hizmetlerine üretici yapay zekâ modellerini entegre ederek bu pazardaki payını pekiştirmesidir[93]. Böyle bir durumda, büyük teknoloji şirketleri pazar konumlarını güçlendirerek büyüyecek, kullanım koşulları ve yaratılan model hakkında kullanıcılara sözleşme koşullarını dayatabileceklerdir.

Bir başka risk olarak ise, üretici yapay zekâ modellerinin arama motoru gibi dijital hizmetlere entegre edilmesiyle bilgiye erişimin kısıtlanması ifade edilebilir. Klasik bir çevrimiçi arama motorunda, kullanıcıya aralarından seçim yapabileceği birden fazla sonuç gösterilmektedir. Üretici yapay zekâ modeli herhangi bir kullanıcı sorusuna tek cevap veren bir çıktı şeklinde tasarlanacak olursa, bu durumda kullanıcıların halihazırda mevcut olan bilgiye erişim imkânı da sınırlandırılmış olur. Yine böyle bir tasarımın çevrimiçi alım satım platformlarında kullanılmasında, çevrimiçi platform kullanıcıların almasını tercih ettiği ürünleri öncelikli olarak tercih etmesi için yeni yollar yaratacaktır. Böyle bir durum ancak bir sohbet robotu uygulamasının kullanıcıya vereceği bir öneriye ne şekilde ulaştığının kontrol edilmesiyle azaltılabilir[94]. Bu durum da ancak şeffaflık düzenlemeleri ile giderilebilir.

Eğitim verisi, günümüzde genel olarak tüm makine öğrenmesi modellerinin üzerinde yükseldiği temeldir[95]. Yapay zekâ modellerinin gelişimini sağlayan "öğrenme" süreci, veri ile gerçekleşmektedir. Özellikle üretici yapay zekâ modellerinde olduğu gibi, insan kullanıcıların doğal dillerinde yaptıkları yönlendirmeler ile çıktı üreten sistemlerin, doğal dili tanıması için gereken öğrenme verisinin miktar ve çeşit yönünden hacmi, sistemin başarısı ile doğru orantılı olacaktır. Bu nedenle, geniş veri setleri öğrenen algoritmaları geliştirenler için altın madeni değerindedir. Bu veri setlerine ulaşmak, her ne kadar veri gizliliğine aykırı yöntemler kullanılarak oldukça kolay olabilse de sosyal medyanın var olmadığı ve hayatın internette bu denli yansımadığı önceki on yıllarda bu arayış oldukça zordu. Bu zorluğun aşılmasının dönüm noktalarından biri bir hukuki süreçle ortaya çıkmıştır.

91 Saffron Huang and Divya Siddarth, 'Generative AI and the Digital Commons' (Arxiv 20 March 2023), >https://doi.org/10.48550/arXiv.2303.11074> accessed 22 July 2023.

92 Norwegian Consumer Council (n 48) 17; Bu etik tartışmalara somut bir örnek; Yeni Zelanda Māori yeri halkının Open AI tarafından geliştirilen bir üretici yapay zekâ modeli Whisper'a karşı çekincelerini dile getirmesi verilebilir. Whisper, Māori dili de dahil olmak üzere internette açık erişimde olan ses dosyalarını eğitim verisi olarak kullandı. Ancak yerli halktan bazı kişiler, kendi etnik ve kültürel kimliklerinin bir parçası olan dillerinin, kendi rızaları olmaksızın bir yapay zekâ modeli geliştirilmesinde kullanılması ile kültürlerinin üzerindeki hakimiyeti kaybederek, veri istismarına uğrayacağı endişelerini dile getirdiler. Rina Chandran, 'Indigenous Groups İn NZ, US Fear Colonisation as AI Learns Their Languages' (03 Ap 2023) <https://www.context.news/ai/nz-us-indigenous-fear-colonisation-as-bots-learn-their-languages> accessed 14 Jul 2023.

93 Norwegian Consumer Council (n 48) 18.

94 Norwegian Consumer Council (n 48) 18.

95 Kate Crawford, *Atlas of AI*, (Yale University Press 2021) 98.

Enron Skandalı olarak bilinen ve ABD'de Enron Şirketi' nin iflas etmesi ile sonuçlanan yolsuzluk soruşturmasında elde edilen 158 çalışana ait yarım milyonu aşan e-postadan oluşan veri tabanı; Federal Enerji Düzenleme Kurulu'nca herkese açık olarak yayınlanmıştır[96]. Böylece, Enron Külliyatı, makine öğrenmesi araştırmacıları için bir dilbilimsel hazine olmuş, fakat aynı zamanda, öğrenen algoritmaların eğitim verisinin elde edilmesinde bir tür yeni norm yaratmıştır: İnternette açık olarak yer alan bir veri setinin içerdiği kültürel, siyasal ve sosyal önyargılar ve veri gizliliği göz önüne alınmaksızın algoritmaları eğitmekte kullanılabilir[97].

Üretici yapay zekâ modellerinde kullanılan kişisel veri niteliği taşıyabilen eğitim verilerinin genellikle veri sahibinin bilgisi veya rızası olmaksızın kullanılabileceğine işaret edilmişti. Bunun en temel nedeni, toplanan verinin çok sayıda kaynaktan gelmesi ve büyük miktarlarda olmasıdır[98]. Şöyle ki, bu tür yapay zekâ modellerini geliştirmek için kullanılan eğitim veri setine, internette alenileştirilen kişisel veriler de dahil edilmektedir. Örneğin, sosyal medya fotoğrafları, içerikleri kişiler tarafından aleni bir biçimde paylaşılsa da kişiler bu kişisel verilerinin bir yapay zekâ modelinin eğitilmesinde kullanılabileceği konusunda bir rıza göstermiş değildir. Yine aynı şekilde kişiler bu şekilde bir ihlal gerçekleştirildiğinin farkında olmayabilir[99].

Bireylerin internet üzerindeki kişisel veri niteliğindeki fotoğraflarının, rızaları ve hatta bilgileri olmadan, yapay zekâ algoritması geliştirilmek için kullanımı yeni değildir. *Microsoft* şirketi tarafından 2016 yılında; ünlü aktörler, politikacılar, gazeteciler, aktivistler, akademisyenler ve sanatçıları da içeren 100.000 kişi ile ilgili yaklaşık 10 milyon fotoğraftan oluşan dünyanın en büyük halka açık yüz tanıma veri seti *MS-Celeb*, bu kişilerin bilgisi ve yüz tanıma sistemlerinde fotoğraflarının kullanımına dair rızaları olmaksızın, internette kazınma *(scrapping)* yolu ile oluşturulmuştu[100]. Bu veri setinde fotoğrafları kullanılan kişilerin rızalarının alınmadığına dair *Financial Times*'da yayınlanan bir soruşturmadan sonra, söz konusu veri seti *Microsoft* tarafından internetten kaldırmıştı[101].

ChatGPT'yi geliştiren Open AI şirketi, sadece ChatGPT ve DALL-E Labs ile ilgili API[102] hizmetleri için sınırlı olmak üzere, gizlilik politikasını 1 Mart 2023 tarihi itibari ile güncellediğini duyurmuştur. Buna göre; ChatGPT kullanıcılar tarafından açıkça belirtilmediği sürece, API verileri yapay zekâ modellerini eğitmek veya iyileştirmek için kullanılmayacaktır. Yine kullanıcı istemlerini ve

96 Jure Leskovec; 'Enron Email Network' (Stanford University) <https://snap.stanford.edu/data/email-Enron.html> accessed 22.7.2023; Bryan Klimt and Yiming Yang, 'Introducing the Enron Corpus' (2004), International Conference on Email and Anti-Spam (Natural Language Server 2021) >https://nl.ijs.si/janes/wp-content/uploads/2014/09/klimtyang04a.pdf> accessed 23.7.2023.

97 Crawford (n 95) 102.

98 Norwegian Consumer Council (n 48) 15.

99 Norwegian Consumer Council (n 48) 32.

100 Kate Crawford and Trevor Paglen, 'Excavating AI The Politics of Images in Machine Learning Training Sets' (2019), <https://excavating.ai/> accessed 22 July 2023.

101 Madhumita Murgia, 'Who's using your face? The ugly truth about facial recognition' (Financial Times, 19.4.2019) <https://www.ft.com/content/cf19b956-60a2-11e9-b285-3acd5d43599e> accessed 22 July 2023.

102 API (Application Program Interface) teriminin Türkçe karşılığı, "uygulama programı arayüzü" olup, yüksek düzeyli bir uygulama programının işletim sisteminin fonksiyon ve verilerini kullanmasını mümkün kılan ara yüz anlamını taşımaktadır. Bkz. Bülent Sankur, *İngilizce Türkçe Ansiklopedik Bilişim Sözlüğü* (Pusula Yayıncılık, 2008).

tamamlamalarını da içeren API verileri, kötüye kullanım ve kötüye kullanımın izlenmesi amacıyla en fazla 30 gün süreyle saklanacak ve yasalar aksini gerektirmedikçe saklama süresinin dolmasından sonra silinecektir[103]. Bu politika değişikliğinden yani 1 Mart 2023'ten önce API'ye gönderilen veriler ise, kullanıcı tarafından veri paylaşımı devre dışı bırakılmamışsa OpenAI tarafından belirtildiği üzere, şirketin yapay zekâ modellerini iyileştirmek için kullanılmış olabilir[104]. 25 Nisan 2023'te ise, kamuoyundan gelen tepkiler üzerine OpenAI, bireysel kullanıcıların da dilerlerse sohbet geçmişini kapatarak kullanım verisinin modelleri eğitmek için kullanılmamasını sağlayabilecekleri özelliği devreye soktuğunu duyurmuştur[105].

Ülkemizde kişisel verilerin korunması ile ilgili temel yasal düzenleme 6698 sayılı Kişisel Verileri Koruma Kanunu'dur (KVKK). Geniş dil modeline dayalı yapay zekâ modellerinin eğitilmesi için kullanılan veriler ile ilgili olarak Türk hukuku bakımından KVKK m. 4'te belirlenen ilkelere uygun bir veri işleme faaliyeti gerçekleştirilmesi gerekir. Bu kapsamda kişisel verilerin işlenmesinde; hukuka ve dürüstlük kurallarına uygun olma, doğru ve gerektiğinde güncel olma, belirli, açık ve meşru amaçlar için işlenme, işlendikleri amaçla bağlantılı, sınırlı ve ölçülü olma, ilgili mevzuatta öngörülen veya işlendikleri amaç için gerekli olan süre kadar muhafaza edilme ilkelerine uyulması gerekir. Aksi durumdaki veri işleme faaliyeti hukuka aykırı olacaktır.

Bununla birlikte, üretici yapay zekâ modellerinin özellikleri göz önüne alındığında, KVKK m. 4'te yer alan veri işleme faaliyetlerinde uyulması gereken ilkelerin, ChatGPT gibi geniş dil modellerine dayalı yapay zekâ modellerinin eğitilmesi sırasındaki veri işleme faaliyetinde uygulanabilmesi kullanılan veri setleri genişledikçe giderek güçleşmekte ve aykırılığın tespiti de aynı oranda zorlaşmaktadır.

KVKK kapsamında, ilgili kişiye, verilerinin münhasıran otomatik sistemler aracılığı ile analiz edilmesi suretiyle kendisi aleyhine bir sonucun ortaya çıkması halinde itiraz etme hakkı tanınmıştır. KVKK m. 11/f.1'in (g) bendinde; işlenen verilerin münhasıran otomatik sistemler vasıtasıyla analiz edilmesi suretiyle kişinin kendisi aleyhine bir sonucun ortaya çıkmasına itiraz etme hakkı olarak ifade edilmiş olan itiraz hakkı, yapay zekâ modelleri kullanılarak bazı alanlarda otomatikleşen karar süreçlerinde uygulama alanı bulabilir. Bu konuyla Kurum tarafından yayınlanan ilgili Rehberde ise, verilerin otomatik sistemler aracılığı ile işlenmesi; bilgisayar, telefon, saat gibi işlemci sahibi cihazlar tarafından yerine getirilen, yazılım veya donanım özellikleri aracılığıyla önceden hazırlanan algoritmalar kapsamında insan müdahalesi olmadan kendiliğinden gerçekleşen işleme faaliyeti olarak tanımlanmıştır[106].

Ancak ilgili kişinin bu şekilde bir itiraz hakkına sahip olduğunun bilmesi için kişisel verilerinin bu modellerde kullanıldığına ilişkin bilgi sahibi de olması gereklidir. Nitekim, Kurumun aydınlatma

103 OpenAI, 'API Data Usage Policies', 14 Haziran 2023, <https://openai.com/policies/api-data-usage-policies> accessed 5 Jul 2023.

104 ibid 1.

105 OpenAI, 'New Ways to Manage Your Data in ChatGPT', (Open AI 25 Apr 2023), <https://openai.com/blog/new-ways-to-manage-your-data-in-chatgpt> accessed 22 Jul 2023.

106 Kişisel Verileri Koruma Kurumu, 'Kişisel Verilerin Korunması Kanununa İlişkin Uygulama Rehberi' (2019) https://www.kvkk.gov.tr/SharedFolderServer/CMSFiles/41784a70-2bac-4e4a-830f-35c628468646.PDF 55 accessed 22 Agu 2023.

yükümlülüğüne ilişkin Tebliğ'de[107] aydınlatma yükümlülüğünün, kişisel verilerin, tamamen veya kısmen otomatik yollardan mı yoksa veri kayıt sisteminin parçası olmak kaydıyla otomatik olmayan yöntemler ile mi işlendiği bilgisinin açıkça belirtilmesini kapsadığının altı çizilmiştir (Tebliğ m.5/i).

Her ne kadar KVKK m. 11 uyarınca, üretici yapay zekâ modellerinin eğitimi için kendi verisinin kullanılıp kullanılmadığı ile ilişkili olarak kişiler veri sorumlusuna başvurabilecek olsalar bile her ilgili kişinin böyle bir hak temelli farkındalığa sahip olması ne yazık ki mümkün değildir.

Bizim de katıldığımız bir görüşe göre, geniş yapay zekâ modellerinin geliştirilmesi faaliyetinde, bu faaliyetlerin kendisinin hukuka aykırı olduğunun bir ön kabul ile benimsenmesi daha faydalı olacaktır[108].

2. Kullanıcı Verisi Açısından Mahremiyet

Günümüzde pek çok farklı alandan toplanan veri ile bu verinin otomatik olarak analizi, bireylerin çok farklı amaçlarla profillenmesine yol açmaktadır. Profilleme; bir tanıma göre, belirli bir kişi sınıfının bir dizi özelliğinin geçmiş deneyimlerden çıkarıldığı ve veriyi elinde bulunduran tarafından bu özellikler kümesine yakın olan bireylerin arandığı bir teknik olarak tanımlanmaktadır[109]. Avrupa Konseyi profillemeyi; bireyin kendisi ile ilgili karar almak veya kişisel tercihlerini, davranış ve tutumlarını analiz veya tahmin için bireyin veya kişisel bilgilerini analiz etmek veya tahmin etmek için kullanıldığını ifade etmiştir[110]. GDPR m.4'te ise profilleme, gerçek bir kişiyle ilgili belirli kişisel özelliklerin değerlendirilmesi, özellikle de iş performansı, ekonomik durumu, sağlığı, kişisel tercihleri, ilgi alanları, güvenilirliği, davranışları, konumu veya hareketleriyle ilgili hususların analiz edilmesi veya tahmin edilmesi amacıyla kişisel verilerin kullanılmasını içeren her türlü otomatik veri işleme süreci olarak tanımlanmıştır[111]. Profillemenin otomatik veri işleme süreci, otomatik veri işleme sürecinin kişisel veriler üzerinden yürütülmesi, bir gerçek kişinin kişisel özelliklerinin değerlendirilme amacı olmak üzere üç unsuru olduğu ifade edilmektedir[112].

Profilleme, yapay zekâ modellerinin geliştirilmesinde günümüzde kaydedilen ilerlemeden çok daha önce var olan bir uygulamadır. Bu kavram, en yalın haliyle, bireylerin farklı amaçlar ile farklı kategorilere ayrılması olarak tanımlanabilir. Gerek devletler gerek şirketler tarafından farklı

107 Kişisel Verileri Koruma Kurumu, Aydınlatma Yükümlülüğünün Yerine Getirilmesinde Uyulacak Usul ve Esaslar Hakkında Tebliğ, R.G. 10.03.2018, Sayı: 30356.

108 El-Mhamdi, Farhadkhani, Guerraoui, et al. (n 85) 19.

109 Roger Clarke, 'Profiling: A Hidden Challenge to the Regulation of Data Surveillance' (1993) Journal of Law and Information Science 4 (2) 405.

110 Council of Europe, The Protection Of Individuals with Regard to Automatic Processing of Personal Data In The Context Of Profiling Recommendation CM/Rec(2010)13 and Explanatory Memorandum (2011) 9 <https://rm.coe.int/16807096c3#:~:text=Recommendation%20CM%2FRec(2010)13%2C%20adopted%20by%20the,Data%20(T%2DPD)>, accessed 21 Ag 2023.

111 GDPR, m.4.

112 Madde 29 Veri Koruma Çalışma Grubu tarafından hazırlanan rehberde profillemenin üç unsuru olduğundan söz edilmektedir. Article 29 Data Protection Working Party, 'Guidelines on Automated individual decision-making and Profiling for the purposes of Regulation 2016/679 (wp251rev.01)' 22.08.2018, <https://ec.europa.eu/newsroom/article29/items/612053/en> s. 6-7, accessed 30 Jul 2023.

amaçlarla bireylerin kategorilere ayrılması yapay zekâ algoritmalarının kullanımından önce de var olan bir uygulamadır[113]. Bu uygulamalar, henüz yapay zekâ modellerinin geliştirilmeden evvel de pazarlama yönetimi alanında önerilmekteydi. Nitekim, *Kotler,* işletmelerin bilgisayar donanımlarına, veri tabanı yazılımlarına, analitik programlara, iletişim ağlarına ve bunlar ile ilgili yetenekleri olan personele yatırım yapmalarını önermişti[114].

Profillemenin yarattığı ucuz depolama maliyetleri ve büyük ölçekli bilgi miktarları veri sorumlularını gelecekte fayda sağlaması ihtimali ile daha fazla kişisel veri toplamaya yöneltmektedir. Bu durum hem KVKK m. 4'te sayılan veri işlemede uyulması gereken ilkelere hem de GDPR bakımından veri minimizasyonu ilkesine aykırılık teşkil edebilir[115]. Bir başka ifadeyle, veri sorumlusunun amaca ulaşmak için gerekli olan en az miktarda ve bu amaca ulaşmak için toplanan kişisel verilerin mutlaka gerekli olup olmamasının saptaması gerekir[116]. Bu yaklaşımın bir sonucu olarak, kişisel verilerin mümkün olduğundan az toplanması ve işlenmesi gereklidir[117].

Bununla birlikte profilleme birçok bakımdan sorun yaratabilir. Örneğin, ticari amaçla yapılan ve bireylerin satın alma alışkanlıklarına göre kategorize edilmesini hedefleyen bir profilleme, söz konusu bireylerin seçmen profillerine de ışık tutabileceği için aynı zamanda politik sonuç ve etkilere de sahip olabilir. Zira tüketici verisi olarak adlandırılan satın alma alışkanlıklarına dair veriler; bireylerin sosyo-ekonomik durumları, aile yapıları, dini inanışları hakkında da bilgi içerebilir. Benzer şekilde bu veriler, kişilerin sağlık durumları hakkında da pek çok çıkarımın yapılmasını sağlayabilmektedir. Sağlık ile ilgili bilgiler ise, sigortacılık sektöründen insan kaynakları sektörüne dek pek çok alanda ayrımcılık başta olmak üzere doğrudan veya dolaylı etkilere sahiptir.

Veri korumaya ilişkin yasal düzenlemeler üretici yapay zekâ modellerinde kullanılan verinin döngüsünde üç yönden uygulama alanı bulabilir. Bunlar; üretici yapay zekâ modelinin eğitilmesinde kullanılan eğitim verileri, üretici yapay zekâ modelinden elde edilen çıktı içerikler ve modelin kendisidir[118]. Bunun yanında, GDPR çerçevesinde altı çizilen ve mahremiyete dair koruma mekanizmalarının ürün ve hizmetlere henüz tasarım aşamasında entegre edilmesini ifade eden tasarımda gizlilik *(privacy by design)* ve varsayılan gizlilik *(privacy by default)* ilkelerinin üretici yapay zekâ modelleri açısından da gözetilmesi gerektiği açıktır[119].

113 1990'larda, işletmelerin hedef kitlelerine ulaşması ve pazar paylarını genişletmeleri için bireylerin tüketim alışkanlıklarına göre profillenmesi yapay zekâ algoritmaları ile olmasa da yapılmaktaydı. Mireille Hildebrandt, 'Profiling and the Rule of Law' (2008) 1 IDIS 57.

114 Öte yandan, pazarlama uzmanı olan *Kotler*'ın işletmeler bakımından bu değerlendirmesi, sadece pazar gücü anlamında değil, mahremiyetle ilgilidir. Zira, bu veri setlerinin oluşturularak bireylerin profillere ayrılması sırasında hem bireylerin mahremiyetinin risk altına girebilir hem de adil ve etik olmayan sonuçların ve hatta sahteciliğin dahi ortaya çıkabilir. Philip Kotler, *Marketing Management* (Pearson Education 2015) 724, 739.

115 Article 29 Data Protection Working Party (n 112) 11.

116 Elif Küzeci, *Kişisel Verilerin Korunması* (On İki Levha, 2020) 225, 237.

117 Norwegian Consumer Council (n 48) 46.

118 Norwegian Consumer Council (n 48) 45.

119 Küzeci, (n 116) 225.

GDPR'ın yapay zekâ modellerinde etkili bir veri koruma düzenlemesi olup olmadığı ise tartışmalıdır. GDPR'ın ChatGPT gibi üretici yapay zekâ modellerinde veri mahremiyeti bakımından tam anlamıyla bir koruma sağladığı söylenemez. GDPR, modellerin eğitiminde kullanılan veriler bakımından ya da çıkan veriler bakımından koruma sağlasa da modelin kendisi bakımından bu korumayı sağlayamamaktadır[120].Bir örnek olarak yürüyüş veya sosyal medya kullanımı gibi görünüşte özel nitelikli olmayan kişisel verilerin bir kişinin zindeliği veya tıbbi durumları hakkında bilgiler gibi sağlığına dayalı özel nitelikli veriye dönüştürme imkânı vardır. Modeller, kullanıcıların kendilerine ilgili bilgileri eklemesi veya bu bilgileri toplamasıyla gelecekte kişiler hakkında tahminlerini daha doğru belirlemeye başlar[121]. Esasında GDPR'ın yeni teknolojiler karşısındaki yeterliliğine dair tartışma, üretici yapay zekâ modellerinden önceki tarihte geleceğe dair bir değerlendirmesinde *Wachter* tarafından dile getirilmiştir. Bu görüşe göre; mevcut veri koruma kurallarının verinin sadece toplandığı ve işlendiği aşamaya göre belirlendiğinin, ancak, verinin toplanmasından sırasındaki kişisel/kişisel olmayan veya hassas/hassas olmayan veri ayrımlarının ve bireyler ile ilgili olarak yapılan çıkarımların yol açtığı mahremiyet ve bunun sonucu olarak ayrımcılık ihlallerini önleyemeyeceğini belirtilmiştir[122]. Yine, GDPR gibi günümüzün veri koruma yasalarında verinin bu şekilde kategorizasyonunun, tamamen verinin toplandığı aşamaya göre belirlendiği dikkat çekicidir[123]. Oysa, verinin toplanmasından sonra bu verilerin analizi ile bireyler profillerine göre çeşitli kategorilere ayrılmakta ve bu kategoriler ise daha sonra işe alım gibi süreçler ile ilgili alınan kararların verilmesinde kullanılmaktadır.

Bu tartışmalar bakımından; ChatGPT ile ilgili ilk yasal değerlendirmeyi yapan İtalyan Veri Koruma Otoritesi'nin (*Garante per la protezione dei dati personali*) kararı önem taşır. 31 Mart 2023 tarihinde ABD merkezli ChatGPT modelini geliştiren OpenAI şirketinin, İtalyan vatandaşlarının verilerinin işlenmesine dair bir sınırlama getirmiştir. Buna göre, Open AI şirketi tarafından uygulamanın kullanıcısı ve verinin öznesi olan kişilere bu verilerin depolanması ve ne şekilde kullanıldığı hakkında bilgi vermiyor oluşu bir eksiklik olarak ifade edilmektedir. Aynı zamanda Open AI şirketinin ChatGPT modelini geliştirmek için kullandığı bu verilerin toplanması ve işlenmesinde de yasal bir dayanak olmadığı ve sunulan hizmetin her ne kadar 13 yaş üstü kullanıcılara yönelik olduğu ifade edilse de herhangi bir yaş doğrulama mekanizması olmadığının da altı çizilmiştir. Yine Open AI şirketinin AB sınırları içerisinde ikametgahı olmaması ve AB için bir temsilci atamamış olması da değerlendirmeler arasındadır[124].

Üretici yapay zekâ modellerinin geliştirilmesi ve yaygınlaştırılması aşamalarında birden fazla paydaşın sürece dahil olması veri gizliliğinin hukuki çerçevesini belirlemeyi güçleştirmektedir. Özellikle veri gizliliği ile ilgili yükümlülük ve sorumlulukların belirlenebilmesi bakımından bu paydaşların

120 Michael Veale, Reuben Binns and Lilian Edwards, 'Algorithms that Remember: Model Inversion Attackand (2018) 376 Data Protection Law' Phil. Trans. R. Soc. A., 3.

121 ibid 3.

122 Sandra Wachter, *'Data Protection in the Age of Big Data*' (2019) 2 Nature Electronics Volume , 6-7.

123 ibid 7.

124 Garante Per La Protezione Dei Dati Personali (31 Mar 2023) <https://www.gpdp.it/web/guest/home/docweb/-/docweb-display/docweb/9870847#english> accessed 21 Ag 2023.

rollerinin belirlenmesine ihtiyaç bulunmaktadır. Bu nedenle Norveç Tüketici Konseyi ChatGPT ile ilgili yayınlanan raporunda; üretici yapay zekâ modelleri sürecine katılan farklı paydaşların, bu yapay zekâ modellerinin tüm döngüsünde GDPR hükümleri ile uyumunu sağlamak için rollerinin açıkça belirlenmesi gerekliliğini ifade etmektedir[125]. Bu öneri kanaatimizce son derece isabetlidir.

3. Mahremiyeti Artırıcı Teknolojilerin Üretici Yapay Zekâ Modelleri Açısından Uygulanmasındaki Zorluklar

Mahremiyeti artırıcı teknolojiler açısından GDPR'ın 25. maddesi ile yasal bir dayanak oluşturmaktadır[126]. GDPR m. 25 gereğince; veri sorumlusu, veri işleme sırasında ve sonrasında son teknoloji kullanır ve takma adla adlandırma gibi uygun olan tüm teknik ve yönetimsel önlemleri almakla yükümlü olup, bu yükümlülük işlenen verinin miktarı, kapsamı, saklama süresi ve erişilebilirlik unsurlarını da kapsar.

Üretici yapay zekâ modellerinde kullanılması yukarıda aktardığımız üzere GDPR 25. Madde uyarınca bir yasal yükümlülük olan mahremiyeti artırıcı teknolojilerin ülkemizdeki yasal karşılığı KVKK 12. Maddesidir. Mahremiyet ve güvenliğin ilişkisine yasal bir örnek olan, "Veri güvenliğine ilişkin yükümlülükler" başlıklı maddenin ilk fıkrasının (c) bendi uyarınca, veri sorumlusunun, verinin saklanmasında her türlü teknik ve idari önlemi alması gerekmektedir[127].

Mahremiyetin korunması ve kişisel verileri koruma kanunlarındaki veri koruma tedbirlerine uyum sağlanabilmesi için alınması gereken teknik tedbirler bulunmaktadır. Bu teknik önlemler arasında sayılabilecek ve ortak özellikleri veri minimizasyonu, merkezi veri depolamadan kaçınmak ve denetime açık olmak olarak gösterilebilecek[128] mahremiyeti artırıcı teknolojilerin (*privacy enhancing technologies – PETS)* en önemlilerinden bazıları diferansiyel mahremiyet[129] *(differential privacy)* ve federe makine öğrenmesi[130] *(federated machine learning*) yöntemleridir. Ancak üretici yapay

125 Norwegian Consumer Council (n 48) 45.

126 Leyla Keser Berber, 'Çapraz Etkileşim: Mahremiyete İlişkin Mevzuat ve Mahremiyet Standartları Arasındaki İlişki', in Leyla Keser Berber and Ali Cem Bilgili (eds) *Güncel Gelişmeler Işığında Kişisel Verilerin Korunması Hukuku* (Oniki Levha Yayınları 2020) 10.

127 GDPR 32. maddesinde de benzer bir hüküm yer almakta ve maddenin (a) fıkrasında *encryption* ifadesi yer almaktadır. Şifreleme (*encryption*), GDPR m. 34/3(a) ile kişisel veriye erişim yetkisi olmayan kişiler için anlaşılmaz kılan yöntemlerden biri olarak açıkça kabul edilmiştir.

128 Seda Gürses and Bart Preneel, 'Cryptology and Privacy in the Context of Big Data' in Bart van der Sloot, Dennis Broeders and Erik Schrijvers (eds) *Exploring the Boundaries of Big Data* (Amsterdam University Press, 2016) 67.

129 Diferansiyel mahremiyet, özellikle istatistik ve makine öğrenmesi analizlerinde kullanılan geniş veri setlerinde yer alan verinin ölçülebilir matematiksel yöntemler ile anonimleştirilmesini sağlar (Elif Üstündağ Soykan, Zeki Bilgin, Mehmet Akif Ersoy and Emrah Tomur, 'Differentially Private Deep Learning for Load Forecasting on Smart Grid' (2019) IEEE Globecom Workshops, 2, <10.1109/GCWkshps45667.2019.902.4520> accessed 22 Jul 2023. Diferansiyel mahremiyet ilk kez, *Dwork et. all* tarafından yayınlanan ve 2017 Gödel Ödülünü kazanan bir çalışmada 2006 yılında ortaya çıkarılmıştır. Bkz. Cynthia Dwork, Frank McSherry, Kobbi Nissim and Adam Smith, 'Calibrating Noise to Sensitivity in Private Data Analysis' in S Halevi and T. Rabin (eds) *Theory of Cryptography* (Springer, 2006) <https://people.csail.mit.edu/asmith/PS/sensitivity-tcc-final.pdf> accessed 22 Jul 2023.

130 Federe makine öğrenmesinde, kullanıcı verileri eğitim verisi olarak kullanılmak üzere bir merkeze gönderilmemekte, eğitim verisi kullanıcının kalmaktadır. Bu yöntem ilk kez 2016 yılında Google araştırmacıları tarafından duyurulmuştur. Jakub Konečný, H. Brendan McMahan, Daniel Ramage and Peter Richtárik, 'Federated Optimization: Distributed

zekâ gibi geniş modellerde bu tekniklerin uygulanması, yapay zekâ alanındaki şirketler arasındaki performansa odaklı yarış nedeniyle giderek zorlaşmaktadır. Şöyle ki; tüm teknik yöntemler, yapay zekâ modellerinin geliştirilmesi için kullanılan veride bir şekilde değişiklik yaratmakta olup bu nedenle de bu modellerin verdiği sonuçlardaki doğruluk (*accuracy*) payı üzerinde olumsuz bir etkiye neden olmaktadır[131]. Bu nedenle, mahremiyeti artırıcı teknolojilerin kullanılması, tek başına bir yasal önlem olarak yeterli olmayacaktır.

C. ÜRETİCİ YAPAY ZEKÂ MODELLERİNDE GÜVENLİK

Veri sızıntıları veri güvenliğine ilişkin bir sorundur. Türk hukuku bakımından, KVKK m. 12'de veri sorumlusunun veri güvenliğini sağlamasına ilişkin bazı yükümlülükler yüklenmiştir. İlgili hükme göre; veri sorumlusu kişisel verilerin hukuka aykırı olarak işlenmesini, erişilmesini önlemek ve muhafazasını sağlamak amacıyla uygun güvenlik düzeyini temin etmeye yönelik gerekli her türlü teknik ve idari tedbirleri almak zorundadır[132]. Nitekim, veri güvenliğinin sağlanması, kişisel verilere ilişkin güvenlik sağlanması açısından da gereklidir[133]. Bu kapsamda; veri sorumlusu tarafından risklerin tespit edilmesi gerekir. Veri sorumlusu, gerçekleşme olasılığının yüksek olduğu riskleri ve gerçekleşmesi durumunda yol açacağı kayıpları doğru bir şekilde belirlemek ve buna uygun tedbirleri almakla yükümlüdür. Veri sorumlusu bu riskleri belirlerken; kişisel verilerin özel nitelikli kişisel veri olup olmadığı, mahiyeti gereği hangi seviyede gizlilik gerektirdiği, güvenlik ihlali halinde ilgili kişi bakımından ortaya çıkarabilecek zararın niteliği ve niceliği dikkate alınmalıdır[134].

ChatGPT ve üretici yapay zekâ modellerinin sağladığı avantajlardan potansiyel yararın sağlanabilmesi için veri güvenliğine ilişkin olan riskler tanınmalı, yönetilmeli ve mümkünse azaltılmalıdır[135]. İfade edilmelidir ki, ChatGPT ve benzer üretici yapay zekâ modelleri veri sızıntısı bakımından risklidir. Nitekim, ChatGPT gibi sohbet botları genellikle çevrimiçi arayüzler üzerinden kullanıcılar ile etkileşim imkanı sağlar. Sohbet botları bakımından ise bazı güvenlik açıkları mevcuttur. Bir kullanıcının istenmeyen eylemleri veya davranışlarını başlatmak için sohbet botunun güvenlik açıklarından yararlanmaya çalışması ya da konuşma botlarına yanlış ve yanıltıcı bilgiler sağlayıp onları yanlış yönlendirmesi, yetkisiz erişim gibi örnekler bu kapsamda verilebilir[136]. Sohbet botları bakımından; veri bütünlüğünü ve gizliliğini sağlamak için uçtan uca şifreleme, kimlik doğrulama gibi önlemler getirilebileceği ifade edilmektedir[137]. Her ne kadar

Machine Learning for On-Device Intelligence' (2016) <https://doi.org/10.48550/arXiv.1610.02527> accessed 22 Jul 2023.

131 Charu C. Aggarwal, *Recommender Systems* (Springer, 2016) 434.

132 Kişisel verilerin işlenmesi sürecinde veri sorumlularının alması gereken teknik ve idari tedbirler konusunda Kişisel Verileri Koruma Kurumu tarafından rehber yayınlanan rehber için bkz. Kişisel Verileri Koruma Kurumu, 'Kişisel Veri Güvenliği Rehberi Teknik ve İdari Tedbirler' (2018) <https://www.kvkk.gov.tr/yayinlar/veri_guvenligi_rehberi.pdf> 28 Jun 2023.

133 Küzeci (n 116) 414.

134 Kişisel Verilerin Korunması Kurumu (n 132) 8.

135 Council Of European Union (n 48) 1.

136 Qammar et al (n 79) 6.

137 Qammar et al (n 79) 13.

bu riskler diğer tüm dijital teknoloji için söz konusu olsa da, özellikle çok farklı alanlarda ve geniş kesimlerce kullanılabilen üretici yapay zeka sistemlerinin yüksek nitelikli analiz kapasitesi ve gelişme hızının öngörülememesi gibi nedenlerle üretici yapay zeka sistemlerinde bu riskler kanımızca çok daha yüksektir.

Siber saldırılara karşı bu modellerin açık olması da veri güvenliğine ilişkin bir başka sorun olarak ortaya çıkabilir. Bu modellere karşı yapılan siber saldırılar hem öğrenme verisinin hem de kullanıcı kitlesinin geniş olması nedeni ile risk faktörünü artırmaktadır. Zira bu modeller, şimdiye dek kullanılan yapay zekâ modellerinden hem girdi verinin hacmi hem de çıktı bakımından oldukça farklıdır. ChatGPT gibi herkesin kullanımına sunulan ve dünyanın her yanından büyük bir kullanıcı kitlesine sahip olan üretici yapay zekâ modellerinde, teknolojinin olası tüm kötüye kullanımının öngörülmesi güç olabilir. Bu kadar risk barındıran bir modelin geniş bir kamuoyu tarafından kullanılması da birçok sorunu beraberinde getirmiştir[138]. Örneğin, makine öğrenmesi modellerine karşı yapılan saldırı türlerinden olan ve belirli bir verinin bir veri seti içinde yer aldığını öğrenme amacı ile gerçekleştirilen üyelik/aidiyet saldırısı (*membership attack*)[139] geniş yapay zekâ modeli ailesinden olan üretici yapay zekâ modellerindeki önemli güvenlik risklerinden birini oluşturmaktadır.

Bu konuda yapılan güncel bir araştırmada, ChatGPT'nin de içinde bulunduğu geniş dil modelleri gibi geniş yapay zekâ modellerinin güvenlik tehditlerine karşı aşırı derecede açık oldukları belirtilmektedir. Buna gerekçe olarak modellerin; kullanıcı tarafından üretilen veriye dayalı olmaları, yüksek ölçekli ezber kullanmaları, oldukça heterojen nitelikteki kullanıcılardan öğrenmeleri unsurlarının tümünün bir arada olması gösterilmektedir[140].

Üretici yapay zekâ modellerinin veri hırsızlığı için de kötü niyetli olarak kullanılabilmesi, bu sistemin yol açtığı veri gizliliği risklerinden biridir. Veri hırsızlığı; ticari sırlar, kişisel bilgiler, şifreler ve yazılım kodları gibi gizli verilere yetkisiz kişi veya kişilerce ağa sızılarak erişilmesidir[141]. Bir başka risk olarak; ChatGPT gibi üretici yapay zekâ modellerine kullanıcı tarafından sorulan soruların, model tarafından her zaman doğru bilgi ile yanıtlanmaması da bilginin doğruluğu bakımından yanıltıcı olmasının yanı sıra aynı zamanda veri gizliliği özelinde risk barındırmaktadır.

Tüm bu risklerin yanında, bu modeller, yasa dışı faaliyetlerde kullanılmak için kötüye kullanılabilir. Özellikle geniş dil modelleri, oldukça ikna edici bir biçimde metinler üretmek için dolandırıcılık gibi suçlarda kullanılabilir. Ya da, siber saldırılar bakımından kötü niyetli bir biçimde kullanılabilir. Bu durumun bir sonucu olarak bir kişinin teknik donanımı olmasa da onun siber suç faaliyetlerinin gerçekleştirmesini kolaylaştırabilir[142].

138 Norwegian Consumer Council (n 48) 12.

139 Üstündağ Soykan et all. (n 129) 5.

140 El-Mhamdi, Farhadkhani, Guerraoui et al. (n 85) 2.

141 T.C. Cumhurbaşkanlığı Dijital Dönüşüm Ofisi, (n 6) 55.

142 Norwegian Consumer Council (n 48) 33.

D. AVRUPA BİRLİĞİ YAPAY ZEKÂ TÜZÜK TASARISINA İLİŞKİN DEĞERLENDİRMELER

AB Komisyonu tarafından ilk kez 21 Nisan 2021'de teklif edilen AB Yapay Zekâ Tüzük Tasarısı[143], teknoloji tarafsız bir anlayışla, ileride ortaya çıkacak olan yeni modelleri de kapsayacak şekilde yapay zekâ sağlayıcılarının AB pazarında ürün pazarlamaları ve hizmet sağlamalarını belirli koşullara tabi kılmayı hedeflemektedir[144]. Henüz yürürlüğe girmemiş olan AB Yapay Zekâ Tüzük Tasarısı'nda belirtildiği üzere; ilgili Tüzük, risk temelli[145] bir yöntem ile hazırlanmıştır. Tüzük Tasarısı, kabul edilemez risk içeren sistemlerin kullanımının yasaklanması, yüksek riskli yapay zekâ sistemlerinin ise belirli uyum koşullarını sağlamaları kaydı ile AB pazarına girebilecekleri öngörülmüştür.

14 Haziran 2023 tarihinde, AB Parlamentosu tarafından ilgili Tüzük Tasarısı'na; ChatGPT gibi üretici yapay zekâ sistemlerinin dahil olduğu temel modelleri (*foundation models*) ile ilgili olarak sunulan değişiklik önerileri kabul edilmiştir[146]. Bu öneriler ile, ilgili Tüzük'e 28b numaralı madde eklenerek, temel modeli sağlayıcısının, modeli piyasada bulundurmadan veya hizmete sunmadan önce, bağımsız bir model olarak mı yoksa gömülü olarak mı sunulduğuna bakılmaksızın, bazı gerekliliklere uygun olduğundan emin olması gerektiği ifade edilmektedir.

Bu gereklilikler:

– Geliştirme öncesinde ve tüm döngü boyunca sağlık, güvenlik, temel haklar (mahremiyet de bu kapsamdadır), çevre, demokrasi ve hukukun üstünlüğüne yönelik makul olarak öngörülebilir risklerin tanımlanmasını, azaltılmasını ve hafifletilmesini uygun tasarım, test ve analiz yoluyla, bağımsız uzmanların katılımı ile belgelenmesi

– Modelin geliştirilmesinden sonra kalan ve azaltılamayan risklerin belgelenmesi (siber riskler, güvenlik de bu kapsamdadır).

– Yalnızca temel modeller için uygun veri yönetişim önlemlerine, özellikle de veri kaynaklarının uygunluğunu ve olası yanlılıklarını (*bias*) ve bu yanlılıkları azaltma önlemlerini incelemeye yönelik önlemlere tabi olan veri kümelerini işlemek ve dahil etmek

– Modelin yaşam döngüsü boyunca bağımsız uzmanların katılımıyla değerlendirmesini sağlama ve temel modeli performans, öngörülebilirlik, yorumlanabilirlik, düzeltilebilirlik, güvenlik ve siber güvenlik seviyelerine ulaşmak için tasarlamak, geliştirmek ve bu esnada test etmek

– Temel modeli madde 60'ta atıfta bulunulan AB veri tabanına kaydettirmek

143 European Commission (n 12)

144 ibid 1.

145 Yapay zekâ konusundaki regülasyonlarda "risk temelli" veya "hak temelli" olarak üzere iki ana bakış açısı bulunmakta olup, AB Yapay Zekâ Tüzük Tasarısı'nda benimsenen "risk temelli yaklaşım", yapay zekâ sistemlerinin regülasyon kapsamına sokulabilmesi için belirli bir oranda hak ihlali riski taşıması koşulunu esas alır. Bkz. Osman Gazi Güçlütürk and Yasin Murat Kadıoğlu, 'Yapay Zekâ ve Regülasyon' Eylem Aksoy Retornaz and Osman Gazi Güçlütürk (eds) *Gelişen Teknolojiler ve Hukuk II: Yapay Zekâ* (On İki Levha Yayıncılık 2021) 99.

146 Artifical Intelligence Act (n 143).

– Üretici yapay zekâ sistemlerinde kullanılan temel modellerin sağlayıcıları ile bir temel modelini üretici bir yapay zekâya dönüştüren sağlayıcıların da ayrıca Tüzüğün 52. maddesinde belirtilen şeffaflık yükümlülüklerine uyması şeklinde ifade edilmektedir.

AB Yapay Zekâ Tüzük Tasarısı m.52/3 ile şeffaflığa ilişkin bazı düzenlemeler de getirmektedir[147]. Tüzük, gerçek kişilerle etkileşime girmeyi ve içerik oluşturmayı amaçlayan belli yapay zekâ modellerinin, Tüzük kapsamında yüksek riskli sınıflandırmasında olup olmadığına bakılmaksızın belirli şeffaflık yükümlülüklerine tabi olması gerektiğini düzenlemektedir[148]. Tüzük Memorandum'unda oluşturdukları belirli manipülasyon riskleri göz önüne alınarak bu yükümlülüğün getirilmiş olduğu ifade edilmektedir. Bu kapsamda şeffaflık yükümlülükleri, insanlarla etkileşime giren, duyguları tespit etmeye yarayan, biyometrik verilere dayalı (sosyal) kategorilerle ilişkiyi belirlemek için kullanılan veya bir içerik üretebilen veya değiştiren modeller bakımından geçerli olacaktır[149].

AB Yapay Zekâ Tüzük Tasarısı m. 52/3 kapsamında, gerçek kişiler, koşullar ve kullanım bağlamından açıkça anlaşılmadığı sürece, bir yapay zekâ sistemi ile etkileşimde bulundukları konusunda bilgilendirilmelidir[150]. Konumuz bakımından ChatGPT, insanlarla etkileşime girebilmekte ve içerik üretebilmekte olup bu kapsamda bu konuda gerekli bildirim yükümlülüğünü haizdir. Şeffaflık yükümlülüğü getirilmesindeki amaç; bir yapay zekâ modelinin etkileşime girdiği kişileri bu konuda bilgilendirmektedir. Nitekim böyle bir uygulama kural olarak, etkileşime giren kişilerin bu durumun farkındalığı ile seçimler yapmasına ve eğer etkileşimde bulunmamayı tercih ederse bundan vazgeçmesine olanak sağlamaktadır[151].

İlgili düzenlemeler her ne kadar alanında bir ilk olsa da üretici yapay zekâ modellerindeki güvenlik ve mahremiyete dair temel sorunlar çözülmeden, ilgili düzenlemelerin de etkili bir yasal çözüm olması mümkün olamayacaktır. Bu sorunların çözülmesi ise; öncelikle yapay zekâ yarışındaki performans ölçütünün güvenlik ve mahremiyete odaklanması ile mümkün olabilir.

Kanaatimizce AB Yapay Zekâ Tüzük Tasarısı'nın bir diğer zayıf noktası; üretici yapay zekâ gibi yüksek riskli yapay zekâ modellerinin sağlayıcıları için öngördüğü ön denetim gibi yükümlülüklerin, mali yükünü karşılamanın küçük ve orta ölçekli sağlayıcılar için oldukça zor olması, bunun da sadece büyük oyuncuların pazarda hakimiyetlerinin daha da artması ile sonuçlanabilecek olmasıdır.

147 Tüzük Tasarısı m.52/3 şu şekildedir: "*Users of an AI system that generates or manipulates image, audio or video content that appreciably resembles existing persons, objects, places or other entities or events and would fasely appeat to a person to be authentic or truthful (deep fake) shall disclose that the content has been artifically generated or manipulated.*"

148 European Commission (n 12), Explanatory Memorandum, 5.2.4. 'Transparency Obligations For Certain AI Systems (Title IV) 34.

149 Explanatory Memorandum (ibid 148) 14; Bu modellerin kolluk kuvvetleri tarafından önleyici ceza hukuku kapsamında kullanılması son yıllarda ceza hukuku bakımından değerlendirilmektedir. Örneğin yüz tanıma teknolojileri ile kişilerin yüz ifadeleri, bakışları, jest ve mimikleri, duygu durumu ve vücut ısısı analiz edilebilmekte ve sisteme önceden aktarılan veriler ile elde edilen yeni veriler karşılaştırılarak potansiyel suçluların icra faaliyetlerine geçmeden tespit edilebileceği ifade edilmektedir. Bu konuda kapsamlı bir çalışma için bkz. Zafer İçer and Elif Dönmez, 'Yüz Tanıma Teknolojilerinin Önleyici Ceza Hukuku ve Ceza Muhakemesi Süreçlerindeki Kullanımı ve Sınırları' (2020) 15 (43) Ceza Hukuku Dergisi, 439.

150 Explanatory Memorandum (ibid 148) 34.

151 Explanatory Memorandum, (ibid 148) 14.

E. ROBOTLARA ENTEGRE GENİŞ DİL MODELLERİNDE MAHREMİYET VE GÜVENLİĞİN HUKUKİ BOYUTLARI

Üretici yapay zekâ modelleri esasında bilgisayar programıdır. Bu programlara uygun donanımların içine yerleştirilmesi de mümkündür. Dolayısıyla robotik sistemlere doğal dil işleme modelleri entegre edilebilir. Esasında insanlar ile sözel iletişim kuran robotlar, modelin içinde buna uygun sohbet robotu yazılımı olan donanımlar olup, üretici yapay zekâ modellerinden önce de toplum tarafından bilinirlik kazanmıştır. Buna bir örnek olarak *Sophia*[152] verilebilir. Bununla birlikte; üretici yapay zekâ modellerinin doğal dil anlama ve kullanma yetkinliklerinin giderek artması ile sıradan bir robot uygulamasından çok daha üstün şekilde insanlar ile dile dayalı etkileşim kurabilen bu modellerin, robotik sistemlere dahil edilmesi halinde veri gizliliğine dair riskler de doğacaktır.

Robotların veri gizliliği ile ilgili riskleri ve endişeleri artırmasının en temel nedenleri; robotların üzerlerinde taşıdıkları sensörler aracılığı ile çevrelerinden veri toplamaları, bu veriyi kaydetme, paylaşma işlevlerinin olmasıdır. Robotlar; hareket etme kabiliyetlerinin yanı sıra, insanın gidebildiği veya gidemediği yerlerde bulunabilmeleri ile aynı zamanda birer gözetim enstrümanı olarak da kullanılmaktadır[153].

Geniş dil modellerine dayalı üretici yapay zekâ modellerinin; veri gizliliği ile ilgili doğurduğu riskler, bu sistemlerin "insansı" (*humanoid*) robotların içine gömülü olması halinde ise daha da artabilir. *Calo* bu durumu, robotların sosyal anlamından ortaya çıkan gizlilik riski olarak nitelendirir[154]. Özellikle çocukların oyuncak veya "arkadaş" olarak etkileşimde bulunabileceği bu türdeki *humanoid* sistemler ile çocuklar herhangi bir oyuncak veya insan ile etkileşimde bulunur gibi, kendileri ve aileleri ile ilgili bilgileri paylaşabilir. İnsan-robot etkileşimi, insan gibi gözüken ve insan gibi konuşan bu tür modeller söz konusu olduğunda, sadece çocukların değil, bu modellerin robot olduğunun farkında olan yetişkinlerin dahi belirli bir yakınlık hissederek bu sistemler ile kişisel paylaşımlarda bulunabilmeleri mümkün olabilir[155]. Bahsedilen tüm bu riskler, başlı başına bir çalışma yapılabilecek kadar geniş olup, çalışma konumuz kapsamı içinde bu modeller bakımından da şeffaflık ile ilgili yasal yükümlülükler getirilmesinin gerekeceğini belirtmekle yetiniyoruz.

152 Sophia, *Hanson Robotics* şirketi tarafından geliştirilen ve ilk kez 14 Şubat 2016 yılında aktive edilen sosyal bir robottur. Ayrıntılı bilgi için bkz. Hanson Robotics, Sophia, <https://www.hansonrobotics.com/sophia/>, accessed 22 Jul 2023.

153 M. Ryan Calo, 'Robots and Privacy' in Patrick Lin, Keith Abney and George A. Bekey (eds) *Robot Ethics: The Ethical and Social Implications of Robotics* (The MIT Press, 2012), 187.

154 ibid 188, 198.

155 Aynı yönde bkz. Ronald C. Arkin and Lilia Moshkina, 'Affect in Human Robot Interaction' in Rafael A. Calvo (ed) *The Oxford Handbook of Affecting Computing* (Oxford University Press, 2015) 483.

III. CHATGPT VE ÜRETİCİ YAPAY ZEKÂ MODELLERİNDE MAHREMİYET VE GÜVENLİĞE DAİR HUKUKİ ÖNERİLER

Hendrycks[156], "*Felakete hazırlanmak aşırı kötümserlikten ziyade, sağduyulu bir harekettir*"[157] sözü ile; ChatGPT gibi üretici yapay zekâ modellerinin yaygın kullanımına başlandıktan sonra, birtakım risklerin çıkması durumunda, modellere ilişkin eylemin artık "tasarlama" aşamasından çıkarak, "yönlendirme" eylemine dönüştüğünü ifade etmektedir. Bu aşamada ise eylemlere ilişkin kapasitemizin giderek düştüğü[158], bu nedenle tüm hükümetlerin bu konuda yapacakları yasal düzenlemeler ile, yapay zekâda performans kadar güvenlik önlemlerine dair önlemleri yürürlüğe koymaları gerektiğini belirtmektedir[159].

Şeffaflık, yapay zekâ politikasının geliştirilmesindeki öncelikli konulardan biridir. Üretici yapay zekâ modellerinin yaygın kullanılmaya başlaması ile şeffaflığa ilişkin kaygılar artmaktadır[160]. Üretici yapay zekâ modellerinin yarattığı en temel sorunlardan biri ürettiği verilerin önyargılı (*bias*)[161] çıkarımlar üretme olasılığıdır. Modelin kendisinin fikir ve düşünceleri olmasa da, modeli geliştirirken kullanılan algoritmalar nesnel bir biçimde tasarlanmamış olabilir[162].

Modelin mahremiyet ve güvenlik endişelerinin giderilmesinin yollarından biri ChatGPT ve benzeri üretici yapay zekâ modellerinin eğitim veri setlerinin ve işleyişinin kamu ile paylaşılması bir başka deyişle şeffaflığı olabilir. Nitekim, kapalı kaynaklı bir üretici yapay zekâ modelinin geliştiricileri yeterli bilgiyi kamuoyu ile paylaşmadıkça modelin ne şekilde çalıştığı veya hangi veri setleri üzerinde eğitilmiş olduğunu bilmek de mümkün değildir[163].

Bunun yanında ChatGPT'nin veri sahiplerinin haklarını sağlamaya yönelik olan talepleri nasıl yerine getirebileceği hâlen belirsizlik taşımaktadır. OpenAI, ChatGPT'ye ilişkin olarak 25 Nisan 2023 tarihinde yayınladığı bilgi metninde; ChatGPT'ye sohbet geçmişini kapatma özelliği getirildiğini, sohbet geçmişinin devre dışı bırakıldığında başlatılan sohbetlerin modellerini eğitmek ve geliştirmek için kullanılmayacağını, bu tercihin tüm kullanıcılar bakımından her zaman değiştirilebilir olduğunu duyurmuştur. Bu özellik kullanıldığında, yeni konuşmaların yalnızca 30 gün boyunca saklanarak

156 Dan Hendrycks, Elon Musk'ın yapay zekâ şirketi XAI'in danışmanı ve Center For AI Safety (Yapay Zekâ Güvenliği Merkezi) direktörüdür. Center For AI Safety, yapay zekâ sistemlerinin toplumsal ölçekli risklerini azaltma misyonu ile kurulmuştur. Center for AI Safety, <https://www.safe.ai/> 22 Jul 2023.

157 Dan Hendrycks, 'Natural Selection Favors AIs Over Humans' (28 Mar 2023) <https://arxiv.org/abs/2303.16200> 22 Jul 2023.

158 ibid 12,13.

159 ibid 32.

160 Jacques de Werra, 'AI Transparency: An Emerging Principle in The IP Ecosystem?' (2023) Journal of Intellectual Property Law& Practice, <https://doi.org/10.1093/jiplp/jpad041> 1.

161 *Bias* terimi, ilk olarak 14. Yüzyılda geometride yatay çizgi anlamında, 19. Yüzyılda ise "yersiz önyargı" anlamına gelecek şekilde kullanılmıştır. Günümüzde makine öğrenmesi alanında bu terimin kullanımı ile kastedilen; bu sistemlerin tümevarım metodu sırasında karşılaştığı tekil örnekleri genelleştirirken yaptığı yanlış sınıflandırmadır. Crawford (n 95) 134. Bu durum, özellikle bireyler ile ilgili yapılan genellemelerde sosyal, ekonomik ve hukuki etkiler doğuracak şekilde ayrımcı sonuçlara yol açabilmektedir.

162 Norwegian Consumer Council (n 48) 15.

163 Norwegian Consumer Council (n 48) 12.

kalıcı olarak silmeden önce yalnızca kötüye kullanımın izleneceğini de ifade etmiştir[164]. Ancak bilgilendirme kapsamında daha evvel kullanıcılar tarafından paylaşılan kişisel veri kapsamındaki bilgilerin değiştirilmesi, silinmesi veya modellerin eğitiminde kullanılmamasına ilişkin bir imkân sağlanmamıştır.

Bu çekincelerin yanında OpenAI'ın kullanıcılar bakımından verilerinin ne kapsamda kullanıldığına ilişkin yeterince şeffaf olmadığı ve bu nedenle GDPR ile uyumunun sağlanmadığı da dile getirilmektedir[165]. Nitekim kişisel verilerin modellerin eğitildiği veri setlerinden silinmesiyle ilgili önemli bir engel veri setinin büyüklüğüdür. Veri setlerinin toplanması, temizlenmesi ve hazırlanması ile ilgili çalışmalar, modeli geliştirenler tarafından genellikle yeterli öncelikte olmayabilir[166].

Veri koruma kurallarının bir başka açmazı; veri koruma hukukunun temelinin kişisel veriye odaklanmış olmasıdır. Bu durumun bir sonucu olarak ilgili kişi ancak bireysel başvuru mekanizmasına bağlı olarak haklarını kullanabilir. Oysa toplumun geniş bir kesimi pek çok farklı nedenle kişisel verinin öneminin farkında değildir. Aksinin geçerli olduğu durumda dahi (bireylerin kendilerinden ne şekilde veri toplandığını, işlendiğini ve bunun önemini bildiklerini varsaydığımızda); bu farkındalığın karşılığı olarak bireylere tanınan hukuki başvuru imkanları yine de tartışılmalıdır. Çünkü bir kişinin kişisel verisi bir başka kişinin veri setinde yer alabilir. Örneğin akıllı telefon uygulamalarının telefon listesindeki kişilerin arkadaş olarak eklenebilmesi için gerektirdiği bir izin olan, telefon rehberine erişim izni verildiğinde, bu izni veren kişinin telefonundaki kişilerin listesi söz konusu kişilerin haberi dahi olmadan ve söz konusu uygulamayı kendileri kullanmasa dahi veri setine dahil olmaktadır. Geniş veri setlerini işleyerek yapay zekâ modellerini geliştiren ve devletler ile yarışacak seviyede ekonomik kazanç elde eden şirketler karşısında bireysel başvuru mekanizması ne kadar etkili ve adil sonuçlar doğurabilir?

Üretici yapay zekâ modellerinin gelişmesi ile bireyler tarafından verilerini korumak gittikçe daha çok zorlaşacak gibi görünmektedir. Nitekim, çok uluslu şirketler karşısında bulunan bireyin veri gizliliğini talep etmesindeki denge büyük ölçüde bozulmuştur. ChatGPT gibi sistemlerin katlanarak gelişmesi ile, bu sistemleri geliştiren şirketlerin yöneticileri; ulusal ve uluslararası alanda yapay zekâ konusunda yasal düzenleme yapacak otoriteleri yasal düzenlemenin nasıl olması gerektiği konusunda ve hatta yasal düzenleme yapılıp yapılmaması konusunda etkileme gücüne kavuşmaktadırlar. Çünkü bir yandan tüm dünyada giderek kötüleşen ekonomik koşullar ve giderek dijitalleşen hizmetler çağında, devletler ve hatta uluslararası kuruluşlar, büyük teknoloji şirketlerinin sunduğu bulut veya veri analitiği gibi hizmetlere giderek daha fazla ihtiyaç duymaktadır. Bu durumun bu şirketler ile yasa koyucular arasında bir tür evrensel menfaat çatışması yarattığı dahi söylenebilir. Şöyle ki; belirli sınırlar ve yaptırımlar ile hukuki bir çerçeve oluşturulacak alanın en büyük temsilcileri, bu hukuki çerçeveyi oluşturacak otoriteler açısından iş birliği yapılması zorunlu aktörler haline gelmektedir.

164 OpenAI (n 105).

165 Norwegian Consumer Council (n 48) 46.

166 Norwegian Consumer Council, (n 48) 46.

Avrupa Konseyi Yapay Zekâ Ad Hoc Komitesi (CAHAI) "Yapay Zekâ Sistemlerinin Regülasyonuna Doğru"[167] başlıklı 2020 tarihli raporundaki "Dijital Güç Yoğunlaşması" bölümünde tam da bu konuda önemli bir uyarıda bulunmaktadır. Veri saklanması ve yapay zekâ teknolojilerinin üretiminin sınırlı sayıdaki aktöre koşulsuz teslimi, insan hakları açısından telafi edilemez zararlar doğurabilecektir: *"Eğer politik güç, toplum yararı yerine hissedarların çıkarlarına öncelik veren birkaç özel kişinin ellerinde yoğunlaşırsa, bu durum demokratik devletlerin otoritesini tehdit edebilir."* Esasen, gerek ülkemizde gerekse de AB'de yürürlükte bulunan veri koruma kurallarının dahi, büyük yapay zekâ modelleri geliştiren şirketler açısından bir tür "öde ve geç" mekanizmasını dolaylı olarak yaratmış olduğu da söylenebilir. Özellikle, devletlerin dahi hizmet sunmada ihtiyaç duyduğu bu tür büyük şirketler için veri koruma yasalarındaki parasal cezalar ne denli yüksek olursa olsun caydırıcı olmamaktadır.

Doktrinde *van der Sloot*; sadece bireysel değil bir gruba ait olmaya dair farklı ayrımcılık düzenlemeleri yapılmasını önermektedir[168]. *Wachter*; bu konuda benzer bir görüş ileri sürerek, benzerlik/ilginlik *(affinity)* konusundaki bir ayrımcılık unsurunun gerek doğrudan gerek dolaylı ayrımcılık açısından var olan hukuki korumanın sürdürülebilir olması için bir çözüm önerisi olarak ortaya koymaktadır[169]. *Hildebrandt* ise; sadece kişisel veriye korunmaya değer bir hukuki statü tanıyan mevcut veri koruma kurallarının kapsamının genişleterek, kişisel veriler ile oluşturulan grup profillerinin de etkili bir hukuki statü kazanması gerektiğini savunmaktadır[170]. Nitekim; bireyi veri koruma düzenlemelerinin merkezi konumuna alan bakış açısının yanında aynı zamanda toplumun tamamını ilgilendiren bazı konularda da verinin nasıl yönetileceği ile ilgili o grubun bireyleri hak sahibi olmalıdır. Tıpkı bireylerin mahremiyetini korumayı amaçlayan düzenlemeler gibi bir grubu oluşturan bireyler de grubu tanımlayan bilgilerin nasıl oluşturulduğu ve kullanıldığı konusunda kolektif bir menfaate sahip olacaktır. Bu anlamda, grup mahremiyeti (*group privacy*) kavramı son yıllarda ortaya çıkan bir yaklaşımdır[171]. Veri toplama ve işleme faaliyetlerinin eskiye oranla daha kolay ve ucuz olması ile büyük ve tanımlanmamış gruplar hakkında veri toplanabilmektedir. Toplanan verilerde her ne kadar kişisel veri niteliğindeki veri yer almasa bile kullanıcıların davranışlarını ve faaliyetlerini bir düzeyde yansıtan davranışsal ve ayrıntılı veriler yer alır[172]. Bu veriler de grubu oluşturan tüm bireyleri ilgilendiren bazı konularda kullanılabilmektedir. Oysa yalnızca kişisel verilerin korunmasına ilişkin düzenlemeler bir grubun veriyle ilişkisini korumaktan uzaktır.

167 CAHAI, Towards Regulation of AI Systems, https://rm.coe.int/prems-107320-gbr-2018-compli-cahai-couv-texte-a4-bat-web/1680a0c17a, 21 Agu 2023.

168 Bart van der Sloot, 'The Individual in the Big Data Era: Moving Towards an Agent-Based Paradigm' in Bart van der Sloot, Dennis Broeders, Erik Schrijvers (eds) *Exploring the Boundaries of Big Data* (Amsterdam University Press, 2016) 196-199.

169 Sandra Wachter, 'Affinity Profiling and Discrimination by Association in Online Behavioural Advertising' (2020) 35(2) Berkeley Technology Law Journal, 369.

170 Bkz. Hildebrandt (n 113).

171 Brent Mittelstadt, 'From Individual to Group Privacy in Big Data Analytics' (2017) 30 Philos. Technol. 476; Linnet Taylor, Luciano Floridi and Bart van der Sloot, 'Introduction: A New Perspective on Privacy', Linnet Taylor, Luciano Floridi and Bart van der Sloot (eds) *Group Privacy New Challenges of Data Technologies*: (Springer, 2017) 2; Veale, Binns ve Edwards, (n 120) 3.

172 Taylor, Floridi and van der Sloot (n 171) 3.

Kanaatimizce de yapay zekâ dil modellerinin geldiği bu seviyede tartışılması gereken en önemli şey; tasarlanan modelin kendisinden evvel, bu modellere gelişmesini sağlayan tüm insanlığın verisini, tekil değil fakat kolektif bir şekilde kategorize ederek, bu kategorilere hukuki bir statü verip vermemek gerektiğidir.

Bireyler, kendi verilerini kullanarak kendileri ile ilgili çıkarımlar ortaya koyan yapay zekâ algoritmalarının sonuçlarını ve kendi verirlerini kontrol edebilir olmalıdır[173]. İnsanlığın, salt veri-üreten yığınlara[174] indirgenmemesi için veri kontrolünün kimde olduğunun hukuken belirlenmesi önemlidir. Bu durum, yasal düzenlemeler yanında, veri korumaya ilişkin toplumda bilinç oluşturulması ve gerekirse veri birlikleri oluşturulması ile çözülebilir. Böylelikle, bireylerin yanı sıra, kişi topluluklarına, örneğin veri işlenmesi sonucu dahil edildikleri kategorilere göre topluluk davası açma gibi etkili hukuki başvuru imkanları sağlanması mümkün olabilir.

Veri işleme faaliyetlerinin çokluğu ve veri tabanlarının büyüklüğü göz önünde bulundurulduğunda; bir bireyin kendi verilerini içerebilecek her veri işleme faaliyetinden haberdar olması, işlemenin ne kadar meşru bir şekilde yapıldığını değerlendirmesi ve eğer değilse veri sorumlusundan faaliyetlerini durdurmasını talep etmesi ya da yasal haklarını kullanması son derece güç bir hâl almaktadır[175]. Bu duruma bir örnek olarak, bir kullanıcının bir sağlık hizmeti sohbet botu ile etkileşime girdiğinde, sunduğu erişim bilgilerinin saklanabileceğini ve üretici yapay zekâ modelinin yeniden eğitilmesi veya diğer ticari amaçlar için kullanılabileceğinin bilincinde olmaması verilebilir. Böyle bir durumda özel nitelikli kişisel verilerini potansiyel kullanımını bilmeden paylaşabilir. Oysa mevcut birçok sohbet botu; kullanıcı verilerini hizmetlerini geliştirmek ve iyileştirmek amacıyla yeniden kullanılmasına izin veren hizmet şartlarına sahiptir. Aynı tartışmalar diğer özel nitelikli kişisel verilerin kullanılabileceği başka alanlardaki sohbet botları için de geçerlidir[176].

Ekonomik gücün az sayıdaki birkaç büyük şirkette yoğunlaşması ile bu şirketler yasama faaliyetleri, hukuki uyuşmazlıklar ve ulusal politikalar üzerinde dahi etkili olacak siyasi bir güce sahip olmaya başlamışlardır. Devlet ve vatandaş ilişkisine artık şirket de eklenmiş ve büyük şirketler bugün toplumun politik bir bileşeni haline gelmiştir[177]. Bu durum *Lyon*'un ifadesiyle yurttaşlıktan tüketiciliğe geçişin tescillenmesini olarak ifade edilmektedir[178]. Öyle ki, bu anlayış son dönemde Avrupa Birliği'nin yeni politika belgelerine veya regülasyon metinlerine de yapısal olarak yansımıştır. Anılan metinlerde, "bireyler" veya "toplum" kavramlarından ziyade "tüketiciler" kavramının kullanılması dikkat

173 Finlandiya devlet televizyonu YLE Yapay Zekâ ve Kişiselleştirme Birimi başkanı Jarno Koponen tarafından "artırılmış temsil" (augmented agency) olarak adlandırılan bu yöntemle, insan odaklı bir kişiselleştirme (çeşitli konularda, her bireyin verileri ile saptanan tercihlerinin göz önüne alınması) sağlanabilecektir. Jarno Koponen, 'Personalization and Augmentation' in Esko Kilpi (ed) *Perspectives on New Work* (Sitra, 2016) 70.

174 Batukan, yapay zekâ ve diğer gelişen teknolojilerin kontrolünün bir veya birkaç kişi veya grubun elinde olması ile insanın doğumdan önce başlamak üzere sürekli denetlenebilir bir varlık haline indirgenebileceği riskine dikkat çekmektedir. Can Batukan, 'XXI. Yüzyılın İktidar Biçimlerine Foucaultcu Bir Bakış' (2020) 18 (69) Eğitim Bilim ve Toplum Dergisi, 16, 35.

175 Taylor, Floridi ve van der Sloot (n 171) 5.

176 Bush (n 48) 5.

177 Wolfgang Streeck, *Satın Alınan Zaman* Kerem Kabadayı (çev) (Koç Üniversitesi Yayınları 2016) 63.

178 David Lyon, *Vesikalı Yurttaş* Barış Baysal (çev) (Kalkedon Yayınları 2012) 217.

çekicidir. Bireyin veya vatandaş kimliğinin tüketici kimliğinde erimesi sadece politik açıdan değil, hukuki açıdan da temel hak ve demokratik topluma zarar vericidir. Zira, teknoloji ile temel haklar arasındaki ilişkiyi üretici ve tüketici ekseninde ele almak, ulus devletlerin bir zamanlar dayatmaya güçlerinin yettiği yasal sınırları tanımayan bir teknoloji politikasına yol açabilecektir[179].

Öte yandan; verinin "alınıp satılabilen"[180] bir ticari nesne olması olgusu da bir sorun olarak karşımıza çıkmakta ve aslında çözümü de içerisinde saklamaktadır. Şöyle ki; veri sırandan bir ticari nesne olmayıp, çalışma konumuzu da oluşturan, dünyayı değiştirici yeni yapay zekâ modellerinin ham maddesidir. Verinin bir mülkiyet hakkı olarak nitelendirilmesi, pek çok sorunun kolayca çözülmesini sağlayabilecek gibi gözükse de bir takım hukuki, ekonomik ve politik nedenler ile aksini ileri sürmek de mümkündür[181]. Şöyle ki, ekonomik bir yaklaşımla kişisel verilerin ticari bir nesneye dönüşmesi halinde; bireyin giderek karmaşıklaşan teknoloji ve bu karmaşık teknolojiye dair hukuki düzenlemeleri anlamaması ve belirli hizmetleri alabilmek için verisini hür iradesi ile "bağışlaması", hatta gelir elde etmek için "satması" dahi mümkün olabilecektir.

SONUÇ

ChatGPT ve üretici yapay zekâ modelleri mahremiyet ve güvenlik riskleri barındırmaktadır. Bu güvenlik riskleri, üretici yapay zekâ modellerinin etki alanları düşünüldüğünde ivedilikle belirlenerek gerekli yasal düzenlemeler yapılmalıdır. İlgili yasal düzenlemeler, modelin kendisi bakımından olduğu kadar, bu modellerin işler hale gelmesindeki en önemli unsur olan veri bakımından önceliklendirilmelidir. Nitekim, ChatGPT gibi üretici yapay zekâ modellerin daha işler olmasını sağlayan şey veridir. Nitekim; birey, toplum ve tüm insanlık için henüz öngörülemeyen olumsuz etkiler yaratabilecek olan ChatGPT gibi üretici yapay zekâ modellerin risklerinin en aza indirgenmesi ile tüm insanlığa faydalı ve güvenilir bir yapay zekâ hedefine varılabilmesi; ancak, mahremiyet ve güvenlik alanında olmak üzere yasal düzenlemelerin ivedilikle alınması ile mümkün olur.

Veri koruma ile ilgili yürürlükte olan yasal düzenlemelerdeki "veri sorumlusu" kavramı ile AB'nin yapay zekâ modellerine dair yasa tasarılarında yer alan ve "yapay zekâ sistemi sağlayıcısı" sıfatlarında öngörülen yükümlülükler bulunmaktadır. Ancak, yapay zekâ modellerini geliştirenler, öngörülen yasal yükümlülüklere uysalar dahi, ChatGPT gibi geniş yapay zekâ modellerinin yaratacağı

179 ibid 217.

180 Bireylerin çevrimiçi olarak yaptıkları ve izledikleri herşey, neredeyse tüm web sitelerinde ve uygulamalarda perde arkasında çalışan ve "gerçek zamanlı teklif" (real time bidding) adı verilen bir çevrimiçi reklamcılık sisteminden toplanır. Bir ticari web sitesine bir sayfa yüklenildiğinde veya bir uygulama kullanıldığında, bireylere hangi ilan veya reklamın gösterileceğini belirleyen perde arkasında yüksek hızlı bir reklam açık artırması gerçekleşir. Bu açık artırma ile, bireylerin çevrimiçi olduklarında yaptıklarına ve nerede olduklarına dair mahrem bilgiler, reklam fırsatları için teklif alınmak üzere pek çok farklı şirket ile paylaşılır. Irish Council For Civil Liberties – ICCL, https://www.iccl.ie/rtb-june-2021/ accessed 20 Feb 2022. Örneğin Google, 9,8 milyon web sitesi üzerinde eş zamanlı teklif vermekte ve kullanıcılardan edindiği verileri kendi açıklamasına göre 1042 farklı şirket ile paylaşmaktadır. Google, Reklam Yöneticisi ve Ad Exchange Program Politikaları Reklam Teknolojisi Sağlayıcıları, <https://support.google.com/admanager/answer/9012903> accessed 13 Mar 2023.

181 Kişisel verilerin korunmasına dair ekonomik hak yaklaşımı ve insan hakkı yaklaşımı konusundaki ayrıntılı inceleme ve farklı görüşler için bkz. Küzeci (n 116) 67 vd.

mahremiyet ve güvenlik riskleri bertaraf edilemeyecektir. Nitekim; bu yasal düzenlemelerdeki yer alan hukuki, teknik ve idari tedbirler hem hukukçuların hem de mühendislerin yasal uyuma fazlası ile odaklanması ile sonuçlanarak asıl meseleden bizleri uzaklaştırmıştır. Okunması mümkün olmayan uzunlukta bilgilendirme metinlerine dayanılarak kişisel verinin işlenmesine rıza gösterilmesinin bir tür yasal uyuma dönüşmesi bunun en çarpıcı kanıtıdır.

Veri koruma düzenlemeleri ve AB Yapay Zekâ Tüzük Tasarısı'nda sorumlular bakımından belirlenen, mali yaptırımların en yükseği dahi, günümüzde bir devletten bile daha büyük ekonomik güce sahip olabilen büyük teknoloji şirketleri için ödenmesi kolay bir bedeldir. Asıl meselemiz, yapay zekâ alanındaki kıyasıya rekabetin performansa odaklanmış olması nedeni ile güvenlik ve mahremiyetin göz ardı edildiği ve birkaç büyük teknoloji şirketi ve bu alandaki büyük oyunculardan olan birkaç devletten oluşan bir azınlık tarafından kuralların belirlendiği bir geleceğe doğru insanlık olarak hızla ilerliyor olduğumuzdur.

Bu çalışmamızda ortaya çıkan gerçeklerden yola çıkarak önerdiğimiz çözüm yollarından belki de yasal düzenlemeler ile en kolay şekilde alınabilecek olanı; veri koruma düzenlemelerine bireylerin yanı sıra insanlığın kolektif menfaatini de koruyacak biçimde düzenlemeler getirilmesi ve kolektif menfaati olanlar bakımından bazı haklar tanınmasıdır. Nitekim; toplumsal açıdan; bireylerin tek başlarına yasal düzenlemeleri anlamaları, yasal haklarının farkında olacak şekilde bilinçlenmeleri, verdikleri veri işlemeye dair rızanın kendileri ve yakın çevreleri ile ilgili yakın veya uzak gelecekte ne tür etki yaratacağını öngörebilmeleri, yasal haklarının ihlal edildiğinin farkına vararak yasal başvuru haklarını kullanmaları gibi eylemler için eğitilerek bilinç sahibi olmaları çok önemli, ancak oldukça zaman alıcı bir eylemdir.

Geniş dil modellerindeki gibi birkaç ayda dahi akıl almaz ilerlemeler kaydeden bir teknoloji karşısında, bireylerin tek başına eğitim ile bilinçlenerek yasal haklarının farkında olmalarının mahremiyet ve güvenlik ile ilgili sorunları tek başına çözebileceği önermesi ne yazık ki artık gerçekçi bir çözüm olmaktan uzaktır. Ancak, veri koruma düzenlemelerindeki veri öznesinin yasal hakları (okuduğunu anlayarak kişisel verisinin işlenmesine rıza göstermesi dahil olmak üzere) tam da böylesi bir eğitimli bilinç ile kullanılabilecek haklardır. Mahremiyet ve güvenliğe dair bireylere eğitim verilmesi zorunludur ve hem kapsam hem de süre bakımından artarak zorunlu eğitimin bir parçası olmalıdır. Ancak, bir yandan bireyler bu şekilde eğitilirken diğer yandan teknolojik gerçeklere uygun farklı yasal düzenlemeler ile teknolojik gerçeklere uygun önlemler alınmalıdır. Daha önce önerilen ve veri işlenmesi ile profillenen bireylerin profilleme kategorilerine göre bir tür "grup mahremiyeti" hakkına sahip olarak yasal haklarını bu grup statüsünde toplu olarak kullanmaları dahi yeterli yasal ve teknolojik bilgi ile gelen bir bilinç seviyesini gerektirmektedir ve bu yaklaşım ne yazık ki gerçekçi değildir.

Mahremiyet ve güvenliğin artık, ivedi ve kolektif bir sorun olması ve buna uygun yasal çözümler getirilmesi teknik gerekçesi ise; "kişisel veri" odaklı yasal düzenlemelerin mahremiyeti ve kişisel verileri korumada artık teknik nedenlerle de yeterli olmamasıdır. Şöyle ki, bir kişinin hassas nitelikteki kişisel verileri dahi, bir başka kişinin veri setinde (telefonunda, sosyal medya içeriklerinde, sağlık

verisinde, DNA'da) bulunabilmekte ve büyük veri setleri ile gelişen geniş dil modelleri gibi yapay zekâ modellerini eğitmekte kullanılabilmektedir. Bu nedenle tek bir bireyin kendi kişisel verilerini korumasına dair farkındalığı veya yasal yollara başvurarak kendi kişisel verisinin ihlaline yönelik tedbirleri talep etmesi halinde dahi tam bir yasal koruma kapsamında olduğunu söylemek artık mümkün olmayacaktır. Bu nedenle mahremiyetin aynı zamanda kolektif bir hak olarak kabulüne dayanan yasal düzenlemelere ihtiyaç bulunmaktadır.

Yapay zekâya ilişkin olan yasal düzenlemeler bakımından ise; bu tür yasal düzenlemeler ile sorumlular bakımından getirilecek zorunlu önlemlere, sorumluların teknik ve hukuki açıdan uyabilmek için getireceği bütçe bakımından değerlendirilmesi gerekir. Nitekim, ilgili sorumluların ayırması gereken bütçenin, küçük ve orta ölçekli şirket düzeyindeki yapay zekâ sistemi sağlayıcıları için karşılanması imkansız nitelikte olması da muhtemeldir. Bu durum da toplumun tekelleşen büyük teknoloji şirketlerinin bulut ve diğeri ilgili hizmetlerine daha bağımlı hale gelmeleri ile sonuçlanabilecek olması gerçeğinin göz önüne alınması gerekliliğidir.

KAYNAKÇA

Aggarwal CC, *Recommender Systems* (Springer, 2016)

Alpaydın, E, *Yapay Öğrenme: Yeni Yapay Zekâ*, Aylin Ağar (çev) (Tellekt 2020).

Alpaydın E, *Machine Learning: The New AI* (The MIT Press, 2016).

Altman S, 'Planning For AGI and Beyond' (24 Feb 2023) <https://openai.com/blog/planning-for-agi-and-beyond> accessed 05 July 2023.

Anderljung M, Barnhart J, Korinek A and Leung J, 'Frontier AI Regulation: Managing Emerging Risks to Public Safety' (11 Jul 2023) <https://arxiv.org/pdf/2307.03718.pdf> 6, accessed 22 Jul 2023.

Angioloni L, Borghuis T and Brusci L (2020) 'Conlon: A pseudo-song generator based on a new pianoroll, wasserstein autoencoders, and optimal interpolations' (Research Gate, October 2020) 876 >https://www.researchgate.net/publication/348909705_Conlon_A_pseudo-song_generator_based_on_a_new_pianoroll_wasserstein_autoencoders_and_optimal_interpolations> accessed 22 July 2023.

Arkin RC and Moshkina L, 'Affect in Human Robot Interaction' in Rafael A. Calvo (ed) *The Oxford Handbook of Affecting Computing* (Oxford University Press, 2015)

Article 29 Data Protection Working Party, 'Guidelines on Automated individual decision-making and Profiling for the purposes of Regulation 2016/679 (wp251rev.01)' 22.08.2018, <https://ec.europa.eu/newsroom/article29/items/612053/en> s. 6-7, accessed 30 Jul 2023

Avrupa Konseyi, 'Guidelines On The Protection of Individuals With Regard to the Processing of Personal Data In A World Of Big Data'<https://rm.coe.int/CoERMPublicCommonSearchServices/DisplayDCTMContent?documentId=090.000.16806ebe7a> accessed 22 July 2023.

Bank of England (2019), 'Machine Learning in UK Financial Services', Working Paper (Bank of England, 16 October 2019) <https://www.bankofengland.co.uk/report/2019/machine-learning-in-uk-financial-services> accessed 22 July 2023.

Batukan C, 'XXI. Yüzyılın İktidar Biçimlerine Foucaultcu Bir Bakış' (2020) 18 (69) Eğitim Bilim ve Toplum Dergisi.

Berber LK, 'Çapraz Etkileşim: Mahremiyete İlişkin Mevzuat ve Mahremiyet Standartları Arasındaki İlişki', in Leyla Keser Berber and Ali Cem Bilgili (eds) *Güncel Gelişmeler Işığında Kişisel Verilerin Korunması Hukuku* (Oniki Levha Yayınları 2020).

Bowman SR, 'Eight Things to Know about Large Language Models' (Arxiv 2 April 2023) <https://arxiv.org/pdf/2304.00612.pdf>, accessed 22 July 2023.

Buchanan B and Wright D, 'The Impact of Machine Learning on UK Financial Services' (2021) 37(3) Oxford Review of Economic Policy

Matt Burgess, 'ChatGPT Has a Big Privacy Problem', (Wired, 4 Nisan 2023) <https://www.wired.com/story/italy-ban-chatgpt-privacy-gdpr/> accessed 22 July 2023.

Bush KE, 'Generative Artificial Intelligence and Privacy: A Primer' (Congressional Research Service, 23 May 2023), USA Congressional Research Service Report, >https://crsreports.congress.gov/product/pdf/R/R47569> accessed 28 June 2023.

CAHAI, Towards Regulation of AI Systems, https://rm.coe.int/prems-107320-gbr-2018-compli-cahai-couv-texte-a4-bat-web/1680a0c17a, 21 Agu 2023

Cahit Arf, 'Makine Düşünebilir mi ve Nasıl Düşünebilir?' Tarık Tuna Gözütok (ed), *Cahit Arf ve Atatürk Üniversitesindeki Halk Konferansları*, 1958-1960 (1959) (Atatürk Üniversitesi Yayınları 1959).

Calo MR, 'Robots and Privacy' in Patrick Lin, Keith Abney and George A. Bekey (eds) *Robot Ethics: The Ethical and Social Implications of Robotics* (The MIT Press, 2012)

Chandran R, 'Indigenous Groups İn NZ, US Fear Colonisation as AI Learns Their Languages' (03 Ap 2023) <https://www.context.news/ai/nz-us-indigenous-fear-colonisation-as-bots-learn-their-languages> accessed 14 Jul 2023.

Clarke R, 'Profiling: A Hidden Challenge to the Regulation of Data Surveillance' (1993) Journal of Law and Information Science 4 (2)

Council Of European Union, 'ChatGPT in the Public Sector – Overhyped or overlooked?'(2023) <https://www.consilium.europa.eu/media/63818/art-paper-chatgpt-in-the-public-sector-overhyped-or-overlooked-24-april-2023_ext.pdf> accessed 22 July 2023.

Council of Europe, The Protection Of Individuals with Regard to Automatic Processing of Personal Data In The Context Of Profiling Recommendation CM/Rec(2010)13 and Explanatory Memorandum (2011) 9 <https://rm.coe.int/16807096c3#:~:text=Recommendation%20CM%2FRec(2010)13%2C%20adopted%20by%20the,Data%20(T%2DPD)>, accessed 21 Ag 2023

Crawford K, *Atlas of AI*, (Yale University Press 2021).

Crawford K and Paglen T, 'Excavating AI The Politics of Images in Machine Learning Training Sets' (2019), <https://excavating.ai/> accessed 22 July 2023

Daniel W, 'Meet IndexGPT, the AI stock picker JP Morgan is developing that may put your financial advisor out of business' (Fortune, 26 May 2023) <https://fortune.com/2023/05/26/jpmorgan-indexgpt-a-i-stock-picker/> accessed 22 July 2023.

David Deutsch, *The Beginning of Infinity: Explanations That Transform the World* (Allen Lane 2011).

Denemeç IS, *To Feed or Not to Feed?:An Analysis of the Copyright Issues Surrounding the Use of Machine Learning Algorithms* (Lykeion 2021)

Dwork C, McSherry F, Nissim K and Smith A, 'Calibrating Noise to Sensitivity in Private Data Analysis' in S Halevi and T. Rabin (eds) *Theory of Cryptography* (Springer, 2006) <https://people.csail.mit.edu/asmith/PS/sensitivity-tcc-final.pdf> accessed 22 Jul 2023.

Ebers M, 'Regulating A and Robotics: Ethical and Legal Challenges, Algorithms and Law' Martin Ebers ve Susana Navas Navarro (eds) *Algorithms and Law* (Cambridge University Press, 2020).

El-Mahamdi E, Farhadkhani S, Guerraoui R et al. 'On the Impossible Safety of Large Language Models', (Arxiv, 9 May 2023), <https://arxiv.org/abs/2209.15259> accessed 22 July 2023

Erwig M, *Once Upon An Algorithm: How Stories Explain Computing* (The MIT Press 2017).

European Commission, Proposal for 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, 21.4.2021, COM(2021) 206 final, 2021/0106 COD <https://eur-lex.europa.eu/legal-content/EN/TXT/PDF/?uri=CELEX:52021PC0206&from=EN> accessed 22 Jul 2023.

Hayden Field, 'The First Fully A.I.-Generated Drug Enters Clinical Trials in Human Patients' (CNBC 29 June 2023) <https://www.cnbc.com/2023/06/29/ai-generated-drug-begins-clinical-trials-in-human-patients.html> accessed 22 July 2023.

Floridi L and Chiriatti M, 'GPT-3: Its Nature, Scope, Limits, and Consequences' (2020) <https://dx.doi.org/10.2139/ssrn.3827044> accessed 22 Jul 2023.

Freud S, *A Complete Introductory Lectures on Psychoanalysis*, James Strachey (çev) (W.W. Norton 1966).

Garante Per La Protezione Dei Dati Personali (31 Mar 2023) <https://www.gpdp.it/web/guest/home/docweb/-/docweb-display/docweb/9870847#english> accessed 21 Ag 2023

Gençoğlu O, 'For Decades of AI Compute' (Laconic 27 Feb 2023) <https://www.laconic.fi/ai-compute/> accessed 22 Jul 2023.

Ginsburg J and Budiardjo LA, 'Authors and Machines' (2019) 34 (2) Berkeley Technology Law Journal <http://dx.doi.org/10.2139/ssrn.3233885> accessed 20.7.2023

Goodfellow I, Bengio Y and Courville A, *Deep Learning* (The MIT Press, 2016)

Google, 'Introducing PaLM 2', (Google 10 Mayıs 2023) <https://blog.google/ technology/ai/google-palm-2-ai-large-language-model/> accessed 22 July 2023.

Google, Reklam Yöneticisi ve Ad Exchange Program Politikaları Reklam Teknolojisi Sağlayıcıları, <https://support.google.com/admanager/answer/9012903> accessed 13 Mar 2023.

Göktaş P, Karakaya G, Kalyoncu AF and Damadoğlu E, Artificial Intelligence Chatbots in Allergy and Immunology Practice: Where Have We Been and Where Are We Going?' (2023) The Journal of Allergy and Clinical Immunology: In Practice, <https://www.sciencedirect.com/science/article/pii/S221.321.9823006414>, accessed 21 Agu 2023

Güçlütürk OG and Kadıoğlu YM, 'Yapay Zekâ ve Regülasyon' Eylem Aksoy Retornaz and Osman Gazi Güçlütürk (eds) *Gelişen Teknolojiler ve Hukuk II: Yapay Zekâ* (On İki Levha Yayıncılık 2021)

Gürses S and Preneel B, 'Cryptology and Privacy in the Context of Big Data' in Bart van der Sloot, Dennis Broeders and Erik Schrijvers (eds) *Exploring the Boundaries of Big Data* (Amsterdam University Press, 2016)

Hanson Robotics, Sophia, <https://www.hansonrobotics.com/sophia/>, accessed 22 Jul 2023.

Harris LA, 'Generative Artificial Intelligence: Overview, Issues, and Questions for Congress' (Congressional Research Service 9 June 2023) <https://crsreports.congress.gov/product/pdf/IF/IF12426> accessed 28 June 2023.

Hendrycks D, 'Natural Selection Favors AIs Over Humans' (28 Mar 2023) <https://arxiv.org/abs/2303.16200> 22 Jul 2023

Hiller LA and Isaacson L, *Experimental Music Composition with an Electronic Computer* (McGraw – Hill Book Company 1959).

Huang S and Siddarth D, 'Generative AI and the Digital Commons' (Arxiv 20 March 2023), >https://doi.org/10.48550/arXiv.2303.11074> accessed 22 July 2023

Hunt EB, *Artificial Intelligence* (Academic Press 1975)

Irish Council For Civil Liberties – ICCL, https://www.iccl.ie/rtb-june-2021/ accessed 20 Feb 2022

Isaacson W, *Geleceği Keşfedenler* Duygu Dalgakıran (çev) (Domingo 2014)

İçer Z and Dönmez E, 'Yüz Tanıma Teknolojilerinin Önleyici Ceza Hukuku ve Ceza Muhakemesi Süreçlerindeki Kullanımı ve Sınırları' (2020) 15 (43) Ceza Hukuku Dergisi.

Kaku M, *Olanaksızlığın Fiziği,* Engin Tarhan (çev) (ODTÜ Yayıncılık 2014)

Kirchengast T, 'Deepfakes and image manipulation: criminalisation and control' (2020) 29 (3) Information& Communications Technology Law.

Kişisel Verileri Koruma Kurumu, 'Kişisel Verilerin Korunması Kanununa İlişkin Uygulama Rehberi' (2019) https://www.kvkk.gov.tr/SharedFolderServer/CMSFiles/41784a70-2bac-4e4a-830f-35c628468646.PDF, accessed 22 Agu 2023.

Kişisel Verileri Koruma Kurumu, Aydınlatma Yükümlülüğünün Yerine Getirilmesinde Uyulacak Usul ve Esaslar Hakkında Tebliğ, R.G. 10.03.2018, Sayı: 30356

Kişisel Verileri Koruma Kurumu, 'Kişisel Veri Güvenliği Rehberi Teknik ve İdari Tedbirler' (2018) <https://www.kvkk.gov.tr/yayinlar/veri_guvenligi_rehberi.pdf> 28 Jun 2023.

Kızrak A, 'Nesne Algılama Yaklaşımlarına Transformer Devrimi' (Medium, 10 Ap 2021) <https://ayyucekizrak.medium.com/nesne-alg%C4%B1lama-yakla%C5%9F%C4%B1mlar%C4%B1na-transformer-devrimi-baf583a29a23> accessed 22 Jul 2023.

Konečný J, Brendan H, Ramage D and Richtárik P, 'Federated Optimization: Distributed Machine Learning for On-Device Intelligence' (2016) <https://doi.org/10.48550/arXiv.1610.02527> accessed 22 Jul 2023.

Koponen J, 'Personalization and Augmentation' in Esko Kilpi (ed) *Perspectives on New Work* (Sitra, 2016)

Kotler P, *Marketing Management* (Pearson Education 2015)

Kulular İbrahim MA, *Robo Danışmanların Hukuken Değerlendirilmesi* (Adalet Yayınevi 2023)

Kurzweil R, *İnsanlık 2.0: Tekilliğe Doğru Biyolojisini Aşan İnsan*, Müge Şengel (çev) (Alfa Yayınları 2017)

Küzeci E, *Kişisel Verilerin Korunması* (On İki Levha, 2020)

Lee A, 'What is a Pre-Trained AI Model?' (NVIDIA, 8 Dec 2022) <https://blogs.nvidia.com/blog/2022/12/08/what-is-a-pretrained-ai-model/> accessed 22 Jul 2023.

Leskovec J; 'Enron Email Network' (Stanford University) <https://snap.stanford.edu/data/email-Enron.html> accessed 22.7.2023; Bryan Klimt and Yiming Yang, 'Introducing the Enron Corpus' (2004), International Conference on Email and Anti-Spam (Natural Language Server 2021) >https://nl.ijs.si/janes/wp-content/uploads/2014/09/klimtyang04a.pdf> accessed 23.7.2023.

Lovelace A, 'Notes of the Translator' for 'Sketch of the Analytical Engine Invented by Charles Babbage' (Taylor and Francis 1842) 722, <https://repository.ou.edu/uuid/6235e086-c11a-56f6-b50d-1b1f5aaa3f5e#page/1/mode/2up> accessed 22 Jun 2023.

Meta AI, 'Meta and Microsoft Introduce the Next Generation of Llama' (18 July 2023) https://ai.meta.com/blog/llama-2/>, accessed 22 July 2023.

Lyon D, *Vesikalı Yurttaş* Barış Baysal (çev) (Kalkedon Yayınları 2012)

Maume P, 'Regulating Robo – Advisory', 55(1) Texas International Law Journal, 51; Dominique Payette, 'Regulating Robo-Advisers in Canada', 33(3) Banking & Finance Law Review.

McCarthy J, 'Programs With Common Sense' in D. Blake and A. Uttelly (eds) Proceedings of the Symposium on the Mechanization of through Processes (H.M. Stationery Office,1959)

Menabrea LF, 'Sketch of the Analytical Engine Invented by Charles Babbage, Esq.' in Richard Taylor (ed), *The Transactions of Foreign Academies of Science and Learned Societies* (Taylor and Francis 1843).

Microsoft, 'Microsoft and OpenAI Extend Partnership' (Microsoft 23 Jan 2023) <https://blogs.microsoft.com/blog/2023/01/23/microsoftandopenaiextendpartnership/> accessed 22 Jul 2023.

Metz C and Weise K, 'Microsoft to Invest $10 Billion in OpenAI, the Creator of ChatGPT' <https://www.nytimes.com/2023/01/23/business/microsoft-chatgpt-artificial-intelligence.html>, accessed 22 Jul 2023.

Mittelstadt B, 'From Individual to Group Privacy in Big Data Analytics' (2017) 30 Philos. Technol.

Murgia M, 'Who's using your face? The ugly truth about facial recognition' (Financial Times, 19.4.2019) <https://www.ft.com/content/cf19b956-60a2-11e9-b285-3acd5d43599e> accessed 22 July 2023

Norwegian Consumer Council, 'Ghost in the Machine: Addressing the Consumer Harms of Generative AI' (Norwegian Consumer Council, June 2023), www.forbrukerradet.no/ai, accessed 26 June 2023.

OpenAI, 'DALL-E 2 Extending Creativity' (OpenAI 14 July 2023) <https://openai.com/blog/dall-e-2-extending-creativity> accessed 22 July 2023.

OpenAI, 'API Data Usage Policies', 14 Haziran 2023, <https://openai.com/policies/api-data-usage-policies> accessed 5 Jul 2023.

OpenAI, 'New Ways to Manage Your Data in ChatGPT', 25 Ap 2023, <https://openai.com/blog/new-ways-to-manage-your-data-in-chatgpt> accessed 22 Jul 2023

OpenAI, 'ChatGPT4 Technical Report' (Arxiv 27 Mar 2023) <https://doi.org/10.48550/arXiv.2303.08774> 10, accessed 22 Jul 2023.

OpenAI, 'Introducing ChatGPT' (OpenAI 30 Nov 2022) <https://openai.com/blog/chatgpt> accessed 22 Jul 2023.

Ozan Özparlak B, *Büyük Veri Çağında Çalışma İlişkilerinde Yapay Zekâ Sistemlerinin Kullanılması: Hukuki Bir Değerlendirme* (On İki Levha Yayınları 2021).

Ozan Özparlak B, 'Yeni Çağın Hukukunu Teknoloji ve Tasarım Şekillendirecek' HBT Dergi Herkese Bilim Teknoloji, 12.1.2018.

Ringe WG ve Ruof C, 'A Regulatory Sandbox for Robo Advice, European Banking Institute' (2018)

Sankur B, *İngilizce Türkçe Ansiklopedik Bilişim Sözlüğü* (Pusula Yayıncılık, 2008).

Silver D, Huang A and Maddison CJ, et al, 'Mastering the Game of Go with Deep Neural Networks and Tree Search' (2016) 529 Nature 484–489 <https://doi.org/10.1038/nature16961> accessed 20 July 2023.

Smits J and Borghuis T, (2022) 'Generative AI and Intellectual Property Rights' in Bart Custers, Eduard Fosch-Villaronga (eds) *Law and Artificial Intelligence*, *Information Technology and Law Series,* (T.M.C. Asser Press 2022).

Streeck W, *Satın Alınan Zaman* Kerem Kabadayı (çev) (Koç Üniversitesi Yayınları 2016)

Taylor L, Floridi L and Sloot B, 'Introduction: A New Perspective on Privacy', Linnet Taylor, Luciano Floridi and Bart van der Sloot (eds) *Group Privacy New Challenges of Data Technologies*: (Springer, 2017)

T.C. Cumhurbaşkanlığı Dijital Dönüşüm Ofisi, 'Chatbot Uygulamaları ve ChatGPT Örneği' (15 March 2023)

TC Cumhurbaşkanlığı Dijital Dönüşüm Ofisi 'Türkiye Cumhuriyeti Ulusal Yapay Zekâ Stratejisi' (Cumhurbaşkanlığı Dijital Dönüşüm Ofisi, Ağustos 2021) >https://cbddo.gov.tr/SharedFolderServer/Genel/File/TR-UlusalYZStratejisi2021-2025.pdf> accessed 22 Jul 2023.

The Center for Research on Foundation Models at the Stanford Institute for Human-Centered Artificial Intelligence, 'On the Opportunities and Risks of Foundation Models' (Arxiv, 16 August 2021) <https://doi.org/10.48550/arXiv.2108.07258>, accessed 22 July 2023.

Touvron H, Martin L; 'Llama 2: Open Foundation and Fine-Tuned Chat Models', (Meta AI, 18 July 2023) >https://ai.meta.com/research/publications/llama-2-open-foundation-and-fine-tuned-chat-models/> accessed 22 July 2023.

Turing AM, 'Computing Machinery and Intelligence' (1950) Mind, Vol. LIX, Sayı 236, 433–460, https://doi.org/10.1093/mind/LIX.236.433 accessed 22.7.2023

UNESCO, 'Ethics of Neurotechnology' (Unesco) <https://www.unesco.org/en/ethics-neurotech> accessed 22 Jul 2023.

United Nations, 'Meet the Robots Who Are Making The World a Better Place' (UN News 06 Jul 2023) <https://news.un.org/en/story/2023/07/1138412> accessed 22 Jul 2023

Üstündağ Soykan E, Bilgin Z, Ersoy MA and Tomur E, 'Differentially Private Deep Learning for Load Forecasting on Smart Grid' (2019) IEEE Globecom Workshops, <10.1109/GCWkshps45667.2019.902.4520> accessed 22 Jul 2023.

Van der Sloot B, 'The Individual in the Big Data Era: Moving Towards an Agent-Based Paradigm' in Bart van der Sloot, Dennis Broeders, Erik Schrijvers (eds) *Exploring the Boundaries of Big Data* (Amsterdam University Press, 2016) 196-199.

Van der Sloot B and Wagensveld Y, 'Deep Fakes: Regulatory Challenges for the Synthetic Society' (2022) 46 (1) Computer Law & Security Review.

Veale M, Binns R and Edwards L, 'Algorithms that Remember: Model Inversion Attackand (2018) 376 Data Protection Law' Phil. Trans. R. Soc. A

Wachter S, '*Data Protection in the Age of Big Data*' (2019) 2 Nature Electronics Volume

Wachter S, 'Affinity Profiling and Discrimination by Association in Online Behavioural Advertising' (2020) 35(2) Berkeley Technology Law Journal

Weizenbaum J, 'ELIZA-A Computer Program For the Study of Natural Language Communication Between Man and Machine' (1966) 9(1) Communications of the ACM.

Werra J, 'AI Transparency: An Emerging Principle in The IP Ecosystem?' (2023) Journal of Intellectual Property Law& Practice, <https://doi.org/10.1093/jiplp/jpad041

Woebot Health, (Woebolt Health 2023) <https://woebothealth.com/> accessed 27 July 2023; Attia Qammar et al., 'Chatbots to ChatGPT in a Cybersecurity Space: Evolution, Vulnerabilities, Attacks, Challenges, and Future Recommendations' (2021) 14(8) Journal of Latex Class Files.