

Yapay Zekâ ve Eğitim İlişkisine Dair Felsefi Bir Değerlendirme

Gülizar Karahan Balya¹

Bu makale, yapay zekâ teknolojilerinin eğitimde kullanımını felsefi ve etik açıdan incelemektedir. Bu doğrultuda, öncelikle yapay zekânın eğitim alanındaki güncel kullanım biçimleri, sunduğu olanaklar ve beraberinde getirdiği riskler ele alınmakta, ardından teknoloji felsefesinde üç temel yaklaşım olan araçsalcılık, teknolojik belirlenimcilik ve ampirik dönüş çerçevesinde yapay zekânın eğitimdeki rolü ile etik boyutları tartışılmaktadır. Yapay zekâ araçsalci yaklaşımda değerini kullanım amaçlarından alan nötr bir araç olarak, belirlenimci yaklaşımda insan yaşamını dönüştüren ve üzerinde sınırlı etkimizin bulunduğu özerk bir güç olarak görülür. Ampirik dönüş sonrasında ise düşünürler teknolojinin bireyi ve toplumu şekillendiren etkisini kabul etmekle birlikte, teknolojik gelişmelerin toplumsal bağlamlardan bağımsız olmadığını ve insanların da teknolojilerin tasarımını, kullanımını ve sonuçlarını şekillendirme kapasitesine sahip olduğunu savunurlar. Makale, bu üç yaklaşımın yapay zekânın eğitimdeki etik sorunları değerlendirme bakımından sunduğu olanak ve sınırlılıkları ortaya koymakta, yapay zekânın eğitsel anlam ve etkilerinin yalnızca teknik özellikleriyle veya bireysel kullanım tercihleriyle açıklanamayacağını, bunların içinde şekillendiği tarihsel, toplumsal ve politik bağlamlarla birlikte değerlendirilmesi gerektiğini savunmaktadır.

Eğitimde Yapay Zekâ ve Olanaklar

Eğitimde yapay zekâdan söz edildiğinde akla ilk gelen büyük dil modellerine dayalı üretken sistemler olsa da, yapay zekâ uygulamaları hem öğretme ve öğrenme süreçlerini desteklemekte hem de ölçme ve değerlendirme aşamalarında yıllardır kullanılmaktadır. Özellikle veri analitiğine ve kişiselleştirme mantığına dayalı çok sayıda ürün eğitim teknolojileri piyasasında

¹ kgulizar@metu.edu.tr

giderek daha fazla yer bulmaktadır. Örneğin, Türkiye’de de uygulamalarına rastladığımız kişiselleştirilmiş öğrenme sistemleri, öğrenci performansını belirlemek amacıyla sınav analizi yaparak desteğe ihtiyaç duyulan konu alanlarını tespit edip öğrenciye özgü soru havuzları oluşturabiliyor. Benzer biçimde, otomatik değerlendirme araçları yoluyla, açık uçlu sorulardan oluşan ya da araştırma içeren sosyal/beşeri bilim ödevleri ve uzun kodlamalar veya problem çözümleri gerektiren temel bilim/mühendislik çalışmaları hızlı bir şekilde değerlendirilerek çok sayıda öğrenciye daha kısa sürede geri bildirim verilebiliyor. Diğer yandan, mevcut ölçme mekanizmalarına yapay zekâ sistemlerinin entegre edilmeye başlandığını da görüyoruz. Türkiye’de Millî Eğitim Bakanlığı’nın 6-22 yaş aralığındaki bireylerin zihinsel beceri ve yeteneklerini belirlemeye yönelik yapay zekâ tabanlı, interaktif bir psikolojik ölçme aracı niteliğindeki Türk Ulusal Zekâ Ölçeği’ni (TUZÖ) kullanıma sunması gibi.² Son olarak, görme, işitme veya öğrenme güçlüğü gibi engel durumları bulunan ya da özel eğitim gereksinimi olan bireyler için geliştirilmiş dijital araçlar (metinleri seslendiren ya da anlık altyazı çevirisi yapan araçlar gibi) yoluyla eğitimde erişilebilirlik ve kapsayıcılık yönünde adımların atıldığından söz edebiliriz.

Yapay zekâ sistemlerinin eğitim alanındaki en büyük vaatleri arasında öğretmenin yükünü hafifletme, öğrenciye geri bildirim verme hızının artması, öğrenmenin kişiselleştirilmesi ve okuldan ayrılma ya da sınıf tekrarı gibi risklere karşı erken uyarı vererek önleyici tedbirlerin alınması yer almakta. Ancak bir sonraki bölümde ele alınacağı üzere, tüm bu olanakların yanında pek çok risk de dillendirilmektedir. Eğitimde yapay zekâyı ilişkin etik değerlendirmelerin daha çok bu olanak ve riskleri (başka bir ifadeyle, tehdit ve fırsatları) tartmayı salık veren yaklaşımlar ve değerlere dayalı çerçeveler üzerinden yapıldığını görmekteyiz. Tehditleri ve fırsatları tartışıp öngörmeye çalışmak ve evrensellik iddiasına sahip ilke ve değerleri belirleyerek yeni gelişen teknolojilerin etik boyutunu bunlar etrafında değerlendirmek elbette önemli. Ancak yapay zekâ etiğine tehdit-fırsat ve ilke-değer çerçeveleri üzerinden yaklaşmanın bazı kısıtlılıkları da var. Makalenin ilerleyen bölümlerinde öncelikle bu çerçevelere kısaca değinmek, ardından eğitimdeki belli başlı riskleri etik ilke ve değerler üzerinden değerlendirmek ve sonraki bölümde alternatif bir yaklaşım olarak eğitimde yapay zekâyı teknoloji felsefesi ekseninde konumlandırarak farklı bir etik yaklaşıma dikkat çekmek istiyoruz.

Riskler, Değerler ve İlkeler

UNESCO ve AB Komisyonu gibi uluslararası kuruluşların ve ülkemizde MEB, YÖK, TÜBİTAK gibi eğitim, yükseköğretim ve araştırmaya yön veren politika yapıcı kurumların yapay zekânın sorumlu kullanımına ilişkin yayımladığı dokümanlara bakıldığında, etik konusunda tehdit ve fırsatlar ile ilke ve değerlere dayalı tavsiyelere yönelik eğilim açıkça görülmektedir. Örneğin, UNESCO’nun 2021’de

² <https://meb.gov.tr/2025-yilinda-egitimde-bircok-yeni-proje-ve-uygulama-hayata-gecirildi/haber/39424/tr>

güncel versiyonunu yayımladığı “*Recommendation on the Ethics of Artificial Intelligence*” (Yapay Zekâ Etiğine İlişkin Tavsiye Dokümanı) başlıklı kılavuzunda insan hakları, onuru ve temel özgürlükler odağında saygı, çeşitlilik ve kapsayıcılık gibi bir dizi değere vurgu yapılmakta. Ayrıca orantısallık, zarar vermeme, ayrımcılık yapmama, hakkaniyet, güvenlik, mahremiyet, sürdürülebilirlik, insan gözetimi, şeffaflık, açıklanabilirlik, sorumluluk, hesap verebilirlik, yapay zekâ okuryazarlığı ve işbirliği ilkelerinin önemine dikkat çekilmektedir. Türkiye’de de YÖK tarafından 2024’te yayımlanan “Yükseköğretim Kurumları Bilimsel Araştırma ve Yayın Faaliyetlerinde Üretken Yapay Zekâ Kullanımına Dair Etik Rehber”³ üretken yapay zekâ sistemlerinin yüksek öğretimde kullanımına ilişkin, etik değerlere dayalı bir dizi tavsiye içermekte. MEB tarafından yayımlanan “Millî Eğitim Bakanlığı Yapay Zekâ Uygulamaları Etik Kurulu Yönergesi” (2025) ile bunun devamı niteliğindeki “Eğitimde Yapay Zekâ Uygulamaları Etik Kılavuzu” (2025) MEB bünyesinde yürütülen yapay zekâ uygulamalarına yönelik başlıca etik ilkeleri belirleyerek, bunların uygulanması, izlenmesi ve değerlendirilmesine ilişkin usul ve esasları açıklamaktadır.⁴ Uluslararası ve ulusal pek çok kılavuz birbirine benzer değer ve ilkelere referansla yazıldıysa da, farklı ülkelerin dokümanlarında uyum (harmoni) veya özen gibi, ortaklaşmayan kültürel değerlere de rastlanmaktadır.⁵

Etik ilke ve değerlere dayalı dokümanlara yöneltilen en büyük eleştiri, bunların yaptırım gücüne sahip olmadığı yönündedir. Ayrıca, bu tür üst düzey dokümanlarda listelenen soyut değerlerin uygulamada ne tür normlara ve düzenlemelere dönüşebileceğinin belirsiz oluşu da bir başka eleştiri konusudur. Örneğin, yapay zekâ kullanımında şeffaflık ilkesinin benimsenmesini tavsiye eden bir kılavuz, bunun uygulamada hangi somut adımlarla mümkün olabileceğine dair örnekler sunmadığı için yeterince yol gösterici olmaması nedeniyle eğitim sistemi düzenlemelerini doğrudan etkileyemeyebilir. Yaptırım gücü ve yol göstericilik meseleleri, etiğe ilişkin temel birer eleştiri noktası olsa da, etik ilke ve değerlerin hukuk düzenlemelerine temel oluşturduğu ve bunların önünü açtığı da bir gerçektir. Bunun yakın bir örneğini, 2024 yılında yürürlüğe giren Avrupa Birliği

³ <https://proje.yok.gov.tr/tr/page/635>

⁴ Yönerge: <https://ttkb.meb.gov.tr/www/quotyapay-zeka-uygulamalari-etik-kurulu-yonergesiquot-yayimlandi/icerik/818>

Kılavuz: <https://ttkb.meb.gov.tr/www/quotegitimde-yapay-zeka-uygulamalari-etik-kilavuzuquot-yayimlandi/icerik/854>

⁵ “Uyum,” aralarında Çin üniversiteleri, araştırma kurumları ve teknoloji şirketlerinin bulunduğu on beş kuruluş tarafından 2019’da yayımlanan Pekin Yapay Zekâ İlkeleri’nde yapay zekâ yönetiminde temel ilkelerden biri olarak yer almaktadır: <https://ai-ethics-and-governance.institute/beijing-artificial-intelligence-principles/> “Özen” ise Yükseköğretim Kurulu tarafından 2024’te yayımlanan Yapay Zekâ Kullanımına Dair Etik Rehber’de “şeffaflık ve dürüstlüğü de içeren bir yükümlülük” olarak geçmektedir: <https://proje.yok.gov.tr/tr/page/635>

Bu makale, kültürel değerlerin farklılaşabileceğine dikkat çekmekle birlikte, söz konusu farklılıklara ilişkin daha kapsamlı çalışmaların da alanyazına önemli bir katkı olacağını değerlendirmektedir.

Yapay Zekâ Tüzüğü'nde⁶ görmek mümkün. Tüzük insan gözetimi, güvenlik, gizlilik, şeffaflık, açıklanabilirlik, çeşitlilik, ayrımcılık karşıtlığı, adillik, toplumsal ve çevresel refah gibi değerlere dayanmakta ve bu değerlerden hareketle yapay zekâ uygulamalarını dört temel risk grubuna ayırarak (kabul edilemez, yüksek riskli, sınırlı riskli ve düşük riskli sistemler) her grup için ayrı düzenlemeler öngörmektedir. Örneğin, öğrencilerin psikolojik durumunu, odaklanmasını veya duygularını ölçmeye yönelik sistemler temel hakları ihlal edebileceği ve ayrımcılığa sebep olabileceği gerekçesiyle kabul edilemez risk grubunda değerlendirilerek yasaklanmaktadır (Madde 5-(1), f, g). Somutlaştırmak gerekirse, sınıflarda veya çevrimiçi sınavlarda kameralar aracılığıyla öğrencilerin yüz ifadelerinden, ses tonlarından veya diğer biyometrik verilerinden hareketle sıkılıp sıkılmadıklarını, kopya çekmeye *niyetlenip niyetlenmediklerini* veya duygusal durumlarını (öfke, mutluluk vb.) tespit etmeye çalışan duygu tanıma sistemleri, geliştirilmesi ve kullanılması yasaklanan uygulamalar arasındadır. Öte yandan, bireylerin eğitim hayatını, mesleki geleceğini ve sosyoekonomik durumunu doğrudan etkileyebilecek sistemler ise yüksek riskli kabul edilmektedir (Gerekçe Maddesi 56). Bunun eğitime bir yansıması olarak, yükseköğretim veya mesleki eğitim kurumlarına başvuran adayların dosyalarını veya sınav sonuçlarını inceleyerek kimin kabul edileceğini ya da adayların hangi bölüme yerleştirileceğini belirleyen filtreleme ve değerlendirme sistemlerinin piyasaya sürülmeden önce ve kullanım süresince risk yönetimi, veri kalitesi standartları, şeffaflık ve sürekli insan gözetimi gibi yükümlülüklerle tâbi olmaları şart koşulmaktadır.

Eğitimde yapay zekânın oluşturduğu diğer belli başlı riskler şöyle sıralanabilir:

1. Algoritmik önyargı / yanlılık

Önyargı ya da yanlılık olarak anılan problem, yapay zekâ teknolojilerinin yaygın bilinen yapısal sorunlarından biridir. Güncel yapay zekâ sistemleri veri kümeleri arasındaki istatistiksel ilişkileri öğrenerek çalışmaktadır. Burada “zekâ” olarak adlandırılan şey, veri içindeki örüntüleri istatistiksel olarak yakalama (“öğrenme”) ve buna dayanarak çıkarım yapma, tahminde ya da öngöründe bulunma becerisidir. Dolayısıyla, bu sistemlerin eğitiminde kullanılan veri setlerinin yanlılıktan uzak olması büyük önem taşımakta, ancak veri setleri çoğu zaman farklı demografik gruplar açısından eşitsizlikler barındırmaktadır. Örneğin, bir ülkenin sağlık sisteminden derlenen veri setlerinde çoğunlukla orta-üst sınıfa mensup, erkek bireylerin temsil ediliyor oluşu, diğer toplumsal sınıf ve cinsiyet kategorilerine mensup bireylerin sağlık durumlarına ilişkin tahminlerde sistematik hataların ortaya çıkmasına ve bu bireyler için dezavantaja yol açabilir.

Eğitim alanından somut bir örnek vermek için maddi olanakları ve dijital altyapısı gelişmiş özel okulların yoğunlukta olduğu bölgelerden toplanan verilerle eğitilmiş bir yapay zekâ aracını ele alalım. Bu veriler, ister istemez belirli bir sosyoekonomik

⁶ <https://artificialintelligenceact.eu/ai-act-explorer/>

düzeydeki semtlerden gelen veli ve öğrenci profillerini yansıtacaktır. Algoritma, böyle avantajlı okullardan gelen verilerdeki okul dışı aktiviteler ve yabancı dil sertifikaları gibi değişkenlerle okul başarısı arasında bir korelasyon tespit edebilir ve aslında sosyoekonomik imkânlarla bağlı olan bu değişkenleri “başarı” kriteri olarak öğrenebilir. Günümüz yapay zekâ sistemleri bu yanlı veri setleriyle eğitildiğinde, sosyoekonomik açıdan dezavantajlı bir bölgede ikamet eden bir öğrenciyi “başarı” göstergelerine sahip olmadığı için “düşük potansiyelli” olarak etiketleyebilir. Böyle bir sistemin yaratacağı tek sorun var olan toplumsal eşitsizlikleri yansıtmaması değil, aynı zamanda “algoritmik meşruiyet” denebilecek bir zeminde eşitsizliği yeniden üretmesidir. Diğer bir deyişle, yanlı eğitim verilerine dayanan yapay zekâ sistemlerinin getirdiği faydalar ve riskler toplumda eşit dağılmayıp, hâlihazırda dezavantaja uğramış olan gruplar daha ağır darbeler alabilir (Coeckelberg, 2022).

2. Veri sömürüsü ve gözetimin olağanlaşması

Yapay zekâ sistemlerinin eğitimde kullanımı, çoğu zaman öğrencilere ait çeşitli kişisel verilerin toplanmasını ve işlenmesini gerektirmektedir. Türkiye’de okul bilgileri, ölçme değerlendirme sonuçları, ses ve görüntü kayıtları, 6698 sayılı Kişisel Verilerin Korunması Kanunu (KVKK) kapsamında kişisel veri niteliği taşır. Zira bu veriler doğrudan belirli bir öğrenciyle ilişkilendirilebildiği gibi, kimlik bilgileriyle birlikte tutulmasalar dahi, başka verilerle eşleştirildiğinde ilgili öğrencinin belirlenmesini mümkün kılabilir. Ayrıca, eğitim verilerinin kurumlar tarafından yasaya uygun şekilde toplanmış olması, onlara bu verileri diledikleri biçimde işleme veya ticarileştirme hakkı tanımaz. Oysa öğrencilere ait birçok kişisel verinin kurumlar tarafından toplanarak eğitim teknolojisi firmalarına verildiği, satıldığı ya da bu firmaların geliştirdiği ürünler aracılığıyla bizzat toplandığı ortaya konulmaktadır. Kişiselleştirilmiş veya kolaylaştırılmış öğrenme vaatleriyle yürütülen bu süreçler, henüz reşit olmayan öğrencileri aslında farkında olmadan ticari veri kaynağına dönüştürmektedir. İnsan Hakları İzleme Örgütü, pandemi döneminde aralarında Türkiye’nin de yer aldığı 49 ülke tarafından onaylanmış olan 163 eğitim teknolojisi platformunu incelemiştir.⁷ Bu çalışma, platformların %89’unun çocuk hakları ve mahremiyetini doğrudan ihlal eden ya da riske atan veri uygulamaları yürüttüğünü ortaya koymuştur. Derlenen rapor, platformların çoğunlukla çocukların ve ebeveynlerin rızası olmadan öğrencileri izleyebildiğini, çocukların kimlikleri, konumları, sınıf içi aktiviteleri, aile ve arkadaş çevreleri gibi pek çok kişisel veriyi eğitim dışı amaçlarla kullandığını göstermektedir. Yapay zekâ teknolojileri bu tür mahremiyet ihlallerini ve gözetim uygulamalarını daha da ileri taşıyarak çok daha fazla veri toplanması, öğrencilerin profillenmesi, davranış ve performanslarını ölçerek öngörme pratiklerini artırma potansiyeline sahiptir.

⁷<https://www.hrw.org/report/2022/05/25/how-dare-they-peep-my-private-life/childrens-rights-violations-governments>

Veri toplamada sınırların aşılmasının yaratabileceği bir diğer risk de eğitim ortamlarında gözetim uygulamalarının normalleşmesidir. Güvenlik, yoklama, derse katılım gibi gerekçelerle kullanılan yüz tanıma ve duygu tespiti sistemleri ile öğrenci profillerine göre oturma düzeni oluşturma gibi uygulamalar, okullarda gözetimi sıradanlaştırarak öğrenmeyi olumsuz yönde etkileyebilir. Buradan hareketle yapay zekâyı, öğrencileri performansları ölçülebilir ve davranışları öngörülebilir bireylere, öğretmenleri de bu ölçüm ve tahminleri kullanarak eğitimi yönlendirmeye çalışan teknisyenlere dönüştürecek yeni bir veri ve gözetim altyapısı olarak değerlendirebiliriz. Ancak sürekli izlendikleri, etiketlendikleri, profillerinin çıkarıldığı ve aralarındaki organik bağın giderek zayıfladığı bir ortamda öğrenciler ve öğretmenler özgün davranışlarını sergilemekte kendilerini ne derece özgür hissedebilirler? Kendilerine sansür uygulamalarının öğrenme ilişkisine ve öğrenci-öğretmen-veli iletişimine etkileri neler olur?

Yapay zekâ desteğiyle sınıf düzeni oluşturmak amacıyla kullanılan bir eğitim teknolojisi örneği.

Kaynak: <https://aklisira.com/app>

3. Öğretmenin pedagojik özerkliği

İnsan gözetimi, yapay zekânın kullanıldığı kritik alanlarda kesinlikle gözetilmesi gerektiği vurgulanan ilkelerin başında gelir. Özellikle İnsan Merkezli Yapay Zekâ (Human-Centered AI) yaklaşımında altı çizilen bu ilke, teknolojinin insanların yerini alması yerine onların yeteneklerini artırmayı, güvenliğini korumayı ve kontrolün daima insanda kalmasını öngörmektedir. Söz konusu alan eğitim ve öğretim olduğunda, öncelikle öğrenci ve öğretmenin, daha sonra da toplumun genel refahının gözetilmesi adına, teknolojinin varlığında öğretmenin özerkliğini

koruyabilmesi önemlidir. Ancak ders planlama, eğitim materyali oluşturma, sınav hazırlama, ölçme-değerlendirme, öğrenci/veli raporları hazırlama gibi konularda öğretmenlerin üstlendikleri görevleri ve karar verme yetkilerini yapay zekâya giderek daha çok devretmesiyle, uzun vadede öğretmenliğin tanımının değişmesi söz konusu olabilir. Kendi mesleki yargısını kullanma pratiğini kaybederek vasıfsızlaştırılan öğretmen, teknolojikleşmiş bir eğitim sisteminde salt bir kullanıcı ya da teknisyen pozisyonuna sürüklenebilir. İnsan faktörü pedagojik ilişkiden silindikçe, öğretmenlik mesleğinin ve öğrenmenin özünde yatan bağ kurmaya ve özene dayalı emek önemsiz ve görünmez hale gelebilir. Bu da hâlihazırda itibarsızlaştırılan öğretmenlik mesleğinin itibarı ve etkisinin daha da azalması, hatta yapay zekâ tartışmalarında sıkça rastlanan bir mesele olan mesleğin ortadan kalkması riskini beraberinde getirebilir.

4. Yapay zekâya aşırı yaslanma

Yapay zekânın eğitimde yaygınlaşmasının önemli risklerinden bir diğeri, öğrencilerin öğrenmede etkili olan bilişsel emeği yapay zekâ sistemlerine devrederek zaman içinde bu ilişkiye bağımlı hale gelmeleridir. Yapay zekâya aşırı yaslanma (*overreliance*) olarak tanımlanabilecek olan bu durum, tıpkı öğretmenin özerkliğini tehdit ettiği gibi, öğrencinin de hem bilişsel gelişimi hem de eleştirel öğrenme serüveni açısından önemli bir risk oluşturur.

Eleştirel pedagoji gibi farklı perspektiflerden bakıldığında, öğrenme yalnızca bir öğretmenin öğrenciye bilgi aktarması olarak tanımlanamaz. Öğrencinin herhangi bir bilgiyle karşılaştığında onu sorgulaması, yeni bağlantılar kurması ve bilgiyi edineceği ve anlamlandıracağı biçime kendi muhakemesi sonucunda ulaşması öğrenmenin temel unsurları olarak anlaşılır. Yapay zekâ araçlarının hazır yanıtlar, kısa özetler ve hatta “yorum” sunma işlevleri, öğrenme sürecinde aktif olarak verilmesi beklenen bilişsel emeği öğrencinin gözünde “gereksiz” kılarak, bazı durumlarda pratik sonuçlar (ödevlerin kısa sürede teslim edilmesi, yüksek sınav notu, vb.) sunabilmektedir. Bu şekilde, örneğin yazma eylemi öğrencinin kendi bilişsel dünyasında kaybolup yolunu kendi bulduğu, özgün düşüncelerini kurup geliştirdiği bir süreç olmaktan çıkarak, yapay zekânın verdiği çıktılar arasından seçim ve düzenleme yaptığı, dolayısıyla düşünmeyi öğrenemediği bir biçime bürünebilir.

Öğrencinin böyle hazır üretilmiş çıktılarla çalışması, uzun vadede bağımsız ve eleştirel düşünme, argüman geliştirme ve hatırlama gibi becerilerin gelişimini ise olumsuz yönde etkileyebilir. MIT Media Lab’dan araştırmacıların büyük dil modellerinin (LLM) beyin üzerindeki etkilerini incelediği 2025 tarihli çalışma, LLM'lere bağımlılığın eğitim üzerindeki etkilerine ilişkin kaygıları gündeme getirmekte ve yapay zekânın öğrenme süreçlerindeki rolünün daha derinlemesine araştırılması gerektiğini göstermektedir.⁸ Çalışmada üç farklı koşulda kompozisyon yazma görevi verilen, 18-25 yaş aralığında 54 öğrencinin beyin

⁸ <https://www.media.mit.edu/projects/your-brain-on-chatgpt/overview/>

aktiviteleri incelenmiştir. İlk grup kompozisyonlarını ChatGPT kullanarak, ikinci grup internet üzerinden arama yaparak, üçüncü grup ise herhangi bir araç kullanmadan yazmıştır. Dört aylık süre boyunca LLM kullanan katılımcılar nöral, dilsel ve davranışsal düzeylerde diğer gruplardan daha düşük performans göstermiş; örneğin, yazılan metnin sahibi hissetme ve metinden alıntı yapma becerisi bakımından en düşük skoru elde etmiştir. Herhangi bir araç kullanmadan yazan grup ise hatırlama fonksiyonu ve yazılan metnin sahibi hissetme bakımından en başarılı grup olmuştur. Bu çalışmanın da işaret ettiği üzere, öğrenme için gerekli olan bilişsel emeğin yerine yapay zekâ geçtiği ölçüde, kritik önemdeki becerilerin ve eğitimin temel amaçlarından olan bağımsız muhakeme, eleştirel düşünme gibi kapasitelerin sekteye uğraması önemli birer risk olarak gözükmektedir.

5. Epistemik güven

Yapay zekâ sistemleri bilgiyle kurduğumuz ilişkiye yeni bir boyut kazandırmaktadır. Örneğin, büyük dil modellerini ele alırsak, bunların bilgiyi seçip filtrelemenin yanı sıra korelasyona dayalı sözcük dizilimleri yoluyla bilgi “üreten” ve “yorum yapan” sistemler olduğunu söylemek yanlış olmaz. Bu demek oluyor ki, yapay zekâ çağında insan bilgiyi dolaylı biçimde, yani bir ‘epistemik aracı’ (*epistemic intermediary*; Neuberger vd., 2025) yoluyla edinme pratiğine sahip. Ancak bu aracının işleyişi tamamen açıklanabilir ve şeffaf değildir. “Kara kutu problemi” olarak adlandırılan bu durum, yapay zekâ çıktılarının hangi hesaplamalar ve ilişkiler sonucunda ortaya çıktığının kullanıcılar hatta bazen geliştiriciler açısından bile tam olarak anlaşılamamasını ifade etmektedir. Bu koşullarda, epistemik aracının güvenilirliği ve meşruiyeti değerlendirilip karar verilecek bir mesele haline gelir ve bilginin kendisi kadar önemli olur (Balázs, 2026).

Bu da demek oluyor ki, yapay zekânın yanlış bilgi üretmesinin yanında çıktılarının hangi gerekçelerle güvenilir kabul edilebileceğine dair belirsizlik de başlı başına bir sorun. Diğer bir ifadeyle, kullanıcıların yapay zekâ tarafından üretilen çıktılarına hangi gerekçelerle güvenebilecekleri, bu güvenin hangi koşullarda haklı bir zemini olduğu ve ne zaman sorgulanması gerektiği soruları önemli birer gündem maddesi. Özellikle eleştirel değerlendirme becerileri henüz yeterli düzeyde gelişmemiş olan öğrenciler açısından, yapay zekânın sunduğu hazır yanıtların doğrulanmaya ihtiyaç duymadan kabul edilmesi büyük bir risk. Benzer şekilde, yapay zekâ destekli karar verme araçlarının sunduğu değerlendirmeler de öğretmenleri ve yöneticileri, bu değerlendirmelerin dayandığı ölçütleri ve üretim süreçlerini, yani kara kutuyu, açmaya çalışmadan kabul etmeye yöneltebilir.

Bu sorunu somutlaştırmak için, bir yapay zekâ sisteminin öğrencilerin moral durumları ile dikkat süreleri arasında bir ilişki tespit ettiğini düşünelim. Böyle bir ilişkinin tespit edilmiş olması, aslında tek başına nedensel bir açıklama sunmaz; moral durumuyla dikkat süresinin birbirini etkilediğini veya bunların başka bir değişkenle ilişkili olduğunu göstermez. Buna rağmen yapay zekâ bu ilişkiyi kesin

bir açıklama gibi sunduğunda, kullanıcılar çıktının ikna ediciliğini onun doğruluğuyla karıştırabilirler. Üstelik aynı konuda öğretmenin gözlemleriyle yapay zekânın değerlendirmesi farklılaştığında, hangi değerlendirmenin neden daha güvenilir kabul edileceği sorusu da ortaya çıkar. Bu nedenle epistemik güven, genel olarak yapay zekâyâ güvenip güvenmemekle değil de, onun sunduğu bilginin dayanaklarını değerlendirebilmekle ilgilidir. Bu değerlendirme yapılmadığı zaman, sorgulama ve doğrulama süreçleri geri plana itilir ve yapay zekânın sağladığı bilişsel kolaylık, doğru bilgi yolunda verilmesi gereken zihinsel emeğin yerini kolayca alır. Böyle bir durumda da eğitim, bilgiye ulaşırken gerekçelendirme ve eleştirel değerlendirme becerilerini geliştiren bir ortam olmaktan uzaklaşır.

Eğitimde Yapay Zekâyı Teknoloji Felsefesi Ekseninde Konumlandırmak

Teknoloji felsefesi, felsefenin diğer alanlarına göre görece geç adı konulmuş ve hâlen gelişmekte olan bir alandır. Antik Yunan ve Aydınlanma dönemlerinde teknoloji üzerine kafa yoran düşünürler bulunsa da günümüzde “klasik” teknoloji felsefecileri olarak anılan isimler Sanayi Devrimi ve sonrasında modern teknolojinin beraberinde getirdiği toplumsal dönüşümü anlamlandırma çabası içinde olmuşlardır. Avrupa Aydınlanması’nın rasyonalite anlayışına ve 19. yüzyılın başat düşüncesi olan endüstriyel ilerleme idealine yönelik eleştirel bir tepki olarak şekillenen klasik teknoloji felsefesi, Aydınlanma’nın iyimser tavrından uzaklaşarak teknolojinin insan varoluşunu olumsuz ve geri döndürülemez biçimde etkilediğine dikkat çekmekteydi. Araçsalcılık (*instrumentalism*) olarak adlandırılan iyimser görüşe göre, teknoloji insan aklının ürettiği ve yine aynı akılsal kapasiteyle denetlenip kontrol edilebilecek, değer yargılarından bağımsız bir olgudur. Diğer bir deyişle, teknoloji nötrdür; ona atfedilecek olumlu ya da olumsuz nitelikler, teknolojiyi kullanan öznenin amaç ve eylemlerine bağlıdır.

Teknolojiye araçsal yaklaşımı reddeden klasik düşünürlerse, daha karamsar bir yaklaşım benimseyerek modern teknolojiyi bireyleri, toplumu, değerleri, çevreyi, ekonomiyi, kısacası dünyayla ilişkimizi bütünüyle şekillendiren dönüştürücü bir güç olarak görmüşlerdir. Örneğin, İkinci Dünya Savaşı öncesi Almanya’sını ve Avrupa’sını izleyen Martin Heidegger (1977), teknolojinin özünün teknolojik olmadığını ileri sürmüştü. Bununla ifade etmek istediği, teknolojinin insanın üretmiş olduğu eski ve yeni aletlerin, barajların, demiryollarının, elektrikli araçların vb. toplamı olarak anlaşılamayacağı, aslında bu tek tek nesnelerin ötesinde, insan yaşamına nüfuz etmiş özerk bir güç olduğu idi. Heidegger’e göre modern teknoloji, insanın dünyayı kullanıma hazır ve emrine amade kaynaklar, hammadde olarak gördüğü ve kendini içinde bulduğu bir düşünme biçimidir; bu nedenle insanın doğaya ve genel olarak varoluşa bakış açısında büyük bir değişime yol açmıştır. Modern teknolojik insan, öteki varlıkların özerk varoluşunu tanımaz, ona saygı beslemez; tüm varlıklar, insanın bu yaşamı sürdürebilmesi için sonsuzca ve sorumsuzca kullanılabileceği, sömürebileceği, emre amade hammaddelelerdir. Örneğin, derin bir orman modern teknolojik insana içinde kaybolacağı bir doğal ortam yerine, keresteye dönüştürülebileceği bir rezerv; coşkuyla akan bir ırmak da izlenesi bir manzaradan önce, suyu şişelenip

satılabilecek bir kaynak olarak görünür. Hatta Heidegger (1977) bu koşullar altında insanın kendisinin bile bir hammadde olarak görülüp görülmeyeceğini tartışmaya açmıştır. Heidegger'le beraber anılan bir diğer klasik düşünür de Jacques Ellul'dür. Ellul de modern teknolojinin, yarattığı dönüşümün sağladığı kolaylıklar ve çözümlerden çok daha fazlası olduğuna dikkat çekiyordu. Ellul, teknolojiyi insan müdahalesinden azade, kendi kendini belirleyen ve kendi mantığı olan otonom bir sistem olarak görür. Ona göre, teknoloji modern çağda insanın içinde yaşamak zorunda olduğu bir ortamdır (*milieu*) ve eski ortam olan doğanın yerini almıştır (Ellul, 1962). İnsanın mecbur olduğu bu ortam, her ne kadar yaşamı kolaylaştırıyor ve çözüm sunuyor gibi görünse de, insanı esir almıştır. Heidegger ve Ellul gibi klasik düşünürler modern teknolojiyi böyle belirlenimci (determinist) bir yaklaşım içinde anlamışlardır.

Klasik düşünürlerin bu yaklaşımına zemin hazırlayan bazı olgulardan söz edilebilir. Charlie Chaplin'in *Modern Zamanlar*'da gözler önüne serdiği gibi, Endüstri Devrimi sonrası ortaya çıkan modern teknolojik düzende insan yaşamı adeta bir Fordist üretim bandına, insan da üretim aracının içindeki bir dişli çarka dönüşmüştür. Yaşam, gitgide ölçülebilen parametrelerle hesaplanır olmuş, çalışma ortamları, şehirler, okullar gibi toplumsal hayatın giderek daha fazla alanı verimlilik ve optimizasyon hedefleri doğrultusunda standart yapılarla, en önemlisi de ölçülebilen zamanla örülmüştür. Klasik teknoloji felsefecilerinin bakışında da modern hayatın bu mekanikleşmiş yanına ilişkin bir eleştiri ve endişe halinden söz edebiliriz.

Teknolojiye ilişkin bu karamsar, belirlenimci felsefi görüş, felsefenin Bilim, Teknoloji, Toplum (diğer adıyla Bilim ve Teknoloji Çalışmaları) başta olmak üzere farklı sosyal bilim alanlarıyla temasının arttığı 1980'li yıllara kadar baskın kalmıştır. Ancak 1980'lerde, teknolojiyi bütün ampirik tezahürlerinin altında yatan, toplumsal hayatı kuşatan ve tekdüze bir biçimde belirleyen bir güç veya sistem olarak gören bu bakış açısının aksine, onun toplumsal olarak inşa edildiği düşüncesi merkeze alınmaya başlamıştır. Bu kez, teknolojiyi anlamak için gerekli olan şeyin ampirik olana, yani tekil teknolojilere ve bunların etrafında örülen ilişkilere odaklanmak ve bunları analiz etmek olduğu görüşü benimsenmiştir. 'Ampirik dönüş' (*empirical turn*; Achterhuis, 2001) olarak anılan bu yeni yönelim, teknoloji ile insan varoluşu arasında karşılıklı bir ilişki olduğuna vurgu yapar; teknolojinin insan yaşamını belirlediği görüşünü reddetmez, ancak insanın da teknolojiye etki kapasitesinin olduğunu savunur. Bu çizgideki düşünürlere göre, teknoloji insan varoluşunu dışarıdan belirleyen özerk (otonom) bir güç olarak kavranamaz. Onun insan yaşamı üzerindeki etkisi tekdüze ve tek biçimli değildir. Yani teknoloji, toplumu mutlaka mekanikleşmeye götüren, zorunlu olarak içine gömüldüğümüz ve etkisinden uzak kalamayacağımız bir sistem değildir. Aksine, teknolojilerin alımlanışı, etkisi ve insan yaşamıyla girdiği ilişkiler farklı farklıdır.⁹

⁹ Bu yaklaşımın yaşayan iddialı savunucularından biri olan Peter-Paul Verbeek'in kitaplarından birinin (2005) başlığında yer verdiği "What Things Do" ifadesi tam da bu bakış açısını yansıtır.

Çünkü teknoloji her zaman belirli bir kültürel, tarihsel ve pratik bağlam içinde şekillenir. Aynı teknoloji, farklı zamanlarda ve farklı toplumsal bağlamlarda farklı etkileşimler içinde bulunabilir. Bu nedenle, insan-teknoloji ilişkisini anlayabilmek için ampirik olanı incelememiz gerekir; yani, spesifik teknolojileri farklı insan pratikleri, anlam dünyaları ve kültürel bağlamlarda kurdukları somut, konumlanmış (*situated*) ilişkiler içinde en iyi şekilde anlayabiliriz.

Teknolojinin anlamı, işlevi ve etkilerinin tek bir (istikrarlı) biçime sahip olmadığını, bu ilişkinin bağlama göre çeşitlenebileceğini görmek için post-fenomenolojik teknoloji felsefesi geleneğinin kurucusu Don Ihde'nin alana kazandırdığı 'çoklu istikrar' (*multistability*) kavramına başvurabiliriz. Ihde'ye göre bir teknoloji, kendi içinde sabit bir "öz" taşımaz; kullanıcının onunla kurduğu ilişki, teknolojinin anlamını, işlevini ve etkisini belirleyen temel unsurdur. Ihde'nin verdiği çarpıcı örneklerden biri, 1930'larda Avustralyalı altın arayıcılarının Yeni Gine'de bıraktığı sardalya kutularının yerli halk tarafından süs nesnesi olarak alna takılarak kullanılmaya başlamasıdır (Ihde, 1990). Batılı gözle "atık" olarak görülebilecek bu metal kutular, farklı bir kültürel bağlamda tamamen farklı bir işlev, anlam ve değer kazanmıştır. Diğer bir deyişle, teknolojik bir nesne olan sardalya kutusunun istikrarlı (*stable*) kullanım biçimi, insanların ona yüklediği anlam ve toplumsal etkisi, kültürel bağlama göre farklılaşmıştır. 'Çoklu istikrar' kavramı, bize teknolojinin araçsalcılığın iddia ettiği gibi nötr ve evrensel bir araç olmadığını, belirlenimciliğin savunduğu gibi de toplumu tek bir biçimde dönüştürme etkisine sahip otonom bir güç olmadığını göstermektedir.

Şimdi yapay zekânın eğitim alanında kullanımını teknoloji felsefesinin bu üç başat görüşü arasında konumlandırmaya çalışalım.

Araçsalcı bakış açısı yapay zekâyı nötr bir araç olarak ele alır. Yapay zekânın kendinden ileri gelen bir değeri, anlamı yoktur. Değer ve anlam, onun pratik kullanımında doğar; yani kullanım amacı ve ortaya çıkardığı sonuçlar üzerinden belirlenir. Yapay zekânın kendisi iyi ya da kötü değil, nötrdür. İyiliği ya da kötülüğüne ilişkin çıkarımlar ancak eğitimde kullanıldığında ne işe yaradığı, hangi amaçlara hizmet ettiği ve eğitimi nasıl etkilediği soruları üzerinden değerlendirilmelidir. Örneğin, yapay zekânın kitlesel eğitimin bireysel farklılıkları yeterince dikkate almaması gibi yapısal problemler karşısında sunabileceği müdahale olanaklarına ilişkin iyimser söylem araçsalcıdır. Buna göre, geleneksel okul yapısında öğretmen aynı sınıftaki çok sayıda öğrenciye büyük ölçüde aynı içeriği, benzer bir hızda ve sınırlı bir kanal yoluyla sunmak durumundadır. Bu yapı, öğrenciler arasındaki hazırbulunuşluk, öğrenme hızı ve yöntemi, ilgi ve öğrenme güçlüğü farklılıklarını karşılamada yetersiz kalabilir. Yapay zekâ desteğiyle öğrencilerin performanslarına ilişkin veriler analiz edilerek eğitim içeriğinin seviyesini, ödev biçimlerini ve geri bildirim yollarını farklı öğrencilere göre uyarlamak mümkün olacaktır. Bir başka klasik örnek vermek gerekirse, bir öğrenci bir üretken yapay zekâ aracını kendi yazdığı metne ilişkin geri bildirim almak ya da metni tartışıp geliştirmek için kullanıyorsa, yapay zekâ öğrenmeye hizmet eden bir pedagojik araç olarak işlev görmüş olur. Öte yandan, öğrenci

üretken yapay zekâya ödevi tümüyle yaptırır ve ortaya çıkan metne kayda değer bir katkı yapmadan teslim ederse, öğrenme amacına hizmet etmediği için yapay zekânın kötüye kullanıldığını ileri sürebiliriz.

Araçsalıcı bakış açısını daha eleştirel kılmak ve derinleştirmek için, teknolojinin kullanımını değerlendirirken neyi önemseyeceğimizi doğru tespit etmek büyük önem taşır. Zira bir teknoloji mevcut bir yapısal soruna çözüm üretirken bile yeni sorunlar ortaya çıkarabilir. Bu nedenle aracın değeri, taşıdığı teknolojik potansiyelden ziyade, eğitim deneyimini dönüştürme biçimi ve ürettiği sonuçlar üzerinden belirlenmelidir. Yapay zekâ araçları öğrencilerin ve öğretmenlerin işini kolaylaştırıyor, onlara zaman kazandırıyor olabilir; ancak eğitim açısından önemseydiğimiz sonuç, ödev tesliminin, not vermenin veya ders hazırlamanın kolaylaşması mıdır, yoksa öğrenmenin desteklenmesi mi? Daha iyi öğrenmeye yol açmayan bir teknoloji kullanımını onaylıyor muyuz? Yapay zekâ ile eğitimi kişiselleştirerek kitlesel eğitimin kronik bir sorununa çözüm üretirken, bir yandan da öğrenciyi sürekli ölçülen, sınıflandırılan ve algoritmalarla yönlendirilen bir nesneye dönüştürmüş olmuyor muyuz?

Öte yandan, önemseydiğimiz sonucu doğru tespit etsek bile, hâlen araçsalıcı yaklaşıma yönelik bir eleştiri getirmek mümkündür. “Yapay zekâyı doğru bir amaçla, yani öğrenmeyi desteklemek amacıyla, nasıl daha verimli ve uygun şekilde kullanırız?” sorusuna odaklanan araçsalıcı yaklaşım, etik sorunu kullanıcı davranışına indirgemeye neden olabilir. Ve bunu yaparken göz ardı ettiği önemli bir boyut vardır: Yapay zekâ ile ilgili etik meseleleri yalnızca kullanımda aramak ve sorumluluğu nihai kullanıcılara at(a)mak. Peki, bu doğru mudur? Yapay zekâya ilişkin etik sorumluluğu nerede ve kimde aramalıyız?

Belirlenimci görüş, yapay zekâyı klasik düşünürlerin hatlarını çizdiği modern teknoloji çizgisinin bir devamı olarak değerlendirir; onun bizi ağına aldığını (veya bir gün alacağını), karşısındaki etkimizin son derece sınırlı olduğunu ve yapay zekânın eğitimi baştan aşağı dönüştürme gücüne sahip olduğunu varsayar. Burada olanaklardan ziyade tehditlere odaklanan bir bakış açısı görürüz. Etik değerlendirme modern teknolojinin kendisinin bir tehdit olması fikrine dayalıdır ve değerlere referansla yapılmaz. Gerçekten de yapay zekâ, Heidegger’in resmettiği modern teknolojinin ve Ellul’un betimlediği ‘teknolojik toplum’un özelliklerini taşımakta ve desteklemektedir. Şöyle ki: Diğer pek çok modern teknolojik aygıt gibi yapay zekâ da enerjiye ve doğal kaynaklara bağımlıdır. Nitekim, yapay zekâdan sıklıkla veriye ve enerjiye aç bir sistem olarak bahsedildiğini görmekteyiz. Yapay zekânın ihtiyaç duyduğu veri, enerji, su ve nadir elementler için insan ve doğa, emre amade hammaddeler olarak görülmekte, büyük ölçekli emek, veri ve doğa talanları yapılmaktadır. Kate Crawford, bu talanı Yeryüzü (*Earth*), Emek (*Labour*) ve Veri (*Data*) şeklinde üç ana bölümden oluşan kitabı *Atlas of AI*’da (2021) detaylı bir şekilde ortaya koymaktadır. Crawford, *Atlas*’ta yapay zekânın görünmeyen katmanlarını ve bedellerini gözler önüne sermeyi amaçladığını ifade etmekte; yapay zekânın soyut bir matematiksel sistem olduğu yönündeki yaygın görüşün aslında politik bir yanılsama yaratma amacı

güttüğünü, bu teknolojinin lityum madenlerinden veri merkezlerine, veri etiketleme emeğinden e-atık çöplüklerine uzanan devasa bir madencilik (*extraction*) ve sömürü ağı olarak işlediğini çarpıcı örneklerle ortaya koymaktadır.¹⁰

Crawford'un tespitlerinden yola çıkarak, Heidegger'in modern teknoloji karşısında insanın kendisinin hammadde haline gelip gelmediğine ilişkin sorusuna, yapay zekânın bu dönüşümü diğer teknolojilerden çok daha açık biçimde gerçekleştirdiğini söyleyerek yanıt verebiliriz. Zira yapay zekâ çağında, teknoloji şirketlerinin gözünde insan, bu sisteme veri sağlayan bir hammadde konumuna indirgenmiştir. Yapay zekânın giderek daha çok nüfuz ettiği eğitim sistemine bu açıdan bakıldığında, onun öğretmen ve öğrencinin elinde birer araç olduğunu savunmak mümkün müdür? Aksine, öğretmen ve öğrenci yapay zekâ teknolojisinin işleyişine eklemlenmiş araçsal öğelere dönüşmüş durumdadır. Öğrenci bu işleyişe veri akışı sağlayan bir kaynak, öğretmen de teknolojik eğitim sisteminde bir kullanıcı konumuna sürüklenmektedir. Modern teknolojinin bir tezahürü olan yapay zekâ, insan faktörünü daha da silikleştirerek eğitimi mekanikleşmeye doğru götürecektir, etkisinden uzak kalamayacağımız bir sistemi beslemektedir.

Peki insan-teknoloji ilişkisinde tek belirleyici olanın teknoloji olduğunu ileri sürmek, başta toplumsal bir aktör olarak insanın, sonra da bu toplumsal aktörün içinde şekillendiği başka tarihsel ve toplumsal yapıları göz ardı etmek anlamına gelmez mi? Bizatihi teknoloji, kendisinin de içinde meydana geldiği bir tarihsellikten ve toplumsallıktan azade düşünülebilir mi? İşte bu soru, bizi yapay zekâyı ampirik olarak incelemeye çağıran görüşe yönlendiriyor. Ampirik dönüş çizgisindeki düşünürlerin insan-teknoloji ilişkisini anlayabilmek için teknolojileri konumlanmış ve farklılaşan bağlamlarda incelemenin önemine işaret

¹⁰ Hakikaten, yapay zekâ geliştirme süreçlerinde telif hakkı ihlalleri, kişisel verilerin rıza alınmaksızın kullanılması ve şeffaf olmayan uygulamalar gibi konularda giderek daha fazla hukuki inceleme ve düzenleme yapılıyor olsa da, çoktan pek çok ihlal gerçekleştirmiş durumda. Örnek olarak, teknoloji devlerinin veri yağmalamalarına ilişkin bazı güncel haberler için şu kaynaklara bakılabilir:

<https://www.reuters.com/technology/french-competition-watchdog-hits-google-with-250-mln-euro-fine-2024-03-20/>

https://en.wikipedia.org/wiki/The_New_York_Times_v._Microsoft_and_OpenAI

<https://www.theguardian.com/technology/2025/jun/25/anthropic-did-not-breach-copyright-when-training-ai-on-books-without-permission-court-rules>

<https://www.aa.com.tr/tr/bilim-teknoloji/amazona-twitch-yayinlarini-yapay-zeka-egitimi-icin-kullandigi-iddiasıyla-toplu-dava-acildi/4036429>

<https://www.reuters.com/sustainability/society-equity/clearview-ai-faces-criminal-complaint-austria-suspected-privacy-violations-2025-10-28/>

<https://haber.sol.org.tr/haber/yapay-zeka-sirketleri-sahafllari-bosaltiyor-milyonlarca-kitabi-alip-tarayip-imha-ettiler>

ettiklerinden söz etmiştik. Bu görüşe göre, tekdüze ve tek biçimli bir etkiye sahip olan bir modern teknoloji den bahsedemeyiz.

İhde'nin 'çoklu istikrar' kavramı, yapay zekânın ne nötr bir araç ne de toplumu tek bir biçimde dönüştüren özerk bir güç olduğuna işaret eder. Yapay zekânın etkileri, anlamları ve kullanım biçimleri, onun tüm yaşam döngüsü boyunca içinde bulunduğu tarihsel ve toplumsal ilişkiler, yapılar ve güç ilişkileri tarafından şekillenmeye açıktır. Bu yaklaşım, yapay zekânın insan yaşamını şekillendiren ve belirleyen gücünü kabul etmekle birlikte, insanların da onun gelişimini ve kullanımını yönlendirebilecek bir etki kapasitesine sahip olduğunu savunur. Peki, bu yaklaşım normatif tutumumuzu, sorumluluk konusundaki bakış açımızı nasıl etkiler?

Karl Popper 1968 yılında Viyana'da düzenlenen Uluslararası Felsefe Kongresi'nde konuşma yaptığında, konuşmasına başlamadan önce salondaki herkesin "Bilim İnsanın Ahlaki Sorumluluğu" başlıklı konuşmasının aslında "biyolojik ve nükleer savaş meselesi için kullanılan bir başka söz" olduğunu açıkça anladığını belirtmiştir. Buna karşılık, günümüzde bilgisayar mühendisleri ve veri bilimcilerinin kendi geliştirdikleri teknolojilere ilişkin ahlaki sorumlulukları konusunda benzer bir ortak uzlaş henüz oluşmuş değildir. Bunun yerine, 'etik tasarım' olarak adlandırdığımız giderek büyüyen bir hareket, yapay zekâ/makine öğrenmesi alanında böyle bir uzlaşmayı geliştirmeye, ya da en azından tartışmanın çerçevesini şekillendirmeye, çalışmaktadır. (Greene v.d., 2019)

Bu alıntının işaret ettiği üzere, etik sorumluluk kullanım aşamasının ötesinde, yapay zekâ teknolojisinin tasarımında da aranabilir. Teknolojiye ilişkin etik sorumluluğun yalnızca nihai kullanıcıların davranışlarında aranması, etik sorunları bireysel kullanıcıların tercihiymiş gibi algılayarak, teknolojinin içinde üretildiği geniş bağlamı görünmezleştirir. Bu açıdan bakıldığında, yapay zekâ sistemlerinin tasarımında ve geliştirilmesinde rol alan aktörlerin de sorumluluk taşıdığı açıktır.

Ancak 'geniş bağlamın' sadece üretim ve tasarımda rol oynayan ve ulaşılabilen kaynakları (veri setleri, donanımlar, insan kaynakları vb.) değil, aynı zamanda tasarımcıların değer yargıları, düşünceleri ve önyargılarını da kapsadığının farkında olmak önemlidir. Bir yapay zekâ ekibinin ulaşabildiği eğitim verisi setleri kadar, hangi değerleri ve ilkeleri önemseyerek ve bunları nasıl yorumlayarak çalıştıkları da teknolojinin etik kullanımına ilişkin sağlam ya da aksak bir zemin oluşturacaktır. Örneğin, dağıtıcı adalet kavramı, eğitimdeki kaynak, fırsat ve risklerin öğrenciler arasında demografik olarak ayırım yapmadan hakkaniyetli bir şekilde paylaştırılmasına dayanır. Tasarım aşamasında bu bakış açısı benimsendiği takdirde, örneğin okul sonrası kulüplere öğrenci yerleştirmede kullanılan bir algoritma, kız ve erkek öğrencileri eşit şansla eleyecek ya da seçecektir. Onarıcı adalet yaklaşımının benimsenmesi durumunda ise yapay zekâ sistemlerine itiraz ve telafi mekanizmaları entegre edilmeye çalışılacaktır.

Sistemin hatalı bir karara yönlendirmesi durumunda, kullanıcının buna itiraz edebileceği veya hatayı düzeltebileceği bir müdahale arayüzünün tasarlanması düşünülebilir.

Yukarıdaki alıntıda Greene ve diğerlerinin ‘etik tasarım’ olarak bahsettiği hareket, günümüzde değer temelli tasarım (*value sensitive design - VSD*) başlığı altında toplayabileceğimiz bir dizi yaklaşımda ifade bulunmaktadır. VSD, teknolojik sistemlerin tasarım ve geliştirme süreçlerine etik değerlerin sistematik biçimde entegre edilmesini amaçlayan bir tasarım yaklaşımıdır. Teknolojinin yalnızca teknik gereksinimler doğrultusunda değil, tasarımdan etkilenen farklı paydaşların gözettiği değerleri ve çıkarları da dikkate alarak geliştirilmesi gerektiğini savunur. Şeffaflık, hesap verebilirlik, adalet ve mahremiyet gibi değerlerin yapay zekâ araçlarının tasarımına dâhil edilmesi elbette önemlidir. Ancak bu değerlerin kendilerinin nötr ve tartışmasız olduğunu varsayabilir miyiz? Hangi değerlerin tasarıma dâhil edileceği, bunların nasıl tanımlanacağı, hangi teknik özelliklere dönüştürüleceği ve aralarındaki çatışmaların nasıl çözüleceği gibi sorular aslında politik nitelik taşır.

Teknolojik çözümçülük (*technological solutionism*) olarak adlandırılan bakış açısı, aslında yapısal nitelikteki toplumsal ve etik sorunları uzmanların düzeltebileceği teknik kusurlar olarak algılama eğilimi olarak tanımlanabilir. Bu bakış açısını benimseyerek, örneğin, algoritmik ayrımcılığın daha iyi veri setleri kullanarak, daha adil algoritmalar yazarak veya daha etkili denetim mekanizmaları kurarak önlenebileceğini varsayarak teknik çözümler üretmeye çalışabiliriz. Ancak bunu yaparken, ayrımcılığın ortaya çıkmasına sebep olan toplumsal, tarihsel veya kurumsal koşulları göz ardı ediyor olabiliriz. Bir algoritmanın teknik açıdan “daha adil” hale getirilmesi, bu algoritmanın hangi toplumsal koşullardan kaynaklandığı, hangi amaçlara hizmet ettiği, kim tarafından geliştirildiği ve hangi güç ilişkileri içinde kullanıldığı gibi soruları ortadan kaldırmaz. Aksine, üretilen teknik çözümler bu toplumsal sorunların ve teknoloji dışındaki olası çözümlerin göz ardı edilmesine yol açabilir.

Bir lisenin öğrencilerin okulu bırakma olasılığını önceden belirlemek amacıyla sınav notları, devamsızlık ve ödev teslimi gibi verilere dayanan bir algoritma kullandığını düşünelim. Sistemin düşük gelirli ailelerden gelen öğrencileri sistematik olarak daha “riskli” olarak sınıflandırdığı ortaya çıksın. Teknolojik çözümçü bir yaklaşım, öncelikle sistemin eğitiminde kullanılan veri setlerindeki yanlılığı gidermeye, sorun yaratan değişkenleri tahmin mekanizmasından çıkarmaya veya algoritmayı daha adil tahminler üretecek şekilde yeniden tasarlamaya yönelecektir. Ancak bu teknik müdahaleler, düşük gelirli ailelerden gelen öğrencilerin neden daha yüksek risk altında oldukları sorusunun göz ardı edilmesine neden olabilir. Örneğin, ödevlerini geciktirip aksatması, bir öğrencinin motivasyon eksikliğinden ziyade bir işte çalışmak zorunda olmasıyla, notlarının düşüklüğü de okul sonrası kurslara erişememesiyle ilgili olabilir. Dolayısıyla, teknik olarak algoritmanın daha doğru tahminler üretmesi sağlansa bile, bu, diğer eşitsizliklerin ortadan kalkmasını sağlamaz.

Hatta öğrenciyi “riskli” olarak tanımlayan bu tür algoritmik sınıflandırmalar, aslında yapısal veya kurumsal olan sorunları öğrencinin bireysel performansına ilişkin bir problem gibi ele alarak sorumluluğun yeniden öğrenciye ya da velisine yüklenmesine yol açabilir. Özetle, burada teknolojik çözümçülüğün gölgede bırakabileceği asıl sorun, öğrencinin hangi toplumsal ve kurumsal koşullar içinde “riskli” hale geldiği ve bu koşullar karşısında sorumluluğun nasıl tanımlanıp üstlenildiğidir.

Sonuç

Etik sorunları yapay zekâyı kullanma biçimimizden önce tasarım aşamasında aramak ve onları değerler üzerinden ya da teknolojik bir yaklaşımla çözmeye çabalamak da aslında meseleye dar bir çerçeveden bakmak anlamına geliyor olabilir. Teknik açıdan daha gelişmiş, daha verimli ya da etik değerleri gözetken bir yapay zekâ tasarımının kendiliğinden daha adil bir eğitim düzeni yaratacağı varsayımı, yapay zekâ sistemlerinin içinde işlediği tarihsel, toplumsal, politik ve ekonomik eşitsizlikleri göz ardı etmemize sebep olabilir. Bu nedenle yapay zekâyı ilişkin etik sorumluluğu belirli bir aktörün (kullanıcının, tasarımcının ya da geliştiricinin) niyet ve davranışlarına indirgemek yerine, yapay zekâyı bir sosyoteknik sistem olarak ele alan bağlamsal bir yaklaşıma ihtiyaç vardır (Johnson ve Verdicchio, 2025).

Johnson ve Verdicchio’nun (2025) vurguladığı üzere, yapay zekâ sistemleri yalnızca teknik bileşenlerden ibaret değildir; onların hem tasarımını hem kullanımını şekillendiren toplumsal normlar, ekonomik çıkarlar, kurumsal tercihler ve güç ilişkileri de bu sosyoteknik sistemin ayrılmaz birer parçasıdır ve etik sorumluluğun doğru şekilde anlaşılması için analiz edilmeleri gerekir. Böylesi bir perspektif, doğru sorgulamayı “Yapay zekâ ile eğitim öğretimi nasıl dönüştürürüz?” ya da “Yapay zekâyı nasıl sorumlu kullanırız?” sorularının ötesine taşıyarak, “Yapay zekânın geliştirilmesine, kullanımına ve sonuçlarının değerlendirilmesine kimler, nasıl karar veriyor ve nasıl çıkar sağlıyor?” şeklinde güncellememizi gerektiriyor. Bu da eğitimcilerin ve öğrencilerin yapay zekâyı sadece “kullanan” konumunda kalmayıp, onun tasarımını, ürettiği çıktıları ve bu çıktıların nasıl değerlendirildiğini sorgulayan bir okuryazarlık geliştirmelerini desteklemekle mümkün olabilir.

Bu bakış açısı, yukarıda değinmiş olduğumuz, yapay zekânın barındırdığı riskleri ve etik yorumlamayı da onları şekillendiren tarihsel, toplumsal ve kurumsal koşullar içerisinde değerlendirmeyi gerektirir. Buradan hareketle geliştirilecek çözümlerin de sadece bireysel kullanım pratiklerine veya teknik düzenlemelere odaklanmayıp, söz konusu koşullarla ilişkili olarak yapılandırılması gerekir. Örneğin, klasik teknoloji felsefesi yaklaşımından hareketle, veri sömürüsü ve gözetimin öğrenci ve öğretmenlerin davranışlarını, ilişkilerini ve öznelliklerini nasıl dönüştürdüğüne odaklanabiliriz. Ancak, bu dönüşümün kimlerin çıkarlarına hizmet ettiğine, mevcut güç ilişkilerini nasıl sağlamlaştırdığına veya kimlerin daha fazla güç kazanmasına yol açtığına dikkat ederek değerlendirilmesi, farklı bir etik-

politik perspektif geliştirmemizi sağlayabilir. Benzer şekilde, öğretmenin ve öğrencinin yapay zekâya aşırı yaslanmasına araçsal bir bakış açısıyla yaklaştığımızda, yalnızca bu bireylerin özerkliği ya da bilişsel becerileri üzerindeki sonuçlarını düşünmekle sınırlı kalabiliriz. Oysa, böyle bir bağımlılığı doğuran koşulları ve bu yönelimle şekillenen eğitim dünyasında kimlerin sistematik olarak daha kırılgan veya bağımlı konuma sürüklendiğini görmeye çalışmak farklı çözüm arayışlarına yönelmemize olanak sağlayabilir.

Kaynakça

- Achterhuis, H. (ed.). (2001). *American Philosophy of Technology: The Empirical Turn*. Bloomington: Indiana University Press.
- Balázs, L. (2026). Epistemic Trust in Generative AI as an Information Source: Development and Validation of the Trust in AI-Epistemic Scale (TAI-E). *Social Sciences*, 15(7), 433. <https://doi.org/10.3390/socsci15070433>
- Coeckelbergh, M. (2022). *The political philosophy of AI: An introduction*. Polity.
- Crawford, K. (2021). *Atlas of AI: Power, Politics, and the Planetary Costs of Artificial Intelligence*. Yale University Press.
- Cressman, D. (2022). A short introduction to the philosophy of technology. In Swierstra, T., Lemmens, P., Sharon, T. & Vermaas, P. (eds.), *The Technical Condition: The Entanglement of Technology, Culture, and Society*. Amsterdam: Boom Uitgevers, p. 43–75.
- Ellul, J. (1964). *The Technological Society*. New York: Vintage Books.
- Greene, D., Hoffmann, A. L., & Stark, L. (2019). Better, nicer, clearer, fairer: A critical assessment of the movement for ethical artificial intelligence and machine learning. *Proceedings of the 52nd Hawaii International Conference on System Sciences*. https://aisel.aisnet.org/hicss-52/dsm/critical_and_ethical_studies/2/
- Heidegger, M. (1977). The Question Concerning Technology. In *The Question Concerning Technology and Other Essays*. NY& London: Garland Publishing.
- Ihde, D. (1990). *Technology and the Lifeworld: From Garden to Earth*. Bloomington: Indiana University Press.
- Johnson, D. G., Verdicchio, M. (2025). The sociotechnical entanglement of AI and values. *AI & Society* 40, 67-76. <https://doi.org/10.1007/s00146-023-01852-5>

Neuberger, C., Langenohl, S., & Nuernbergk, C. (2025). Epistemic authority in the digital public sphere: An integrative conceptual framework and research agenda. *Communication Theory*. Advance online publication.

<https://doi.org/10.1093/ct/qtae020>

UNESCO (2021). Recommendation on the Ethics of Artificial Intelligence.

<https://www.unesco.org/en/articles/recommendation-ethics-artificial-intelligence>

Verbeek, P. P. (2005). *What Things Do. Philosophical Reflections on Technology, Agency and Design*. Penn State University Press.

Yükseköğretim Kurulu (2024). *Yükseköğretim Kurumları Bilimsel*

Araştırma Ve Yayın Faaliyetlerinde Üretken Yapay Zekâ Kullanımına Dair Etik

Rehber. <https://proje.yok.gov.tr/tr/page/635>