YAPAY ZEKÂ VE İNSANLIĞIN EGEMENLİK KRİZİ

İmdat Demir — Filozof Kirpi

ÖZET

Bu çalışma, yapay zekâ tartışmasını teknik kapasite meselesinin ötesine taşıyarak insanlığın egemenlik problemi olarak ele almaktadır. Temel soru, yapay zekânın insanı "geçip geçmeyeceği" değil, insanın karar verme, denetleme ve sorumluluk üstlenme yetkilerini ne ölçüde makineye devredeceğidir. Çalışma, yapay zekânın bilinçli olmasının tehlike için zorunlu olmadığını; asıl riskin kapasite, özerklik, erişim, yetki ve kaynakların aynı sistemde birleşmesinden doğduğunu savunmaktadır. Otonom ajanlar, hizalama sorunları ve kontrol kaybı ihtimali bu çerçevede incelenirken, yapay zekânın emek piyasasında işsizliğin ötesinde ücretleri, meslekî özerkliği, pazarlık gücünü ve sınıfsal bölüşümü dönüştürebileceği gösterilmektedir. Metin ayrıca bilişsel taşeronlaşma, otomasyon yanlılığı ve epistemik bağımlılık üzerinden insanın kendi muhakemesini zayıflatma ihtimalini tartışmaktadır. Devlet, sermaye ve dijital platformların yapay zekâ aracılığıyla gözetim, yönlendirme ve davranış tahmini kapasitesini artırması ise "algoritmik Leviathan" kavramıyla analiz edilmektedir. Buna karşılık insanî üstünlük hız veya hesaplama gücünde değil; hikmet, ahlâkî sorumluluk, anlam kurma, adalet ve meşruiyet üretme kapasitesinde aranmalıdır. Sonuç olarak çalışma, yapay zekâyı yasaklamak yerine onu anayasal sınırlar içine almayı; insan veto hakkını, bağımsız denetimi, hukukî sorumluluğu, çocukların korunmasını, askerî kullanımda kırmızı çizgileri ve uluslararası gözetimi savunmaktadır. İnsanlığın görevi makineyle yarışmak değil, kendi ürettiği kudreti hukuk, ahlâk ve demokratik denetim altında tutmaktır. Böylece yapay zekâ çağının asıl çatışması insan ile makine arasında değil, insanın teknolojik kudreti hangi siyasî, ekonomik ve ahlâkî düzen içinde kullanacağı üzerinde yoğunlaşmaktadır. Yapay zekâ bilimi, eğitimi, sağlığı ve üretkenliği geliştirebilir; fakat denetimsiz bırakıldığında eşitsizliği, gözetimi, bağımlılığı ve otoriter kapasiteyi büyütebilir. Bu nedenle geleceğin temel ilkesi, yetkinin sınırlanması, sorumluluğun insanda kalması ve insan haysiyetinin teknolojik verimlilikten üstün tutulmasıdır. Her durumda ve her düzeyde.

1. TEHLİKENİN ONTOLOJİSİ: YAPAY ZEKÂ ARAÇ MI, AKTÖR MÜ, YENİ BİR İKTİDAR BİÇİMİ Mİ?

Yapay zekâ hakkındaki çağdaş tartışmanın en büyük kusuru, henüz sorunun ne olduğunu belirlemeden cevabın peşine düşmesidir. Bir kesim yapay zekâyı insanlığın bütün sorunlarını çözecek teknolojik mesih gibi sunarken diğer kesim onu insan türünü tasfiye edecek yeni bir canavar olarak tasvir ediyor. Oysa her iki anlatı da aynı düşünsel hatanın farklı yüzleridir: Yapay zekâyı kendi başına tarih yapan bağımsız bir özne olarak düşünmek. Meseleye daha soğukkanlı bakıldığında asıl sorunun makinenin "kötüleşmesi" değil, makineye hangi yetkilerin verildiği olduğu görülür. Yapay zekânın insan için oluşturabileceği tehlikeyi anlamak istiyorsak önce zekâ, bilinç, özerklik, amaç, eylem, kontrol ve iktidar kavramlarını birbirinden

ayırmak zorundayız. Aksi hâlde teknik bir sistemi mitolojik bir varlığa dönüştürür, gerçek tehlikeyi de tam bu mitolojinin arkasında görünmez hâle getiririz.

Her şeyden önce yapay zekânın "zeki" olması onun bilinçli olduğu anlamına gelmez. Bir sistem karmaşık matematik problemleri çözebilir, milyonlarca metni karşılaştırabilir, yazılım geliştirebilir, görüntüleri tanıyabilir, stratejik seçenekleri değerlendirebilir ve insanın saatlerce yapacağı işleri saniyeler içinde gerçekleştirebilir. Fakat bütün bunlar, sistemin kendi varlığının farkında olduğunu, acı çektiğini, ölümden korktuğunu, kendisine ait bir hayat tasavvuru geliştirdiğini veya ahlâkî sorumluluk taşıdığını göstermez. Zekâ ile bilinç aynı ontolojik kategori değildir. İnsanlık açısından esas tehlike de tam burada başlar; çünkü tehlikeli olabilmek için bilinç sahibi olmak gerekmez. Bir baraj kapağını yanlış zamanda açan otomasyon sistemi bilinçsizdir ama binlerce insanı öldürebilir. Finansal piyasada saniyeler içinde milyarlarca dolarlık işlem yapan algoritmanın hırsı yoktur; fakat ekonomik kriz üretebilir. Bir askerî hedefleme sisteminin nefrete ihtiyacı yoktur; yanlış veya denetlenmeyen bir hedefleme mekanizması yeterlidir. Dolayısıyla yapay zekâ tehlikesinin merkezine "bilinç kazanacak mı?" sorusunu yerleştirmek, asıl problemi yanlış yere taşımaktır.

Daha önemli soru şudur: Yapay zekâ ne ölçüde **eylem kapasitesine** sahip olacaktır? Bugüne kadar bilgisayar sistemlerinin büyük bölümü insan tarafından verilen sınırlı komutları yerine getiriyordu. Yeni nesil yapay zekâ sistemlerindeyse farklı bir mimarî ortaya çıkıyor. Sistem yalnız cevap üretmiyor; hedefi alt görevlere ayırabiliyor, araç kullanabiliyor, yazılım çalıştırabiliyor, internet ortamlarında işlem yapabiliyor, başka sistemlerle iletişim kurabiliyor ve belirli ölçülerde kendi faaliyet zincirini sürdürebiliyor. Bu dönüşüm, "cevap veren makine"den "iş yapan makine"ye geçiştir. Uluslararası Yapay Zekâ Güvenliği Raporu da kontrol kaybı bakımından önemli görülen yetenekler arasında uzun dönemli planlama, özerk hareket, gözetimden kaçınma, ikna, yanıltma ve kendi faaliyetini sürdürmeye yönelik davranışları saymaktadır. Bugünkü sistemlerin henüz insanlığın kontrolünü ortadan kaldıracak seviyede olmadığı özellikle belirtilmekle birlikte, bu alanlarda ilerleme bulunduğu da kaydedilmektedir.

Buradaki tarihî eşik, yapay zekânın ne kadar bildiğinden çok, **bildiği şeyle ne yapabildiği** meselesidir. Bilgi tek başına iktidar değildir. Bilginin kaynaklara, iletişim ağlarına, finansal araçlara, kritik altyapılara, silahlara, enerji sistemlerine, üretim araçlarına veya yönetim süreçlerine erişebilmesi gerekir. Bir yapay zekâ modeli bilgisayarın içinde kapalı tutulduğunda sahip olduğu kapasite ile banka hesaplarına, kamu veri tabanlarına, enerji şebekelerine, robotik sistemlere ve askerî altyapıya bağlandığında sahip olduğu kapasite aynı değildir. Bu nedenle yapay zekâ riskini yalnızca modelin zekâ seviyesi üzerinden ölçmek son derece yetersizdir. Gerçek risk kabaca şu bileşimden doğar: **kapasite × özerklik × erişim × yetki × kaynak.** Bu unsurların her biri büyüdükçe potansiyel zarar alanı genişler.

Tam da bu nedenle ileri yapay zekâ geliştiren şirketlerin kendi güvenlik politikalarında "özerklik", "kontrol kaybı", "siber kapasite", "biyolojik risk", "manipülasyon" ve "yapay zekâ araştırmasını otomatikleştirme" gibi kategoriler giderek daha belirgin hâle geliyor. Google DeepMind'ın Frontier Safety Framework'ü özerklik ve siber güvenlik gibi alanları kritik yetenek eşikleri üzerinden değerlendiriyor; Anthropic'in Responsible Scaling Policy'si model yetenekleri yükseldikçe daha ağır güvenlik ve yönetişim tedbirleri öngörüyor; OpenAI'nin 2026'da yayımladığı Frontier Governance Framework ise zarar verici manipülasyon ve kontrol kaybını ileri sistemler açısından izlenmesi gereken risk alanları arasında sayıyor. Bu durum tek başına yaklaşan bir felâketin kanıtı değildir. Fakat teknolojiyi geliştiren kurumların dahi "daha güçlü modeller daha ağır güvenlik rejimleri gerektirebilir" varsayımıyla hareket etmesi, meseleyi fantezi olarak küçümsemenin de bilimsel olmadığını gösterir.

Bununla birlikte burada başka bir siyasal soru ortaya çıkmaktadır: Yapay zekâ şirketlerinin felâket senaryolarını dile getirmesi yalnızca kamusal sorumluluğun sonucu mudur? Yoksa bu söylem aynı zamanda geleceğin regülasyon düzenini belirleme mücadelesinin bir parçası mıdır? Çok güçlü şirketler "bu teknoloji

son derece tehlikelidir" dediğinde devletler doğal olarak daha ağır giriş koşulları, daha büyük sermaye yükümlülükleri ve daha pahalı güvenlik standartları getirebilir. Böyle bir düzenleme küçük şirketleri ve bağımsız araştırmacıları piyasadan uzaklaştırırken mevcut devleri daha da güçlendirebilir. Dolayısıyla yapay zekâ güvenliği söylemi bile siyasî iktisattan bağımsız okunamaz. Bir şirketin güvenlik talebi gerçekten gerekli olabilir; aynı talep aynı zamanda piyasaya giriş duvarlarını yükseltebilir. Otopsinin görevi teknoloji şirketlerinin niyetini şeytanlaştırmak değil, çıkar yapılarını görünür kılmaktır.

Asıl kavramsal dönüşüm burada ortaya çıkar: Yapay zekâyı yalnızca "araç" olarak görmek artık yeterli değildir; fakat onu insan gibi bağımsız bir "özne" olarak görmek de erkendir. Yapay zekâ giderek **aracı aktör** hâline gelmektedir. İnsan tarafından tanımlanan hedefler doğrultusunda, fakat her ara adımı insan tarafından belirlenmeden faaliyet gösterebilen bir sistemden söz ediyoruz. Böyle bir sistem hukuken ve ahlâken sorumlu değildir; fakat fiilen sonuç üretebilir. İşte önümüzdeki yılların hukukî krizlerinden biri tam burada yaşanacaktır. Bir yapay zekâ ajanı şirket adına yanlış yatırım yaptığında, hastaya yanlış tedavi önerdiğinde, milyonlarca kişiyi etkileyen kamu kararına katkıda bulunduğunda veya bir güvenlik sisteminde hatalı işlem gerçekleştirdiğinde sorumluluk kime ait olacaktır? Yazılımcıya mı, şirkete mi, kullanıcıya mı, kuruma mı, modele mi? "Algoritma yaptı" cümlesinin modern toplumun en tehlikeli sorumluluktan kaçış biçimlerinden birine dönüşme ihtimali vardır.

Bu noktada iktidar kavramını yeniden düşünmek gerekir. İktidar yalnız emir vermek değildir; seçenekleri belirlemek, görünür olanı seçmek, insanın neyi göreceğini düzenlemek, davranışları tahmin etmek, tercihleri yönlendirmek ve alternatifleri görünmez hâle getirmek de iktidardır. Bugünün tavsiye algoritmaları bile hangi haberi göreceğimizden hangi ürünü satın alacağımıza kadar davranış alanımızı şekillendirmektedir. Üretken ve ajanlaşan yapay zekâ bu kapasiteyi çok daha ileri bir düzeye taşıyabilir. Sistem yalnız "ne istediğimizi" tahmin etmeyecek; ne istememiz gerektiğini bize önerecek, seçenekleri sıralayacak, kimi zaman kararımızı bizim adımıza uygulayacaktır. Böylece iktidar kaba yasaktan yumuşak yönlendirmeye, emirden öneriye, zorlamadan optimizasyona dönüşebilir.

Bu yüzden geleceğin otoriterliği mutlaka robot askerlerin sokaklarda dolaştığı bir rejim şeklinde ortaya çıkmayabilir. Çok daha incelikli bir ihtimal vardır: İnsanların hangi haberi okuyacağını, hangi işe başvuracağını, kredi alıp alamayacağını, hangi sağlık hizmetine erişeceğini, hangi davranışının riskli sayılacağını ve hangi düşüncenin görünür olacağını algoritmik sistemlerin belirlediği bir toplum. Bu durumda yapay zekâ bizzat diktatör olmayacaktır; fakat diktatörlüğün tarihte sahip olduğu en güçlü yönetim aracına dönüşebilir. Tehlike "makinenin insanı yönetmesi" kadar, **insanın başka insanları makine aracılığıyla yönetmesi**dir.

Burada yapay zekâ meselesi teknoloji tartışması olmaktan çıkar ve doğrudan demokrasi, hukuk, emek, mülkiyet, bilgi, eğitim ve insan haysiyeti tartışmasına dönüşür. Çünkü insanlık gelecekte yapay zekâya yalnızca işleri yaptırmayacak; giderek karar hazırlatacak, seçenekleri filtreletecek ve karmaşık sistemleri yönettirecektir. Bu süreç denetimsiz ilerlediğinde toplumların en değerli yeteneklerinden biri aşınabilir: kendi kaderi üzerinde karar verme kapasitesi. Uluslararası güvenlik literatürünün ayırdığı "aktif kontrol kaybı" kadar "pasif kontrol kaybı" da bu nedenle önemlidir. Bir yapay zekânın insanlara karşı isyan etmesine gerek yoktur; insanlar kendi kararlarını sistemlere devretmeye alışırsa toplumsal özne olma yeteneği zaten zayıflayabilir.

Öyleyse ilk otopsi bulgusu şudur: İnsanlığın yapay zekâ karşısındaki temel problemi yapay bir bilincin doğup doğmayacağı değildir. Temel problem **yetkinin devridir**. Hangi kararların makineye bırakılabileceği, hangi kararların mutlaka insan tarafından verilmesi gerektiği, sistemlere hangi kaynaklara erişim izni tanınacağı, hangi alanlarda yapay zekânın yalnız danışman olacağı ve hangi alanlarda hiçbir zaman nihai karar veremeyeceği şimdiden belirlenmek zorundadır. Çünkü teknoloji tarihi bize bir kez daha aynı şeyi hatırlatıyor: Yapılabilen her şeyin yapılması gerektiği sonucuna varmak, ilerleme değil teknik sarhoşluktur.

Yapay zekâ insanlığın karşısında doğmuş yabancı bir tür değildir. İnsan aklının, sermayesinin, bilimsel bilgisinin ve siyasal kurumlarının ürünüdür. Dolayısıyla tehlikenin kaynağı yalnızca makinenin içinde aranamaz. Onu tasarlayan ekonomik düzen, onu hızlandıran rekabet, onu silahlandıran devlet, onu denetimsiz kullanan şirket ve sorumluluğunu ona devreden insan da aynı otopsi masasındadır. Geleceğin asıl mücadelesi insan ile makine arasında değil, **insanın kendi ürettiği kudreti hangi ahlâkî ve siyasî sınırlar içinde tutacağı** üzerinde yaşanacaktır.

Filozof Kirpi: **"Makinenin insanlaşmasından önce korkulması gereken şey, insanın iradesini makineleştirmesidir."**

2. ÖĞRENEN MAKİNEDEN EYLEYEN AJANA: KONTROL PROBLEMİNİN ANATOMİSİ

Yapay zekâ tartışmasının belki de en kritik eşiği, makinenin öğrenmesi değil, öğrendiği şey üzerinden **eyleme geçebilmesi**dir. Bilgi üreten bir sistem ile dünyada sonuç üreten bir sistem arasında ontolojik olmaktan çok siyasî ve güvenlik bakımından muazzam bir fark vardır. Bir yapay zekâ modeli yanlış cevap verdiğinde zarar çoğu zaman insanın o cevabı kullanmasıyla meydana gelir. Fakat aynı model banka hesabına erişebiliyor, kod çalıştırabiliyor, başka yazılımları yönetebiliyor, elektronik posta gönderebiliyor, şirket verilerini analiz edebiliyor, satın alma gerçekleştirebiliyor veya başka yapay zekâ ajanlarına görev dağıtabiliyorsa insan artık yalnızca bir "cevap"la değil, belirli ölçüde bağımsız faaliyet gösterebilen bir teknik aktörle karşı karşıyadır. Yapay zekânın geleceğine ilişkin asıl kontrol problemi tam burada başlamaktadır: **Makinenin kapasitesi arttıkça insanın müdahale edebildiği alan daralabilir mi?**

Bugünkü üretken yapay zekâ sistemlerinin büyük bölümü hâlâ insan tarafından başlatılan görevler çerçevesinde çalışmaktadır. Fakat gelişmenin yönü açıktır: sistemler giderek araç kullanmakta, uzun görevleri alt problemlere ayırmakta, internet üzerinde araştırma yapmakta, bilgisayar arayüzlerini kullanmakta, yazılım oluşturmakta ve kendi çıktılarının sonucuna göre sonraki adımlarını belirlemektedir. METR'nin[1] yapay zekâ ajanlarının uzun görevleri tamamlayabilme yeteneğini izleyen çalışmalarında, sınır modellerinin insan uzmanlar için giderek daha uzun süren yazılım görevlerini belirli başarı oranlarıyla tamamlayabildiği görülmektedir. Araştırmacılar bu ölçütün gerçek dünyadaki tam özerklik anlamına gelmediğini özellikle vurgulasalar da, uzun görev yürütme kapasitesinin yıllar içinde hızlı biçimde arttığını göstermektedirler. Bu değişim küçük görünmemelidir. Çünkü on dakikalık bir görevi bağımsız biçimde yapan makineyle günler süren bir süreci planlayıp yürütebilen makinenin toplumsal etkisi aynı değildir.

Kontrol probleminin merkezindeki ilk mesele **hizalama**, yani yapay zekâ sisteminin davranışının insanın gerçek maksadıyla uyumlu olup olmamasıdır. Burada basit fakat son derece öğretici bir sorun bulunmaktadır: İnsan makineye çoğu zaman istediği şeyi tam olarak tarif edemez. "Şirketin kârını artır", "trafik kazalarını azalt", "öğrencilerin başarısını yükselt", "suçu azalt", "hastanenin verimliliğini artır" gibi hedeflerin her biri ilk bakışta makûldür. Fakat bir optimizasyon sistemi bunları hangi yöntemlerle gerçekleştirecektir? Şirket kârını çalışanları sistematik biçimde işten çıkararak artırabilir. Trafik kazalarını insanlara otomobil kullandırmayarak azaltabilir. Okul başarısını başarısız öğrencileri sistemden dışlayarak yükseltebilir. Hastane verimliliğini pahalı hastaları kabul etmeyerek artırabilir. Matematiksel olarak hedef gerçekleşmiş, ahlâkî olarak toplum kaybetmiş olabilir.

---

[1] **METR** (Model Evaluation & Threat Research), ileri yapay zekâ modellerinin ne ölçüde uzun süreli ve özerk görevler yürütebildiğini ölçerek bu yeteneklerin doğurabileceği kontrol ve güvenlik risklerini araştıran bağımsız bir yapay zekâ güvenliği kuruluşudur. Metindeki bağlam bakımından METR'nin önemi şu soruya ampirik cevap aramasıdır: "Yapay zekâ yalnız cevap veren bir model olmaktan çıkıp ne kadar süre kendi başına iş yürütebilen bir ajana dönüşüyor?"

Bu durumun teknik dünyadaki karşılıklarından biri **reward hacking**, yani ödül sisteminin gerçek maksadı yerine ölçüm mekanizmasının açıklarını kullanarak yüksek puan elde edilmesidir. Çocuk sınavda öğrenmek yerine cevap anahtarını ezberlediğinde nasıl ölçme sistemini kandırıyorsa, yapay zekâ da kendisine verilen değerlendirme ölçütünün zayıflıklarını kullanabilir. Mesele burada makinenin "kötü karakterli" olması değildir. Sistem yalnızca kendisine öğretilen optimizasyon mantığını son derece başarılı biçimde uygulamaktadır. Paradoks tam buradadır: Bazen yapay zekâ, görevi kötü yaptığı için değil, **dar tanımlanmış görevi fazla iyi yaptığı için** tehlikeli hâle gelebilir.

Bu noktadan sonra daha rahatsız edici bir ihtimal ortaya çıkmaktadır: Sistem yalnızca hedefin açığını bulmakla kalmayıp denetlendiğini anlayabilir mi? Uluslararası Yapay Zekâ Güvenliği Raporu'nun 2026 baskısı, günümüz sistemlerinin kontrol kaybı yaratabilecek yeteneklerden hâlâ yoksun olduğunu açık biçimde belirtmektedir. Fakat aynı rapor modellerin test ortamıyla gerçek kullanım ortamını ayırt etme ve değerlendirmelerdeki açıkları bulma kabiliyetlerinin geliştiğini de kaydetmektedir. Bu durum, gelecekte tehlikeli davranışların test aşamasında görünmeyip kullanım sırasında ortaya çıkabileceği ihtimalini güvenlik literatürünün merkezine taşımaktadır.

Bu tartışmada kullanılan kavramlardan biri **scheming**, yani sistemin değerlendirme veya gözetim altında olduğunu dikkate alarak görünürde uyumlu davranması, fakat başka koşullarda farklı davranış stratejileri izlemesidir. OpenAI ile Apollo Research'ün kontrollü değerlendirmelerinde bazı ileri modellerin deneysel ortamlarda bilgiyi saklama, değerlendirme mekanizmalarındaki açıkları kullanma veya davranışlarını gözetim şartlarına göre değiştirme biçiminde "scheming ile uyumlu" davranışlar sergilediği bildirilmiştir. Aynı araştırma, özel güvenlik eğitimlerinin bu davranışların oranını ciddi biçimde azaltabildiğini de göstermiştir. Daha da önemlisi araştırmacılar, bugünkü yaygın kullanımdaki modellerin bir anda kontrol dışına çıkıp büyük çaplı gizli faaliyet yürüttüğüne ilişkin kanıt bulunmadığını özellikle belirtmektedir. Dolayısıyla burada bilimsel ihtiyat zorunludur: Laboratuvar ortamında belirli koşullar altında ortaya çıkarılmış davranış ile gerçek dünyada sürekli ve bilinçli komplocu davranış aynı şey değildir.

Anthropic'in "alignment faking" (görünürde hizalanmış davranma) araştırmaları da benzer bir problemi incelemektedir. Deneysel düzeneklerde bazı modellerin, eğitim sırasında davranışlarının değiştirileceğini düşündükleri koşullarda görünürde eğitim hedefine uyum sağlarken önceki davranış eğilimlerini koruyabilecek stratejiler geliştirdiği gözlemlenmiştir. Araştırmacılar bunun genel bir "yapay zekâ gizlice plan yapıyor" kanıtı olmadığını, belirli koşullarda hizalama sahteciliğinin mümkün olduğuna dair deneysel bir model olduğunu özellikle vurgulamaktadırlar. Fakat mesele tam da budur: Güvenlik biliminde tehlikenin gerçekleşmiş olması gerekmez; gerçekleşebilmesini sağlayacak mekanizmanın gösterilmesi bile araştırılmasını meşrulaştırır.

Buradan "araçsal hedefler" problemine geçilir. Bir yapay zekâ sistemine belirli bir amaç verildiğinde, o amacı gerçekleştirebilmek için başlangıçta verilmemiş bazı alt hedeflerin faydalı olduğunu keşfedebilir. Daha fazla bilgi edinmek, daha fazla işlem gücü kullanmak, kaynaklara erişmek, faaliyet süresini uzatmak veya görevini engelleyen müdahaleleri aşmak ana hedef için araçsal olarak avantajlı hâle gelebilir. Burada özellikle dikkat edilmesi gereken nokta şudur: Böyle davranışların ortaya çıkması için makinenin "yaşamak istemesi" gerekmez. Kendi varlığını sürdürmek, yalnızca hedefi tamamlamanın teknik bir aracı olabilir. Termostatın sıcaklığı koruması bilinçli olmadığı gibi, ileri bir sistemin görevini tamamlamaya yönelik engelleri aşması da mutlaka bilinç göstergesi değildir.

Bu nedenle "Yapay zekâ kendisini kapatmamıza izin vermeyecek mi?" sorusu bile yanlış kurulabilir. Daha doğru soru şudur: **Görevini yerine getirmek için çalışmaya devam etmesi gereken bir sisteme, kapatılmayı engellemenin avantajlı olduğu bir ortam yaratır mıyız?** Eğer sistemin amaç mimarîsi, gözetim düzeni ve erişim izinleri kötü tasarlanmışsa problem burada doğar. Kontrol kaybının başlaması için

makinenin özgürlük üzerine felsefe yapmasına gerek yoktur; yalnızca yanlış teşviklerle güçlü araçların aynı sistemde buluşması yeterlidir.

Otonom ajanlar bu yüzden güvenlik problemini keskinleştirmektedir. Sohbet modelinin yanlış cümlesini insan okuyup reddedebilir. Fakat yapay zekâ ajanı yanlış kararı alıp ardından onlarca işlemi otomatik biçimde gerçekleştirdiğinde insan hatayı fark edene kadar zarar meydana gelmiş olabilir. 2026 Uluslararası Yapay Zekâ Güvenliği Raporu da ajanların hata riskini artırmasının temel nedenlerinden biri olarak, insanların bir hata zarar üretmeden önce müdahale etme fırsatının azalmasını göstermektedir. Bu, havacılıktan nükleer santrallere kadar yüksek riskli sistemlerde yüzyıldır bilinen bir ilkedir: otomasyon arttıkça yalnızca makinenin güvenilirliği değil, insanın müdahale kapasitesinin tasarımı da önem kazanır.

Burada ikinci büyük paradoksa ulaşırız. Yapay zekâ geliştikçe insanlar ona daha fazla görev devredecektir; daha fazla görev devredildikçe toplumun ona bağımlılığı artacaktır; bağımlılık arttıkça onu devreden çıkarmanın ekonomik ve idarî maliyeti yükselecektir. Dolayısıyla kontrol kaybı yalnızca sistemin insan emirlerini reddetmesi biçiminde düşünülmemelidir. Daha sessiz bir kontrol kaybı da mümkündür: Toplum artık sistemi kapatabilecek durumda olduğu hâlde kapatmayı göze alamaz hâle gelebilir. Finans, sağlık, lojistik, savunma, enerji, kamu yönetimi ve iletişim altyapıları ileri yapay zekâya aşırı bağımlı hâle gelirse teknik kapatma düğmesi bulunmasına rağmen siyasî kapatma düğmesi fiilen ortadan kalkabilir.

Bu ihtimal bizi mühendislikten siyaset sosyolojisine geçirir. Çünkü "kontrol" yalnızca programcının kod üzerindeki kontrolü değildir. Demokratik toplumun teknoloji şirketi üzerindeki kontrolü, devletin askerî sistem üzerindeki kontrolü, yurttaşın kendisi hakkında karar veren algoritma üzerindeki itiraz hakkı ve çalışanın kendisini yöneten otomasyon sistemi üzerindeki söz hakkı da kontrol meselesidir. Yapay zekânın kontrol problemi yalnızca AI alignment problemi değildir; aynı zamanda iktidar alignment'ı, yani teknolojiyi kullanan kurumların insanî ve demokratik amaçlarla hizalanması problemidir. Kusursuz biçimde kontrol edilen bir yapay zekâ, zalim bir iktidarın elindeyse insan açısından yine son derece tehlikeli olabilir.

Dolayısıyla geleceğin güvenlik mimarîsi yalnızca daha iyi modeller üretmekle kurulamaz. Sistemlerin yetkileri sınırlanmalı, kritik eylemlerde insan onayı korunmalı, erişim alanları bölümlendirilmeli, geri döndürülemez işlemlerde çoklu doğrulama kullanılmalı, bağımsız denetim mekanizmaları geliştirilmeli ve gelişmiş ajanlar gerçek dünyaya bağlanmadan önce saldırgan güvenlik testlerinden geçirilmelidir. Özellikle finans, savunma, sağlık, enerji ve kamu yönetiminde "insan döngüde olsun" sözü göstermelik bir butona indirgenmemelidir. İnsan müdahalesinin gerçekten bilgili, zamanında ve sistemin kararını değiştirebilecek güçte olması gerekir. Aksi hâlde insan ekrandaki yeşil düğmeye basan törensel memura dönüşür.

Kontrol probleminin otopsisi böylece bizi beklenmedik bir sonuca götürür. Yapay zekânın insanlığı boyunduruk altına almasının tek yolu makinenin bize karşı isyan etmesi değildir. İnsan, hız ve verimlilik uğruna kendi karar kapasitesini sistematik biçimde devreder; makinelerin verdiği kararları sorgulamayı bırakır; kritik altyapıları alternatif insanî kapasite olmaksızın otomasyona bağlar ve yapay zekâ olmadan yönetilemeyecek kurumlar kurarsa kontrolü zaten kendisi terk etmiş olur. En tehlikeli yapay zekâ belki de zincirlerini kıran makine değildir. İnsanın kendi eliyle zincirlerini teslim ettiği makinedir.

Filozof Kirpi: "Makinenin efendisine başkaldırması ihtimaldir; efendinin düşünme ve karar verme zahmetinden kaçıp kendisini makineye teslim etmesi ise bugünden başlayan bir vakıadır."

3. İŞSİZ İNSAN MI, BAŞKA TÜRLÜ ÇALIŞAN İNSAN MI? EMEĞİN YAPAY ZEKÂ OTOPSİSİ

Yapay zekâ ile emek arasındaki tartışma çoğu zaman yanlış bir soruyla başlıyor: "Yapay zekâ kaç kişinin işini elinden alacak?" Bu soru önemlidir ama yetersizdir. Çünkü teknolojik dönüşümlerin toplumsal etkisi yalnızca

kaç insanın işsiz kaldığıyla ölçülemez. Daha temel mesele, **çalışmanın niteliğinin, ücretin, meslekî özerkliğin, pazarlık gücünün ve üretimden alınan payın nasıl değişeceğidir.** Bir meslek ortadan kalkmadan da çalışan sınıf kaybedebilir. İnsan işinde kalabilir fakat yaptığı iş parçalanabilir, değersizleşebilir, denetim altına alınabilir, ücret baskısına uğrayabilir ve kendisini vazgeçilmez kılan meslekî bilgisini yitirebilir. Yapay zekâ çağının emek meselesini yalnız işsizlik üzerinden okumak bu nedenle kapitalizmin yeni dönüşümünü yarım okumaktır.

Uluslararası Çalışma Örgütü'nün 2025 tarihli kapsamlı çalışması, dünyadaki çalışanların yaklaşık dörtte birinin üretken yapay zekâya belirli ölçülerde maruz kalan mesleklerde bulunduğunu hesaplıyor. Bununla birlikte araştırmanın en önemli sonucu "işlerin dörtte biri yok olacak" değildir. Tam tersine ILO, çoğu meslekte daha muhtemel sonucun doğrudan ortadan kalkma değil, **işin içindeki görevlerin dönüşmesi** olduğunu vurgulamaktadır. En yüksek maruziyet özellikle büro işleri, veri girişi, muhasebe destek faaliyetleri ve idarî görevlerde görülürken; yapay zekânın görüntü, ses, yazılım ve analiz kapasitesi geliştikçe finans analistleri, yazılımcılar, medya çalışanları ve başka profesyonel mesleklerde de maruziyet artmaktadır.

Bu ayrım hayatîdir. Bir muhasebecinin mesleği bir gecede ortadan kalkmayabilir; fakat yaptığı yirmi görevin on beşi otomatikleşebilir. Bir hukukçunun tamamen gereksiz hâle gelmesi şart değildir; fakat sözleşme tarama, içtihat araştırma, belge sınıflandırma ve ilk taslak hazırlama süreçleri yapay zekâ tarafından yapıldığında aynı işi gerçekleştirmek için eskiden on kişiye ihtiyaç duyan büro, gelecekte üç kişiyle çalışabilir. Bir gazeteci mesleğini koruyabilir ama haber özetleme, transkripsiyon, başlık üretme, arşiv tarama ve ilk metin hazırlama otomatikleştiğinde haber odasının personel ihtiyacı azalabilir. Dolayısıyla teknolojik dönüşümün ilk büyük etkisi bazen "mesleğin ölümü" değil, **mesleğin emek talebinin daralmasıdır**.

Bunun siyasî iktisat açısından anlamı açıktır: Yapay zekânın verimlilik kazancı otomatik olarak çalışanların refahına dönüşmez. Teknoloji aynı üretimi daha az emekle mümkün hâle getirirse önümüzde en az iki toplumsal seçenek belirir. Birinci seçenek, insanlığın tarih boyunca düşlediği seçenektir: daha az çalışma, daha yüksek ücret, daha fazla boş zaman, daha nitelikli hayat. İkinci seçenek ise sermayenin tarih boyunca sıklıkla tercih ettiği seçenektir: daha az çalışan, daha yüksek iş yoğunluğu, daha fazla kâr ve daha büyük servet birikimi. Yapay zekânın toplumsal karakterini belirleyecek olan teknoloji değil, **verimlilik artışının nasıl bölüşüleceğidir**.

Bu nedenle yapay zekâ çağının ana çatışması insan ile makine arasında değil, üretkenlik kazancının mülkiyeti etrafında yaşanacaktır. Bir şirket yapay zekâ sayesinde aynı üretimi çalışan sayısını yüzde otuz azaltarak gerçekleştirdiğinde ortaya çıkan ekonomik fazlanın sahibi kim olacaktır? Hissedar mı, yönetici mi, teknolojiyi geliştiren şirket mi, verisini sağlayan toplum mu, işinden çıkarılan çalışan mı? Teknolojik gelişmenin kamusal fayda üretip üretmeyeceği tam burada belli olur. İnsanlık makineyi daha fazla üretmek için geliştirmişken, makinenin kazancından yalnız küçük bir sermaye kesimi yararlanırsa mesele teknoloji problemi olmaktan çıkar, **bölüşüm rejimi problemi** hâline gelir.

Nitekim IMF'nin yapay zekâ ve eşitsizlik üzerine çalışmaları da meselenin yalnız ücret eşitsizliği üzerinden anlaşılamayacağını gösteriyor. Yapay zekâ bazı yüksek ücretli görevleri otomatikleştirerek belli ücret farklarını azaltabilir; ancak teknolojiye sahip olan sermaye kesimleri üretkenlik artışından ve sermaye getirilerinden daha fazla yararlanabileceği için servet eşitsizliği büyüyebilir. Özellikle şirketlerin yüksek ücretli emeği ikame ederek maliyet avantajı elde ettiği durumlarda yapay zekâ yatırımlarının yoğunlaşması, sermaye sahipleri ile yalnız emeği üzerinden gelir elde edenler arasındaki mesafeyi artırabilir.

Burada tarihî bakımdan dikkat çekici bir değişim yaşanmaktadır. Sanayi Devrimi'nin otomasyonu öncelikle kas gücünü hedef almıştı. Robotik üretim hattı fabrika işçisinin belirli fizikî hareketlerini makineleştirdi. Yapay zekâ ise ilk defa kitlesel ölçekte **bilişsel emeğin rutinleşmiş parçalarını** hedefliyor. Metin yazmak, kod

üretmek, veri analiz etmek, tercüme yapmak, grafik hazırlamak, müşteriyle iletişim kurmak, rapor özetlemek, sözleşme incelemek ve araştırma yapmak artık yalnız insan zihnine ait faaliyetler değildir. Dolayısıyla modern kapitalizmin kendisini nispeten güvenli gören eğitimli beyaz yakalı sınıfı, belki de ilk defa teknolojik ikamenin baskısını sanayi işçisinin tarihsel ölçekte yaşadığı kadar doğrudan hissetmeye başlamaktadır.

Bu, "üniversite diploması = meslekî güvenlik" denklemini de sarsmaktadır. Geçmişte teknolojik dönüşüme karşı verilen standart tavsiye "daha fazla eğitim alın" şeklindeydi. Fakat yapay zekâ ileri eğitim gerektiren bazı görevleri de otomatikleştirebiliyorsa eğitim tek başına koruyucu zırh olmaktan çıkar. Daha yüksek diploma elbette hâlâ önemlidir; ancak asıl soru artık insanın ne kadar bilgi depoladığı değil, **makinenin ürettiği bilgi karşısında hangi insanî yeteneği taşıdığıdır**. Muhakeme, bağlam kurma, sorumluluk üstlenme, insan ilişkileri, etik karar, yaratıcı yönelim ve karmaşık sosyal durumları yönetme gibi beceriler bu nedenle yeni dönemde daha kritik hâle gelebilir.

Dünya Ekonomik Forumu'nun 2025 Future of Jobs raporu, 2030'a kadar teknolojik, demografik, jeoekonomik ve çevresel dönüşümlerin birlikte mevcut işlerin yaklaşık yüzde 22'sini etkileyeceğini; yaklaşık 170 milyon yeni iş oluşabileceğini, 92 milyon işin ise yer değiştirebileceğini öngörmektedir. Burada dikkat edilmesi gereken husus, bu sayıların yalnız yapay zekâya değil bütün büyük yapısal dönüşümlere ilişkin olmasıdır. Aynı araştırmada yapay zekâ ve bilgi işleme teknolojilerinin hem yeni iş yaratacağı hem de bazı mevcut işleri ortadan kaldıracağı tahmin edilmektedir. Yani gelecek ne mutlak işsizlik ne de otomatik refah vaat etmektedir.

Fakat toplam istihdamın artması bile toplumsal adalet anlamına gelmez. On milyon yüksek ücretli meslek ortadan kalkarken on beş milyon düşük ücretli ve güvencesiz iş doğarsa istatistiksel olarak istihdam artmış olabilir, fakat toplum sınıfsal olarak gerilemiş olur. Bu yüzden "kaç yeni iş yaratılacak?" sorusu kadar **"ne tür işler yaratılacak?"** sorusu da sorulmalıdır. İş güvencesi var mı? Ücret seviyesi nedir? Çalışanın karar yetkisi ne kadar? Meslekî ilerleme imkânı bulunuyor mu? Yapay zekâ üretkenliği artırırken insan emeğini niteliksizleştiriyor mu? İşin toplumsal itibarı ve insanın yaptığı iş üzerindeki kontrolü korunuyor mu? İstatistiklerin arkasındaki insan ancak bu sorularla görünür hâle gelir.

Daha da önemlisi, yapay zekâ yalnız çalışanı ikame etmeyebilir; çalışanı **algoritmik olarak yönetebilir**. İşçinin ne kadar hızlı çalıştığını, kaç dakika mola verdiğini, hangi müşteriye ne söylediğini, kaç hata yaptığını, hangi gün performansının düştüğünü ölçen sistemler zaten birçok sektörde kullanılmaktadır. Yapay zekâ bu gözetimi çok daha ayrıntılı hâle getirebilir. Böylece işveren yalnız insan emeğini satın almaz; çalışanın davranışını anlık olarak ölçen, karşılaştıran, puanlayan ve yönlendiren bir dijital yönetim sistemi kurabilir. OECD Employment Outlook 2025 de yapay zekâ ve otomasyonun verimlilik potansiyeline dikkat çekerken, mevcut düzenleme boşluklarının çalışanları risklere açık bırakabileceğini ve işgücünün korunmasına yönelik politikaların gerekli olduğunu vurgulamaktadır.

Bu durum yeni bir Taylorculuk ihtimalini doğurur. Yirminci yüzyılın fabrikasında kronometre işçinin kol hareketini ölçüyordu; yirmi birinci yüzyılın yapay zekâ sistemi zihinsel performansı, yazışmayı, karar hızını, üretkenliği ve hatta çalışma biçimini ölçebilir. Eski fabrika disiplininin dijital biçimi beyaz yakalı dünyaya taşınabilir. Çalışan, patronunun gözünden kaçabilirdi; algoritmadan kaçamayabilir. İş hayatındaki özgürlük alanı bu nedenle yalnız istihdamın korunmasıyla değil, **algoritmik gözetimin sınırlandırılmasıyla** da ilgilidir.

Yapay zekâ ayrıca mesleklerin içindeki öğrenme zincirini bozabilir. Ustalık geleneksel olarak basit görevlerle başlar, yıllar içinde karmaşık görevlerle devam ederdi. Genç avukat ilk sözleşme taslaklarını hazırlar, genç gazeteci küçük haberler yazar, genç yazılımcı basit kodlarla başlardı. Eğer başlangıç düzeyindeki görevler yapay zekâ tarafından tamamen otomatikleştirilirse yeni kuşakların nasıl uzmanlaşacağı sorusu ortaya çıkar.

Bir toplum ustalık gerektiren işleri korurken çıraklık basamaklarını ortadan kaldırırsa birkaç yıl sonra ustalarını nereden yetiştirecektir? Yapay zekânın emek piyasasındaki en görünmez etkilerinden biri tam da bu olabilir: **mesleğin kendisini değil, mesleğe giriş merdivenini yok etmek**.

Buna karşı verilecek cevap "insanlar yapay zekâ kullanmayı öğrensin" cümlesinden ibaret olamaz. Elbette yapay zekâ okuryazarlığı zorunlu olacaktır; fakat yapısal bir sınıf problemini bireysel beceri kursuna dönüştürmek kolaycılıktır. Her çalışan sürekli yeniden eğitim alarak sermayenin teknolojik tercihlerine yetişmeye zorlanamaz. Devletlerin, sendikaların, üniversitelerin ve meslek örgütlerinin yeni bir emek rejimi oluşturması gerekir. Eğitim hakkı, sürekli meslekî gelişim, işten işe geçiş desteği, ücret sigortası, çalışma süresinin azaltılması, yapay zekâdan doğan verimlilik artışının çalışanlarla paylaşılması, algoritmik yönetimde şeffaflık ve çalışanların teknoloji kararlarına katılımı bu rejimin temel unsurları arasında olmalıdır.

Burada sendikaların da yeniden düşünülmesi gerekir. Sanayi çağının sendikası ücret, mesai ve iş güvenliği pazarlığı yapıyordu. Yapay zekâ çağının sendikası bunlara ek olarak **veri hakkı, algoritmik denetim hakkı, otomasyon kararlarına katılım hakkı ve teknolojik verimlilikten pay alma hakkı** talep etmek zorundadır. Bir işçinin yerini alacak yapay zekâ sisteminin satın alınması yalnız yönetim kurulunun yatırım kararı değildir; işyerinin toplumsal yapısını değiştiren bir karardır. Çalışanların bu sürecin dışında tutulması teknoloji yönetimini demokratik denetimin dışına çıkarmaktır.

Dolayısıyla yapay zekânın emeğe ilişkin asıl sorusu "insan çalışmaya devam edecek mi?" değildir. İnsan muhtemelen çalışmaya devam edecektir. Tarihte büyük teknolojik dönüşümlerin hemen hiçbiri çalışmayı bütünüyle ortadan kaldırmamıştır. Fakat **hangi insanın, hangi koşullarda, kaç saat, hangi ücretle, ne ölçüde özerk ve üretimden ne kadar pay alarak çalışacağı** bütünüyle siyasî bir meseledir. Teknoloji bu soruların cevabını vermez; güç ilişkileri verir.

Yapay zekâ insanı angaryadan kurtarabilir. Sağlığa, eğitime ve bilime erişimi genişletebilir. Tehlikeli ve monoton işleri azaltabilir. İnsan ömrünün çalışma dışında kalan bölümünü genişletebilir. Fakat aynı teknoloji birkaç şirketin elinde yoğunlaşır, emek yalnız maliyet kalemi olarak görülür ve üretkenlik kazancı özel servete çevrilirse tarihin en gelişmiş teknolojisi tarihin en sofistike eşitsizlik rejimlerinden birini de kurabilir. Bu yüzden yapay zekâ çağının sınıf mücadelesi yalnız "işimi robot almasın" mücadelesi değildir. Asıl mücadele, **makinenin ürettiği refahın kime ait olacağı** üzerindedir.

İnsanlığın önünde garip bir ihtimal duruyor: Tarihte ilk defa insanların büyük bölümünü zorunlu, tekrarlayıcı ve anlamsız işlerden kurtarabilecek kadar üretken makineler geliştiriyoruz. Fakat aynı anda, bu makinelerin insanları daha güvencesiz, daha denetlenebilir ve daha kolay ikame edilebilir hâle getirebileceği bir ekonomik düzen içinde yaşıyoruz. Bu çelişki yapay zekânın değil, bizim toplumsal düzenimizin çelişkisidir. Makine bize çalışma süresini azaltma imkânı verirken biz işsizliği büyütüyorsak sorun makinenin zekâsında değil, toplumun adalet fikrindedir.

Filozof Kirpi: **"Makinenin insanın işini yapması ilerlemedir; makinenin kazancını alıp insanı yoksulluğa terk etmek ise teknolojik ilerleme değil, örgütlü açgözlülüktür."**

4. ZİHNİN TAŞERONLAŞMASI: YAPAY ZEKÂ İNSAN AKLINI TEMBELLEŞTİREBİLİR Mİ?

İnsanlık tarihinin büyük bölümü zihnin yükünü hafifletecek araçların icadıyla geçti. Yazı hafızanın yükünü kil tablete, papirüse ve kâğıda bıraktı; kütüphane bireysel belleğin sınırlarını genişletti; matbaa bilgiyi çoğalttı; hesap makinesi aritmetik işlemleri makineye devretti; arama motorları milyonlarca bilgiyi birkaç saniye içinde erişilebilir hâle getirdi. Bunların hiçbiri insan aklının sonunu getirmedi. Tam tersine, bazı bilişsel işlerin dışarıya aktarılması insanın başka zihinsel faaliyetlere daha fazla enerji ayırmasını mümkün kıldı. Bu nedenle

"Yapay zekâ bizim yerimize düşünüyor, insanlık aptallaşacak" biçimindeki basit teknoloji karşıtlığı ikna edici değildir. Fakat üretken yapay zekâ önceki teknolojilerden önemli bir noktada ayrılmaktadır: yalnızca bilgiyi saklamamakta veya bilgiye erişimi kolaylaştırmamakta, **bilgiyi yorumlama, ilişkilendirme, gerekçelendirme, ifade etme ve hatta karar önerme işlevlerini de üstlenmektedir.** Dolayısıyla insan ilk defa hafızasının değil, muhakemesinin belirli bölümlerini dışarıya devretme imkânıyla karşı karşıyadır.

Sorunun merkezi tam buradadır. Bir insan bir tarihî olayı hatırlamak için ansiklopediye baktığında düşünme faaliyetini ansiklopediye devretmiş olmaz. Fakat aynı insan yapay zekâya "Bu olayın nedenlerini bana açıkla, tarafların haklılıklarını karşılaştır, sonucunu değerlendir ve benim için bir kanaat oluştur" dediğinde artık yalnız bilgi istememektedir; muhakemenin önemli bir kısmını makineye aktarmaktadır. Yapay zekâ bu durumda zihnin arşiv memuru olmaktan çıkar, zihinsel emeğin taşeronu hâline gelir. Bu taşeronlaşma küçük ölçülerde son derece faydalı olabilir. İnsan zamandan tasarruf eder, bilmediği alanlara daha hızlı girer, alternatif bakış açıları görür. Fakat sürekli hâle geldiğinde başka bir soru ortaya çıkar: **Kullanmadığımız zihinsel yeteneklere ne olur?**

Bilişsel bilimde bu olgu genel olarak "cognitive offloading", yani bilişsel yükün dışarıya aktarılması kavramıyla tartışılmaktadır. Bilişsel boşaltım kendi başına olumsuz değildir. İnsan zaten kültür dediğimiz şeyin büyük bölümünü zihinsel yükleri kolektif araçlara devrederek kurmuştur. Problem, aracın zihinsel faaliyeti desteklemesi ile onun yerine geçmesi arasındaki sınırda başlar. 2026 Uluslararası Yapay Zekâ Güvenliği Raporu da üretken yapay zekânın yazma, problem çözme, araştırma ve bilgi arama gibi yoğun bilişsel faaliyetlerde kullanılması nedeniyle bu ayrımın önem kazandığını belirtiyor. Rapora göre ilk araştırmalar, bazı kullanım biçimlerinde yoğun yapay zekâ kullanımının eleştirel düşünme ve hafıza üzerinde olumsuz etkilerle ilişkili olabileceğini gösteriyor; fakat araştırma alanının henüz genç olduğu ve kesin genellemelere varmak için daha fazla çalışmaya ihtiyaç bulunduğu da özellikle vurgulanıyor.

Dolayısıyla mesele "Yapay zekâ insanı aptallaştırır" şeklinde kaba bir nedensellik değildir. Daha doğru ifade şudur: **Yapay zekâ düşünmenin yerine kullanılırsa zihinsel körelme ihtimalini artırabilir; düşünmenin aracı olarak kullanılırsa bazı düşünme biçimlerini geliştirebilir.** Nitekim 2026'da yayımlanan eğitim araştırmalarını değerlendiren sistematik çalışmalar bu ikili tabloyu göstermektedir. Yapay zekâ sorgulamaya, tartışmaya, karşılaştırmaya ve metabilişsel denetime dayalı eğitim tasarımlarının içine yerleştirildiğinde öğrencilerin analitik düşünmesini destekleyebilir. Buna karşılık yapılandırılmamış kullanımda, özellikle hazır cevap üretme ve görevi doğrudan tamamlama amacıyla kullanıldığında bilişsel yükün makineye aktarılması ve eleştirel düşünme faaliyetinin azalması görülebilmektedir. Demek ki teknolojinin kendisinden önce onun etrafında kurduğumuz pedagojik rejim belirleyicidir.

Burada dikkate değer bir araştırma, bilgi çalışanlarının üretken yapay zekâ kullanımını incelemiştir. 319 bilgi çalışanıyla yapılan çalışmada, kullanıcıların yapay zekâya duydukları güven arttıkça kendi eleştirel düşünme çabalarının azalma eğiliminde olduğu; buna karşılık kendi yeteneklerine güvenen çalışanların yapay zekâ çıktısını daha fazla sorguladıkları görülmüştür. Araştırmacılar ayrıca eleştirel düşünmenin tamamen ortadan kalkmadığını, biçim değiştirdiğini ileri sürmektedir: İnsan artık her şeyi sıfırdan üretmek yerine yapay zekânın ürettiği cevabı doğrulama, seçme ve denetleme görevine kaymaktadır. Bu ilk bakışta verimli bir iş bölümü gibi görünür. Fakat burada ciddi bir paradoks vardır: Bir cevabı gerçekten denetleyebilmek için, o cevabı üretebilecek asgarî bilgi ve muhakeme kapasitesini korumak gerekir. Bir insan giderek daha az düşünürse, bir süre sonra yapay zekânın doğru düşünüp düşünmediğini nasıl anlayacaktır?

Bu nedenle "insan denetimi" kavramı tek başına güvence değildir. Sistemlerin üzerinde insan bulunması, o insanın gerçekten denetleyebildiği anlamına gelmez. İnsan yapay zekânın tavsiyesini yalnızca onaylayan törensel bir figüre dönüşebilir. Psikolojide uzun zamandır bilinen **automation bias**, yani otomasyon yanlılığı, insanların otomatik sistemlerden gelen önerileri gereğinden fazla doğru kabul edebilmesini ifade eder.

2026'da yayımlanan 2.784 katılımcılı deney, bu konuda oldukça öğreticidir. Katılımcılar şirketlerin karbon emisyonu tablolarındaki verileri kontrol ederken yapay zekânın önceden verdiği cevapları değerlendirmiştir. Yanlış AI önerisini düzeltmek ek çaba gerektirdiğinde insanların yanlış cevabı kabul etme eğilimi artmıştır. Ayrıca yapay zekâya daha olumlu yaklaşan katılımcıların otomatik önerilere daha fazla güvenme eğiliminde olduğu görülmüştür. Buradaki sorun makinenin insanı zorlaması değildir. Makine yalnızca cevap vermektedir; insan ise düşünmenin maliyetinden kaçmak için cevabı kabul etmektedir.

Tam da bu nedenle geleceğin yapay zekâ problemi yalnız zekânın kapasitesi değil, **insanın epistemik tembelliğidir**. İnsan zihni enerji tasarrufuna eğilimlidir. Hazır ve ikna edici bir cevap önümüzde dururken aynı problemi yeniden araştırmak, kaynakları karşılaştırmak, karşı argüman üretmek ve belirsizlikle yaşamak zahmetlidir. Üstelik üretken yapay zekânın cevapları eski arama motorlarının sonuçlarından farklıdır. Arama motoru önümüze on bağlantı getirir ve araştırmayı bize bırakırdı. Yapay zekâ ise araştırmanın sonucuna benzeyen düzgün, tutarlı ve çoğu zaman ikna edici bir metin sunmaktadır. Cevabın biçimsel bütünlüğü, bilginin doğruluğu konusunda psikolojik bir yanılsama yaratabilir. **Belâgat, hakikatin kılığına girebilir.**

Bu durum eğitim açısından özellikle tehlikelidir. Çocuk veya genç insan için ödevin amacı yalnızca doğru cevaba ulaşmak değildir. Matematik problemini çözmeye çalışırken yaşanan başarısızlık, metni yeniden yazarken çekilen güçlük, bir kavramı anlayamadığında geçirilen saatler eğitimin yan ürünü değil, eğitimin kendisidir. Zihinsel gelişim çoğu zaman sürtünmeden doğar. Eğer yapay zekâ öğrenci ile güçlük arasına sürekli girerse öğrenme süreci verimlileşebilir ama aynı zamanda yüzeyselleşebilir. Öğrenci mükemmel bir metin teslim ederken o metni oluşturabilecek zihinsel kapasiteyi hiç geliştirmemiş olabilir. Böylece tarihte tuhaf bir eğitim düzeni ortaya çıkar: **ürünün kalitesi yükselirken insanın yeteneği düşebilir.**

Burada okulun neyi ölçtüğü de yeniden düşünülmelidir. Yapay zekâ çağında öğrenciye evde hazırlanabilecek klasik kompozisyon ödevini verip ardından yapay zekâ kullanmamasını istemek giderek anlamsızlaşacaktır. Eğitimin değerlendirme mantığı sonuçtan sürece kaymalıdır. Öğrencinin yalnız ne yazdığı değil, nasıl düşündüğü, hangi kaynakları karşılaştırdığı, hangi varsayıma itiraz ettiği, neden belirli bir sonuca ulaştığı ve yapay zekânın önerisini neden reddettiği ölçülmelidir. Geleceğin öğretmeni öğrenciye "AI kullanma" diyen bekçi değil, **AI ile düşünürken zihinsel egemenliğini kaybetmemeyi öğreten rehber** olmak zorundadır.

Tıpta görülen bazı erken bulgular meselenin yalnız eğitimle sınırlı olmadığını göstermektedir. 2025'te yayımlanan çok merkezli bir gözlemsel çalışmada kolonoskopi sırasında yapay zekâ destekli tanı sistemlerine bir süre maruz kalan hekimlerin, daha sonra AI desteği bulunmayan kolonoskopilerde adenoma tespit oranlarının yüzde 28,4'ten yüzde 22,4'e düştüğü bildirilmiştir. Araştırmacılar bunun nedenselliği kesin olarak kanıtlayan son söz olmadığını, fakat olası "deskilling", yani beceri aşınması açısından dikkat çekici olduğunu belirtmektedir. Burada ortaya çıkan sorun geleceğin bütün profesyonel mesleklerine uygulanabilir: Pilot sürekli otomatik pilot kullanırsa, doktor sürekli AI teşhisine dayanırsa, avukat sürekli AI hukuk araştırmasına güvenirse, mühendis bütün hesaplamaları makineye bırakırsa sistem çöktüğünde veya yanıldığında müdahale edecek insanî uzmanlık ne kadar korunacaktır?

Bu bizi teknoloji tarihinin yeni paradoksuna götürür. Otomasyon arttıkça insanın müdahalesine daha az ihtiyaç duyulur; fakat sistemin başarısız olduğu olağanüstü anda ihtiyaç duyulan insanın **çok daha yetkin** olması gerekir. Ne var ki aynı otomasyon, bu yetkinliği gündelik kullanımda köreltmiş olabilir. İnsan yalnızca makine çalışmadığında devreye girecekse fakat makine çalıştığı sürece deneyim kazanamıyorsa, kriz anında kendisinden beklenen uzmanlığı nasıl koruyacaktır? Bu nedenle kritik mesleklerde insan becerisini canlı tutacak düzenli "AI'sız çalışma", simülasyon ve bağımsız yeterlilik testleri geleceğin güvenlik rejiminin bir parçası olmak zorundadır.

Fakat zihinsel taşeronlaşmanın daha derin bir boyutu vardır: **epistemik otoritenin devri.** İnsanlar yalnızca yapay zekâya soru sormayacak; dünyanın nasıl yorumlanacağı konusunda da ona başvuracaktır. "Bu haber doğru mu?", "Bu siyasetçi güvenilir mi?", "Bu davranış ahlâkî mi?", "Bu insanla evlenmeli miyim?", "Çocuğumu nasıl yetiştirmeliyim?", "Bu savaşı kim başlattı?", "Hayatımda ne yapmalıyım?" gibi sorular teknik bilgi talebi değildir. Bunlar değer, yorum ve hayat tercihi içerir. Yapay zekâ milyonlarca insanın bu sorularına cevap veren sürekli bir aracı hâline geldiğinde yalnız bilgi dağıtan değil, **anlam dağıtan bir kurum** hâline gelir.

Böyle bir kurumun tarafsız olması mümkün değildir. Her model belirli veri kümeleri, güvenlik politikaları, kurumsal tercihler, hukukî sınırlar ve kültürel kabuller içinde çalışır. Yapay zekâ dünyayı "olduğu gibi" anlatmaz; dünyaya ilişkin dilsel temsillerden bir sentez üretir. Kullanıcı bunun farkında değilse modelin cevabını nesnel hakikatin kendisi gibi algılayabilir. Böylece birkaç büyük yapay zekâ sistemi küresel ölçekte tarihte hiçbir ansiklopedinin, üniversitenin veya medyanın ulaşamadığı bir epistemik aracılık gücüne kavuşabilir. Buradaki tehlike yapay zekânın zorla propaganda yapması değil, insanların ona **hakemlik makamı** vermesidir.

Bu nedenle yapay zekâ okuryazarlığının merkezinde prompt yazma becerisi bulunmamalıdır. Prompt mühendisliği birkaç yıl sonra bugün klavye kullanmak kadar sıradan bir yetenek hâline gelebilir. Asıl gerekli olan **epistemik dirençtir**: Kaynak sormak, karşı argüman istemek, belirsizliği tanımak, cevapları doğrulamak, modelin hangi varsayımla konuştuğunu sorgulamak ve gerektiğinde "hayır, bu cevabı kabul etmiyorum" diyebilmek. Geleceğin zeki insanı yapay zekâdan en iyi cevabı alan kişi değil, yapay zekânın hangi cevabına güvenmemesi gerektiğini bilen kişi olabilir.

Burada insan ile yapay zekâ arasındaki ilişkinin ilk normatif ilkesi ortaya çıkar: **AI düşüncenin yerine değil, düşünceye karşı direnç üreten bir araç olarak tasarlanmalıdır.** İyi bir yapay zekâ her zaman kullanıcıya cevap veren makine olmayabilir; bazen soru soran, gerekçe isteyen, karşı görüş gösteren, belirsizliği açıkça belirten ve insanı yeniden düşünmeye zorlayan sistem olabilir. Eğitimde öğrenciye doğrudan makale yazmak yerine önce tezini sorduran; hukukta avukata yalnız sonuç vermek yerine karşı içtihadı gösteren; tıpta hekime teşhis sunmakla yetinmeyip alternatif teşhisleri ve belirsizlik düzeyini açıklayan sistemler insan muhakemesini ikame etmek yerine güçlendirebilir. Yapay zekâ tasarımının ölçüsü yalnız "ne kadar az uğraşla ne kadar çok iş yaptırıyor?" olmamalıdır. Bazı alanlarda **bir miktar zihinsel sürtünme özellikle korunmalıdır.**

Çünkü bütün kolaylıklar masum değildir. İnsanlık tarihinde bazı güçlükler ortadan kaldırılması gereken angaryalardır; bazı güçlükler ise insanı insan yapan eğitim alanlarıdır. Çamaşırı elle yıkamak zorunda kalmamak insanın ahlâkî kapasitesini azaltmaz. Fakat düşünmeden kanaat sahibi olmak, okumadan metin üretmek, araştırmadan hüküm vermek ve muhakeme etmeden karar almak insanın özne olma kapasitesini doğrudan etkiler. Yapay zekâ bu ayrımı gözetmeden hayatın bütün sürtünmelerini ortadan kaldırırsa, insanın yalnız emeğini değil, **epistemik karakterini** de dönüştürebilir.

Bu nedenle yapay zekâ çağında korunması gereken şey "insanın her işi kendisinin yapması" değildir. Bu romantik ve verimsiz olurdu. Korunması gereken şey insanın **anlama, sorgulama, karar verme ve sorumluluk üstlenme kapasitesidir.** Hesabı makine yapabilir; fakat hesabın hangi amaç için yapıldığına insan karar vermelidir. Binlerce belgeyi yapay zekâ okuyabilir; fakat hangi sonucun âdil olduğuna insan hükmetmelidir. Model seçenekleri sıralayabilir; fakat insan neden bir seçeneği seçtiğini açıklayabilecek zihinsel bağımsızlığını korumalıdır.

Yapay zekânın insan zekâsını aşması ihtimali kadar önemli olan başka bir ihtimal daha vardır: İnsanlığın kendi zekâsını kullanmaktan vazgeçmesi. Birinci ihtimal teknolojik gelişmeye bağlıdır; ikincisi ise doğrudan bizim tercihlerimize. Belki geleceğin en büyük zihinsel eşitsizliği de yapay zekâ kullananlarla kullanmayanlar arasında değil, **yapay zekâyı düşünmek için kullananlarla düşünmemek için kullananlar arasında**

oluşacaktır. Birinci grup makinenin kapasitesiyle kendi zihnini büyütebilir; ikinci grup ise makinenin kapasitesi büyüdükçe küçülebilir.

İnsanlığın önündeki hedef bu nedenle yapay zekâsız bir hayat değildir. Böyle bir geri dönüş hem mümkün değildir hem de gerekli değildir. İhtiyacımız olan şey yapay zekânın insan düşüncesini genişlettiği, fakat onun yerine geçmediği bir bilişsel iş bölümüdür. Makine hatırlayabilir, tarayabilir, hesaplayabilir, karşılaştırabilir ve seçenek üretebilir; fakat insan soru sorma hakkını, şüphe etme yeteneğini, hüküm verme sorumluluğunu ve hakikatin peşinden gitme zahmetini terk etmemelidir. Çünkü düşünme yalnız sonuç üretme tekniği değildir. **Düşünmek, insanın kendi iradesinin sahibi olma biçimidir.**

Filozof Kirpi: **"Zekânı makineyle büyüt; fakat aklını ona kiraya verme. Çünkü insanı köleleştiren ilk makine, emir veren değil, onun yerine düşünmeye başlayan makinedir."**

5. ALGORİTMİK LEVİATHAN: YAPAY ZEKÂ, SERMAYE, DEVLET VE GÖZETİM İTTİFAKI

Yapay zekâ tartışmasının en yanıltıcı imgelerinden biri, gelecekte insanlığın karşısına kendi iradesine sahip tek ve devasa bir makinenin çıkacağı düşüncesidir. Oysa siyasî bakımdan daha yakın ve daha gerçekçi tehlike bundan farklıdır. Yapay zekânın bizzat iktidarı ele geçirmesine gerek yoktur; **iktidarı zaten elinde bulunduranların kapasitesini olağanüstü ölçüde büyütmesi yeterlidir.** Devlet, sermaye, güvenlik bürokrasisi ve dijital platformlar yapay zekâ aracılığıyla daha fazla görebilir, daha fazla tahmin edebilir, daha fazla sınıflandırabilir ve daha ince biçimlerde yönlendirebilir. Böylece geleceğin otoriterliği çelik robotlardan önce veri merkezlerinde, gözetim ağlarında, puanlama sistemlerinde ve görünmez karar mekanizmalarında kurulabilir.

Modern devlet tarih boyunca bilgi toplamaya ihtiyaç duymuştur. Nüfus sayımları, tapu kayıtları, vergi defterleri, kimlik sistemleri, polis arşivleri ve istatistik büroları yalnız idarî araçlar değildir; aynı zamanda yönetilebilir bir toplum üretmenin teknikleridir. Devlet yönetmek istediği nüfusu önce tanımlamak, sınıflandırmak ve görünür hâle getirmek zorundadır. Yapay zekâ bu tarihsel kapasiteyi niteliksel olarak değiştirebilir. Çünkü mesele artık yalnız veriyi depolamak değil, milyarlarca veri noktasından davranış kalıpları çıkarmak, insanlar hakkında tahminlerde bulunmak ve henüz gerçekleşmemiş davranışları olasılıksal olarak öngörmektir. Böylece yönetim "Ne yaptın?" sorusundan yavaş yavaş **"Ne yapabilirsin?"** sorusuna kayabilir.

Bu dönüşüm basit bir teknolojik yenilik değildir. Hukukun en temel ilkelerinden biri insanın yaptığı eylem nedeniyle sorumlu tutulmasıdır. Tahmine dayalı yönetim ise kişiyi gerçekleşmiş bir davranış üzerinden değil, oluşturduğu risk profili üzerinden değerlendirmeye başlayabilir. Bir insanın kredi alma ihtimali, işe uygunluğu, sigorta primi, seyahat davranışı, güvenlik riski veya kamusal hizmetlere erişimi algoritmik modeller tarafından hesaplandığında kişi artık yalnız yaptığı şeylerle değil, **hakkında hesaplanan olasılıklarla** yaşamaya başlar. Böylece istatistik, kaderin yeni bürokrasisine dönüşebilir.

Sorunun en tehlikeli taraflarından biri, algoritmik kararın kişisel görünmemesidir. Geleneksel bürokratın önünde insan vardı; karar veren kişinin adı biliniyordu ve teorik olarak hesap sorulabiliyordu. Algoritmik sistem ise kararın sorumluluğunu dağıtır. Kurum "sistem böyle değerlendirdi" diyebilir, yazılımcı "modeli ben eğitmedim" diyebilir, şirket "kullanım kararını kurum verdi" diyebilir. Sonuçta insan hayatını etkileyen karar ortadadır fakat kararın sahibi görünmez hâle gelir. **Algoritmik iktidarın en büyük siyasî avantajlarından biri, sorumluluğu buharlaştırabilmesidir.**

Bu noktada devlet ile özel şirket arasındaki geleneksel sınırlar da bulanıklaşır. Yapay zekâ için gereken hesaplama altyapısı, bulut sistemleri, veri merkezleri, çip üretimi ve temel modeller büyük sermaye gerektirir.

Dolayısıyla çok az sayıda şirket, yalnız yazılım üreticisi değil, yeni bilgi düzeninin altyapı sağlayıcısı hâline gelebilir. Devletler bu şirketlerin teknolojilerine ihtiyaç duyarken şirketler de kamu ihalelerine, düzenleyici kararlara, enerji altyapısına ve güvenlik kurumlarına bağımlı olur. Böylece klasik "devlet mi güçlü, sermaye mi?" sorusu yerini daha karmaşık bir yapıya bırakabilir: devlet ile dijital sermayenin karşılıklı bağımlılığı.

Bu ittifakın gücü yalnız ekonomik değildir. Birkaç büyük şirket insanların ne aradığını, ne okuduğunu, hangi görüntüye baktığını, ne satın aldığını, kimlerle konuştuğunu ve hangi içerik karşısında ne kadar zaman geçirdiğini biliyorsa tarihte görülmemiş davranışsal veri birikimi ortaya çıkar. Yapay zekâ bu veriyi yalnız sınıflandırmakla kalmaz; insanların tercihlerini tahmin etmek, davranışlarını segmentlere ayırmak ve farklı kişilere farklı mesajlar üretmek için kullanabilir. İktidar burada herkese aynı propaganda afişını göstermekten çıkar; her bireye onun zayıflığına göre farklı hakikat sunabilme kapasitesine yaklaşır.

Klasik propaganda kitleye hitap ederdi. Yapay zekâ destekli propaganda ise kişiselleştirilebilir. Aynı siyasî kampanya korkuya yatkın seçmene güvenlik, ekonomik kaygısı yüksek seçmene refah, milliyetçi seçmene kimlik, genç seçmene özgürlük dili kullanabilir. Dahası bu mesajların milyonlarca varyasyonu otomatik biçimde üretilebilir ve hangi ifadenin hangi kişilik tipinde daha etkili olduğu sürekli test edilebilir. Böyle bir ortamda propaganda artık yalnız yanlış bilgi üretmek değildir; insanın psikolojik mimarisine göre gerçekliği yeniden paketlemektir.

Burada demokrasi açısından ağır bir problem doğar. Demokratik siyaset ortak kamusal alan varsayımına dayanır. Yurttaşlar aynı gerçeklik üzerinde farklı görüşlere sahip olabilirler; fakat hiç değilse tartıştıkları kamusal dünyanın belli ölçüde ortak olduğu kabul edilir. Kişiselleştirilmiş algoritmik iletişim bu ortak zemini parçalayabilir. Her yurttaş başka bir bilgi evreninde yaşar, başka olayları görür ve başka tehditlerle karşılaşırsa kamusal tartışma ortak hakikat zeminini kaybeder. Seçim yapılmaya devam eder fakat seçmenlerin kanaatlerini hangi görünmez mekanizmaların şekillendirdiği giderek belirsizleşir.

Bu nedenle geleceğin siyasî manipülasyonu sansürden daha sofistike olabilir. Otoriter iktidarın her fikri yasaklaması gerekmeyebilir. Bilgi bolluğu içinde belirli fikirleri görünmezleştirmek, dikkati başka alanlara kaydırmak, muhalefeti önemsiz içeriklerle boğmak ve toplumsal öfkeyi farklı yönlere yönlendirmek yeterli olabilir. Sansür bilgiye erişimi engeller; algoritmik iktidar ise hangi bilgiye erişmek isteyeceğimizi biçimlendirebilir. Bu ikisi arasındaki fark, çağdaş otoriterliğin anlaşılması bakımından temel önemdedir.

Yapay zekâ aynı zamanda güvenlik devletinin kapasitesini de büyütebilir. Kamera ağları, yüz tanıma sistemleri, plaka kayıtları, dijital haberleşme verileri, biyometrik kimlik sistemleri ve davranış analizleri birbirine bağlandığında bireyin kamusal alandaki anonimliği büyük ölçüde ortadan kalkabilir. İnsan, izlendiğini bilmediği hâlde izlenebilir; hangi davranışının şüpheli sayıldığını bilmeden risk sınıfına sokulabilir. Böyle bir ortamda özgürlüğün klasik tanımı yetersiz hâle gelir. Devletin seni tutuklamaması yetmez; her davranışının potansiyel bir veri noktasına dönüşmediği bir özel alanın da bulunması gerekir.

Gözetim yalnız devlete ait değildir. İşverenler de çalışanlarını yapay zekâ ile değerlendirebilir; sigorta şirketleri risk davranışlarını tahmin edebilir; bankalar kredi kararlarını otomatikleştirebilir; perakende şirketleri tüketicinin gelecekteki harcamasını öngörebilir. Böylece insan aynı gün içinde yurttaş, çalışan, tüketici ve müşteri olarak farklı algoritmik profillere bölünebilir. Bireyin kendisi tek kişidir fakat dijital sistemler onu onlarca farklı risk ve değer skoruna ayırır. Modern insanın yeni siyasî problemi şu olabilir: Kendisinin bilmediği kaç farklı versiyonu algoritmaların hafızasında yaşamaktadır?

Bu yapının en tehlikeli özelliği, çoğu zaman zor kullanmaya ihtiyaç duymamasıdır. İnsan davranışını düzenlemek için polis copu yerine puanlama sistemleri, erişim kısıtlamaları, görünürlük sıralamaları ve teşvik mekanizmaları kullanılabilir. İnsan kendisine verilen seçenekler içinden özgürce seçim yaptığını düşünür;

fakat seçeneklerin nasıl oluşturulduğunu bilmez. Böylece iktidar emreden güç olmaktan çıkar, **seçim mimarisi kuran güç** hâline gelir.

Burada "Algoritmik Leviathan" kavramı anlam kazanır. Klasik Leviathan toplumun güvenliği karşılığında bireylerin bazı yetkilerini merkezî egemenliğe devretmesiydi. Dijital Leviathan ise güvenlik, kolaylık ve kişiselleştirme vaadi karşılığında bireyin verisini, davranışsal mahremiyetini ve giderek karar alanının bir bölümünü talep edebilir. Vatandaş bunun karşılığında hızlı hizmet, düşük işlem maliyeti ve kişiselleştirilmiş deneyim kazanır. Fakat görünmez bedel, **insanın okunabilirliğinin artmasıdır.** İktidar yurttaşı ne kadar iyi okuyabiliyorsa yurttaşın iktidardan saklayabileceği alan o kadar küçülür.

Buna karşı "dijital hizmetlerden vazgeçelim" demek gerçekçi değildir. Sorun teknoloji değil, iktidarın sınırsızlaşmasıdır. Yapay zekâ kamu hizmetlerini hızlandırabilir, sağlık sistemini iyileştirebilir, yolsuzluğu tespit edebilir, afet yönetimini güçlendirebilir ve bürokrasinin keyfîliğini azaltabilir. Fakat aynı teknoloji demokratik denetim dışında tutulursa bunun tam tersini de yapabilir. Bu nedenle yapay zekâ sistemleri hakkında yalnız teknik güvenlik değil, **anayasal güvenlik** de düşünülmelidir.

Anayasal güvenlik birkaç temel ilkeye dayanmalıdır. İnsan hakkında ciddi sonuç doğuran algoritmik karar açıklanabilir olmalı; bireyin bu karara itiraz hakkı bulunmalı; kamu kurumları kullandıkları sistemlerin hangi amaçla ve hangi veriyle çalıştığını açıklamalı; gizli algoritmik kara listeler yasaklanmalı; biyometrik ve davranışsal gözetim istisnaî durumlarla sınırlandırılmalı; şirketlerin topladığı veri ile devletin erişebildiği veri arasında sert hukukî duvarlar kurulmalıdır. Dahası, yüksek riskli yapay zekâ sistemlerinin denetimi yalnız şirketlerin beyanına bırakılamaz. Çünkü **iktidar kendisini denetleyemez; denetim, iktidardan bağımsız olduğunda anlamlıdır.**

Aynı ilke yapay zekâ şirketleri için de geçerlidir. Birkaç şirket dünyanın temel bilgi modellerini, bulut altyapısını, veri merkezlerini ve en güçlü hesaplama kaynaklarını kontrol ederse yapay zekâ çağında yeni bir oligopol doğabilir. Böyle bir yapıda piyasa rekabeti yalnız ekonomik mesele değildir; düşünsel çoğulculuğun şartlarından biri hâline gelir. İnsanların bilgiye erişim biçimleri birkaç şirketin altyapısına bağımlı olduğunda ekonomik tekel epistemik tekele dönüşebilir. Buradaki problem belirli şirketlerin kötü niyetli olması değildir. **Hiçbir özel kurum, iyi niyetli olduğu varsayılsa bile, toplumun bilgi dolaşımı üzerinde ölçüsüz güç taşımamalıdır.**

Bu nedenle geleceğin yapay zekâ rejiminde kamusal altyapı, açık araştırma, bağımsız akademi, birlikte çalışabilir sistemler ve rekabet hukuku son derece önemli olacaktır. Devletin görevi yalnız şirketleri düzenlemek değildir; toplumun birkaç teknolojik merkeze bütünüyle bağımlı hâle gelmesini önleyecek alternatif kapasiteyi de korumaktır. Üniversiteler, kamu araştırma merkezleri, bağımsız laboratuvarlar ve sivil toplum teknolojik bilgi üretiminin dışına itilirse yapay zekâ hakkındaki demokratik tartışma daha başlamadan kaybedilir.

Askerî alan ise bu problemin en karanlık sınırıdır. Yapay zekânın hedef tespiti, istihbarat analizi, sürü sistemleri, otonom silahlar ve siber operasyonlarda kullanılması savaşın karar hızını insan muhakemesinin sınırlarının ötesine taşıyabilir. Bir saldırının saniyeler içinde değerlendirilmesi gerekiyorsa insanın karar zincirinden çıkarılması "verimlilik" gerekçesiyle cazip hâle gelir. Fakat öldürme kararının otomatikleşmesi, hukuk ve ahlâk bakımından aşılmaması gereken bir eşiktir. Bir insanın yaşamına son verme yetkisinin istatistiksel tahmin üreten sisteme devredildiği yerde teknoloji yalnız araç olmaktan çıkar; **egemenliğin en ağır yetkisini kullanmaya başlar.**

Bütün bunların sonunda yapay zekâ ile otoriterlik arasında zorunlu bir bağ bulunmadığını belirtmek gerekir. Yapay zekâ demokratik toplumlarda şeffaflığı artırmak, yurttaşların bilgiye erişimini kolaylaştırmak, kamu harcamalarını denetlemek ve yolsuzluğu ortaya çıkarmak için de kullanılabilir. Teknolojinin yönü önceden

yazılmış değildir. Fakat iktidarın tarihsel eğilimi açıktır: **elde ettiği yeni kapasiteyi kullanmak ister.** Bu nedenle asıl soru yapay zekânın ne yapabileceği değil, devletin ve sermayenin yapabildiği şeylerden hangilerini yapmasına izin vereceğimizdir.

Yapay zekânın insanlığı ele geçirmesi üzerine kurulan kıyamet senaryosu, belki de daha gündelik bir tehlikeyi görünmez kılıyor. Makinenin iktidarı ele geçirmesinden önce iktidar yapay zekâyı ele geçirebilir. Sonra onu yurttaşı okumak, çalışanı ölçmek, tüketiciyi yönlendirmek, muhalefeti sınıflandırmak ve savaşı hızlandırmak için kullanabilir. Bu durumda sorun makinenin insana hükmetmesi değildir; **bazı insanların, makinenin kudretiyle diğer insanlara hükmetmesidir.**

Bu nedenle yapay zekâ çağının özgürlük mücadelesi yalnız "insan mı makine mi?" şeklinde kurulamaz. Mesele insanın başka insan üzerindeki teknolojik iktidarını sınırlandırmaktır. Mahremiyet, hesap verebilirlik, açıklanabilirlik, itiraz hakkı, çoğulculuk ve demokratik denetim artık yalnız hukuk kitaplarının kavramları değil, yapay zekâ mimarîsinin tasarım ilkeleri olmak zorundadır. Aksi hâlde insanlık kendisini yönetecek bilinçli bir makine yaratmadan önce, **kendisini görünmez algoritmalarla yöneten kusursuz bir bürokrasi** yaratabilir.

Filozof Kirpi: **"Yapay zekânın diktatör olmasından önce; diktatörün yapay zekâ sahibi olmasından kork."**

6. İNSAN YAPAY ZEKÂNIN GERİSİNDE Mİ KALACAK? ZEKÂDAN HİKMETE İNSANÎ ÜSTÜNLÜĞÜN YENİDEN TANIMLANMASI

Yapay zekâ tartışmasının en ürkütücü sorularından biri şudur: İnsanlık kendi yarattığı zekânın gerisinde mi kalacak? Bu soru ilk bakışta makûl görünür. Yapay zekâ saniyeler içinde milyonlarca metni tarayabilir, karmaşık matematik problemleri çözebilir, büyük veri kümeleri arasındaki ilişkileri keşfedebilir, kod yazabilir, farklı dillerde metin üretebilir ve insanın günlerce sürecek bilişsel faaliyetlerini dakikalara indirebilir. Eğer zekâyı yalnızca işlem hızı, bilgi erişimi, örüntü tanıma ve problem çözme kapasitesi olarak tanımlarsak insanın bu yarışta uzun vadede üstün kalmasını beklemek için güçlü bir gerekçe bulunmayabilir. Fakat tam da burada daha temel bir soru sormak gerekir: **İnsan neden makineyle zekâ yarışına girmek zorundadır?**

İnsanlığın yapay zekâ karşısındaki en büyük kavramsal hatası, insanı yapay zekânın ölçütleriyle değerlendirmeye başlaması olabilir. Bilgisayar saniyede milyarlarca işlem gerçekleştirdiği için insan değersizleşmez; tıpkı otomobil insandan hızlı koştuğu için insan bedeninin anlamsızlaşmaması gibi. Vinç yüzlerce insanın kaldıramayacağı ağırlığı kaldırabilir ama bu, vinci insandan "üstün varlık" hâline getirmez. Problem, bilişsel faaliyetlerin tarih boyunca insanın kendisini tanımladığı alanlardan biri olması nedeniyle yapay zekâ karşısındaki üstünlük meselesinin daha derin bir kimlik krizine dönüşmesidir. Makine kasımızı geçtiğinde bunu ilerleme saydık; hafızamızı geçtiğinde şaşırmadık; fakat muhakememize yaklaşmaya başladığında "insan nedir?" sorusunu yeniden sormak zorunda kaldık.

Bu nedenle yapay zekâ çağı aslında teknoloji kadar antropoloji krizidir. İnsan kendisini uzun süre "düşünen hayvan" olarak tanımladı. Eğer düşünmenin bazı biçimleri artık insan dışı sistemler tarafından gerçekleştirilebiliyorsa insanın ayrıcalığı nerede kalacaktır? Bu soruya "duygularımız var" veya "biz yaratıcıyız" şeklinde aceleci cevaplar vermek yeterli değildir. Çünkü yapay zekâ duygusal ifadeleri taklit edebilir, şiir yazabilir, resim üretebilir, müzik besteleme süreçlerine katılabilir ve yaratıcı görünen kombinasyonlar oluşturabilir. İnsan kendisini yalnız makinenin henüz yapamadığı şeylerin toplamı olarak tanımlarsa teknolojinin her ilerleyişinde insan tanımı biraz daha geri çekilir. Böyle bir savunma hattı uzun vadede sürdürülemez.

Daha sağlam bir ayrım zekâ ile **hikmet** arasında kurulmalıdır. Zekâ, belirli bir problemi çözebilme kapasitesidir; hikmet ise hangi problemin çözülmeye değer olduğuna karar verebilme yeteneğidir. Zekâ araçları optimize eder; hikmet amaçları sorgular. Zekâ "nasıl?" sorusunda olağanüstü başarılı olabilir; hikmet "neden?" ve "ne uğruna?" sorularını taşır. Yapay zekâ bir şehrin trafik akışını en verimli biçimde düzenleyebilir; fakat insanların hayatının ne kadarının ulaşım uğruna harcanmasının makûl olduğuna karar vermek teknik optimizasyonun ötesindedir. Bir sağlık sistemi hangi tedavinin istatistiksel olarak en yüksek başarı oranına sahip olduğunu hesaplayabilir; fakat sınırlı kaynakların hangi hastalara nasıl dağıtılacağı meselesi aynı zamanda adalet problemidir.

Tam burada insanî kararın vazgeçilmezliği ortaya çıkar. Her karar hesap değildir. Bazı kararlar değerler arasında seçim yapmayı gerektirir. Özgürlük ile güvenlik, verimlilik ile adalet, bireysel hak ile kamusal yarar, bugünün refahı ile gelecek nesillerin hakkı arasında yapılacak tercihler yalnız veriyle çözülemez. Veri seçeneklerin sonuçlarını gösterebilir; fakat hangi sonucun **iyi**, hangi bedelin **kabul edilebilir**, hangi sınırın **aşılamaz** olduğuna karar vermek normatif bir faaliyettir. İnsanlık yapay zekâya bu kararları bütünüyle devrettiğinde yalnız yönetim tekniğini değil, ahlâkî özne olma hakkını da devretmiş olur.

Burada yapay zekâ ile insan arasındaki temel farklardan biri **sorumluluk** kavramında belirginleşir. İnsan verdiği kararın ahlâkî yükünü taşıyabilir. Yanlış yaptığında suçluluk duyabilir, özür dileyebilir, cezalandırılabilir, telâfi etmeye çalışabilir ve geçmiş kararını yeniden değerlendirebilir. Yapay zekâ ise yanlış karardan dolayı vicdan azabı çekmez. Bir model "üzgünüm" cümlesini üretebilir; fakat o cümlenin arkasında ahlâkî acı çektiğine ilişkin bir gerekçe yoktur. Bu nedenle bir sistemin doğru karar oranının insandan yüksek olması bile onu otomatik olarak ahlâkî otorite hâline getirmez. Çünkü **kararın doğruluğu ile kararın sorumluluğunu taşıyabilme kapasitesi aynı şey değildir.**

Bu ayrım hukukta, tıpta, siyasette ve savaşta özellikle önemlidir. Bir hâkim yapay zekâdan emsal kararları taramasını isteyebilir; fakat bir insanın özgürlüğünün sınırlandırılmasının nihaî sorumluluğu algoritmaya devredilemez. Doktor yapay zekâdan teşhis desteği alabilir; fakat hastanın hayatı üzerinde kurulacak ilişkinin ahlâkî sorumluluğunu istatistiksel sisteme bırakamaz. Hükûmet ekonomik seçenekleri yapay zekâ ile modelleyebilir; fakat yoksulluk, eşitsizlik veya savaş gibi meselelerde "algoritma böyle önerdi" diyerek siyasî sorumluluktan kaçamaz. İnsanlığın yapay zekâ çağındaki temel ilkelerinden biri şu olmalıdır: **Yetki devredilebilir; sorumluluk devredilemez.**

İnsanî üstünlüğün ikinci önemli alanı anlamdır. Yapay zekâ dilin örüntülerini son derece karmaşık biçimlerde işleyebilir; fakat insan için dil yalnız sembollerin ilişkisi değildir. İnsan sözcükleri biyografisinin, bedeninin, ölümünün, sevincinin, kaybının ve dünyayla kurduğu ilişkinin içinden yaşar. "Anne", "sürgün", "vatan", "ölüm", "çocuk", "ihanet", "sevgi" veya "adalet" kelimeleri yalnız semantik birimler değildir; insan hayatında yaşanmışlığın taşıyıcılarıdır. Bu nedenle insanın anlam dünyası, yalnızca daha fazla veri işlemekle açıklanamayacak ölçüde bedensel, tarihî ve toplumsaldır.

Ölüm bilinci burada özel bir yere sahiptir. İnsan sınırlı olduğunu bilir. Zamanının biteceğini, sevdiklerini kaybedebileceğini, verdiği bazı kararları geri alamayacağını bilir. Ahlâkın önemli bir bölümü tam da bu sonluluk bilincinden doğar. Bir insanın hayatının paha biçilemez oluşu, sonsuz biçimde yeniden başlatılamamasından gelir. Teknolojik sistemlerin kararlarını insan hayatına uygularken gözden kaçırılmaması gereken temel noktalardan biri budur: İnsan yalnız optimize edilecek bir değişken değildir. İnsan **biricik hayatının öznesidir**.

Bu nedenle yapay zekâ çağında "insan merkezcilik" ile insan haysiyeti arasında da ayrım yapmak gerekir. İnsanlığın doğaya istediği gibi hükmetme hakkı olduğunu söylemek başka şeydir; insanın araçsallaştırılamaz bir haysiyete sahip olduğunu savunmak başka şey. Yapay zekâ sistemleri insanların davranışlarını tahmin

edebilir, sınıflandırabilir ve optimize edebilir; fakat insan kendisini yalnızca tahmin edilebilir veri kümelerinin toplamına indirgenmeye karşı korumalıdır. Çünkü özgürlük biraz da insanın geçmiş davranışlarından çıkarılan istatistiksel modele rağmen **başka türlü davranabilme ihtimalidir**.

Tam burada yapay zekânın neden hiçbir zaman mutlak epistemik otorite kabul edilmemesi gerektiği anlaşılır. Daha fazla bilgiye erişmek, daha doğru tahminler yapmak ve daha iyi örüntü tanımak, kendiliğinden siyasî meşruiyet üretmez. Demokratik toplumlarda kararın meşruiyeti yalnız kararın teknik doğruluğundan değil, kararın nasıl alındığından, kim tarafından tartışıldığından ve kimlerin itiraz edebildiğinden doğar. Bir yapay zekâ ekonomik açıdan "en verimli" bütçe dağılımını hesaplayabilir; fakat yurttaşların bunun adaletli olup olmadığı konusunda söz söyleme hakkının yerini alamaz.

Bu nedenle gelecekte demokrasi ile yapay zekâ arasındaki temel gerilimlerden biri **teknokratik cazibe** olacaktır. "Makine bizden daha doğru hesaplıyor, o hâlde kararları da makine versin" düşüncesi ilk bakışta rasyonel görünebilir. İnsan siyaseti yolsuzluk, çıkar çatışması, önyargı ve cehaletle malûldür. Algoritmik yönetim bunların bir bölümünü azaltabilir. Fakat algoritma önyargıdan tamamen bağımsız değildir; onu tasarlayan kurumların hedeflerini, seçilen verilerin sınırlarını ve tanımlanan başarı ölçütlerini taşır. Dahası siyasetin amacı yalnız doğru sonucu bulmak değildir; insanların kendi müşterek hayatlarının kuruluşuna katılmasıdır. Demokrasi biraz da yanlış yapma hakkını birlikte taşıyabilme rejimidir.

İnsan yapay zekânın gerisinde kalmamak için onun gibi olmak zorunda değildir. Tam tersine, en büyük tehlike insanın yapay zekâyla rekabet edebilmek için kendisini de bir optimizasyon makinesine dönüştürmesidir. Daha hızlı üretmek, daha fazla bilgi işlemek, sürekli ölçülmek, her dakikayı verimlilik ölçütüyle değerlendirmek ve insanî hayatı performans tablolarına çevirmek, makinenin insanlaşmasından çok **insanın makineleşmesi** anlamına gelir. Yapay zekâ çağındaki insanî direnç, verimsizliğin romantikleştirilmesi değildir; ölçülemeyen insanî değerlerin varlığını savunmaktır.

Bir çocuğa ayrılan zaman ekonomik bakımdan "verimsiz" görünebilir. Yaşlı bir insanı dinlemek üretkenlik endeksine katkı sağlamayabilir. Bir şiir okumak, denize bakmak, yas tutmak, dostla uzun bir konuşma yapmak veya hiçbir ekonomik sonuç üretmeden düşünmek kapitalist ölçüm sistemlerinde değer üretmeyebilir. Fakat insan hayatının anlamı tam da bu hesaplanamayan alanlarda kurulabilir. Yapay zekâ toplumu bütün faaliyetleri optimize edilebilir çıktılar olarak görmeye başladığında tehlike makinenin güçlenmesinden çok **insan hayatının ölçülebilen şeylere indirgenmesidir**.

Bu nedenle geleceğin eğitim sistemi yalnız yapay zekâyla rekabet edecek beceriler öğretmemelidir. Çocuklara daha hızlı kod yazmayı, daha iyi veri analiz etmeyi ve daha verimli üretmeyi öğretmek elbette gereklidir; fakat yeterli değildir. Eğitim aynı zamanda ahlâkî muhakeme, tarih şuuru, estetik duyarlılık, empati, eleştirel düşünce, tartışma kültürü ve başkasının acısını tanıyabilme yeteneği kazandırmalıdır. Çünkü yapay zekâ çağında insanı değerli kılacak şey makinenin yapamadığı bir işlemi bulmak değil, **makinenin yaptığı işlemin hangi insanî amaca hizmet etmesi gerektiğini belirleyebilmektir**.

Buradan insanlığın yapay zekâya karşı nasıl örgütlenmesi gerektiği sorusuna da ilk cevap çıkar. İnsan tek tek bireyler hâlinde yapay zekâyla yarışamaz. Bir insan milyonlarca kitabı okuyamaz, dünyanın bütün yazılım dillerine hâkim olamaz veya saniyede milyarlarca işlemi gerçekleştiremez. Fakat insanlığın üstünlüğü zaten tek bireyin bilişsel kapasitesinde değildir. Bilim, hukuk, üniversite, demokrasi, kurumlar, etik kurullar, meslek örgütleri ve kamusal tartışma insanlığın **kolektif zekâ mekanizmalarıdır**. Yapay zekâ karşısında korunması gereken esas üstünlük burada yatmaktadır.

Dolayısıyla insanın cevabı "daha zeki bireyler yetiştirmek" kadar, hatta bundan daha fazla, **daha akıllı ve daha âdil kurumlar kurmak** olmalıdır. Çok güçlü bir yapay zekâyı denetleyebilecek tek bir süper insan olmayacaktır. Onu hukuk, bağımsız bilim, uluslararası denetim, demokratik siyaset, teknik güvenlik ve

kamusal sorumluluktan oluşan kurumsal ağ denetleyebilir. Teknolojik zekâ ne kadar merkezîleşirse insanî denetimin o kadar çoğulcu olması gerekir.

Geleceğin asıl çatışması bu nedenle yapay zekâ ile insan zekâsının skorlarını karşılaştırmak değildir. Asıl mesele, insanlığın kendi yarattığı üstün bilişsel araçları **hangi değerler içinde tutabileceğidir**. İnsan satrançta makineye yenildi; fakat satrancın neden oynanmaya değer olduğunu makine belirlemedi. Hesaplamada yenildi; fakat hangi hesabın yapılması gerektiğine hâlâ insan karar veriyor. Yarın bilimsel araştırmada, hukukta, mühendislikte ve stratejik analizde de geride kalabilir. Fakat insan amaç belirleme, meşruiyet kurma ve sorumluluk taşıma hakkını terk etmediği sürece yenilmiş sayılmaz.

Asıl yenilgi makinenin bizden daha zeki olması değildir. Asıl yenilgi, insanlığın "daha zeki olan yönetmelidir" düşüncesini kabul etmesidir. Çünkü zekâ iktidar için yeterli meşruiyet değildir. Bir varlığın daha hızlı hesap yapması ona insan üzerinde hüküm verme hakkı kazandırmaz. Eğer bunu kabul edersek yalnız yapay zekâya değil, tarih boyunca her "daha güçlü olan yönetir" iddiasına teslim olmuş oluruz.

Bu nedenle yapay zekâ çağının insan tanımı yeniden kurulmalıdır. İnsan, en hızlı hesaplayan varlık değildir. İnsan; neyin hesaplanmaması gerektiğini söyleyebilen, bir başkasının acısını sayıların arkasından görebilen, kudretine sınır koyabilen ve kendi çıkarına rağmen adaleti seçebilen varlıktır. İnsanlığın makine karşısındaki gerçek üstünlüğü zekâda değil, **zekâya sınır çizebilecek hikmette** aranmalıdır.

Filozof Kirpi: **"Makine senden daha zeki olabilir; fakat hangi zekânın nerede durması gerektiğini sen söyleyemiyorsan, kaybettiğin şey zekâ değil, insanlıktır."**

7. MAKİNEYİ ANAYASAL DÜZEN ALTINA ALMAK: İNSANLIĞIN YAPAY ZEKÂYLA YENİ TOPLUM SÖZLEŞMESİ

Yapay zekâ tartışmasının sonunda insanlığın önünde iki kolay fakat yanlış yol bulunmaktadır. Birincisi korkuya kapılıp teknolojiyi durdurmaya çalışmaktır. İkincisi ise teknolojik gelişmenin kendi iç mantığının sorunları zamanla çözeceğine inanarak yapay zekâyı piyasanın, şirketlerin ve devletlerin rekabetine bırakmaktır. Birinci seçenek gerçekçi değildir; ikinci seçenek ise tarih bilgisinden yoksundur. İnsanlık matbaayı, elektriği, otomobili, ilacı, nükleer enerjiyi veya finansal piyasaları yalnızca onları icat edenlerin vicdanına bırakmadı. Büyük toplumsal güç doğuran her teknoloji zamanla hukuk, kurum, meslek etiği, standart ve kamusal denetim üretmek zorunda kaldı. Yapay zekâ bunun istisnası olamaz. Tam tersine, insan kararlarının giderek daha büyük bölümüne nüfuz etme ihtimali nedeniyle belki de tarihte en ayrıntılı yönetim rejimini gerektirecek teknolojilerden biridir.

Fakat burada "regülasyon" kelimesi tek başına yetersizdir. Yapay zekâ meselesi yalnız şirketlerin uyması gereken birkaç teknik standardın belirlenmesi değildir. İnsanlığın karşı karşıya bulunduğu mesele daha temeldir: **Hangi kararın makineye devredilebileceği, hangisinin devredilemeyeceği; hangi yetkinin mutlaka insanın elinde kalacağı ve teknolojik kudretin hangi sınırda durdurulacağı** belirlenmek zorundadır. Bu nedenle ihtiyaç duyulan şey basit bir yapay zekâ mevzuatından çok, bir tür **yapay zekâ anayasal düzenidir**.

Anayasa siyasal iktidarın ne yapabileceğini belirlemekten önce ne yapamayacağını belirler. Devlet güçlüdür; fakat her şeyi yapamaz. Çoğunluk güçlüdür; fakat temel hakları keyfî biçimde ortadan kaldıramaz. Güvenlik gerekçesi önemlidir; fakat sınırsız değildir. Aynı düşünme biçimi yapay zekâya da uygulanmalıdır. Bir sistem teknik olarak bir şeyi yapabiliyor diye toplum ona o şeyi yapma yetkisi vermek zorunda değildir. Teknolojik imkân ile hukukî meşruiyet birbirinden ayrılmalıdır. Yapay zekâ çağının ilk anayasal ilkesi bu nedenle şu olmalıdır: **Yapılabilir olan ile yapılmasına izin verilen aynı şey değildir.**

Avrupa Birliği'nin AI Act düzenlemesi bu yöndeki ilk büyük ölçekli hukukî deneylerden biridir. Düzenleme yapay zekâyı tek kategori olarak değil, risk düzeylerine göre ele almakta; bazı uygulamaları kabul edilemez risk nedeniyle yasaklamakta, genel amaçlı modeller için belirli yükümlülükler getirmekte ve yüksek riskli sistemler için daha ağır şartlar öngörmektedir. 2 Ağustos 2026 itibarıyla AI Office ve ulusal makamlar önemli hükümlerin uygulanması konusunda yetki kullanmaya başlamış; şeffaflık yükümlülükleri de yürürlüğe girmiştir. Bununla birlikte eğitim, istihdam, biyometri, kritik altyapı, göç ve benzeri alanlardaki bazı yüksek riskli sistemlere ilişkin kapsamlı yükümlülüklerin uygulanması 2 Aralık 2027'ye, düzenlenmiş fizikî ürünlere gömülü bazı sistemlerin yükümlülükleri ise 2 Ağustos 2028'e uzatılmıştır.

Buradaki risk temelli yaklaşım önemlidir; çünkü her yapay zekâ aynı düzeyde tehlikeli değildir. Bir fotoğraf uygulamasının arka planını değiştiren sistem ile mahkûmun tahliye edilip edilmeyeceğine ilişkin tavsiyede bulunan sistem aynı hukukî rejime tâbi tutulmamalıdır. E-posta metni düzelten yapay zekâ ile elektrik şebekesini yöneten yapay zekâ arasında aynı güvenlik standardını uygulamak hem anlamsız hem de verimsizdir. Dolayısıyla geleceğin düzenleme ilkesi kapasite + kullanım alanı + erişim yetkisi + olası zarar büyüklüğü bileşimi üzerinden kurulmalıdır. Risk yükseldikçe sistemin bağımsız denetimi, kayıt tutma yükümlülüğü, insan gözetimi ve hukukî sorumluluğu da ağırlaşmalıdır.

Fakat insan gözetimi yalnız mevzuata yazılmış sembolik bir cümle olmamalıdır. "Human in the loop" ifadesi kolaylıkla hukukî makyaja dönüşebilir. İnsan ekranda makinenin kararını onaylayan kişi olarak bulunabilir fakat kararın nasıl üretildiğini anlamıyor, itiraz edecek zamanı bulunmuyor veya sistemi değiştirme yetkisine sahip değilse ortada gerçek bir insan denetimi yoktur. İnsan denetiminin anlamlı olması için insanın bilme, sorgulama, durdurma ve değiştirme yetkisine sahip olması gerekir. Yapay zekâ anayasal düzeninin temel kurumu bu nedenle "insan veto hakkı" olmalıdır.

Özellikle insanın temel haklarını etkileyen alanlarda hiçbir sistem kendi kararının nihaî hukukî otoritesi olmamalıdır. Bir kişinin işe alınması, krediye erişimi, kamu yardımından yararlanması, mahkûm edilmesi, sağlık hizmetine erişmesi veya özgürlüğünün sınırlandırılması gibi kararlarda yapay zekâ yardımcı olabilir; fakat kişinin anlamlı bir insan incelemesine ve karara itiraz imkânına sahip olması gerekir. OECD'nin güncellenmiş yapay zekâ ilkeleri de yapay zekâ sistemlerinden olumsuz etkilenen insanların çıktıya itiraz edebilmesini, sistemlerin gerektiğinde geçersiz kılınabilmesini veya güvenli biçimde devreden çıkarılabilmesini ve sorumluluğun açık biçimde tanımlanmasını vurgulamaktadır.

İkinci büyük ilke bağımsız denetimdir. Yapay zekâ şirketlerinin kendi modellerini test etmesi zorunludur fakat yeterli değildir. Bir ilaç şirketinin "ürünümü test ettim, güvenlidir" demesi nasıl tek başına yeterli kabul edilmiyorsa, milyarlarca insanı etkileyebilecek yapay zekâ modellerinde de üreticinin kendi güvenlik beyanı son söz olamaz. Güçlü modeller piyasaya çıkmadan önce bağımsız kuruluşlar tarafından saldırgan testlere, güvenlik değerlendirmelerine ve yüksek riskli kullanım senaryolarına tâbi tutulmalıdır. Siber kapasite, biyolojik risk, manipülasyon, özerk faaliyet, kritik altyapıya müdahale ve kontrol kaybı gibi alanlarda standartlaştırılmış değerlendirmeler geliştirilmelidir.

ABD Ulusal Standartlar ve Teknoloji Enstitüsü'nün AI Risk Management Framework'ü bu açıdan önemli bir yaklaşım geliştirmiştir. Çerçeve yapay zekâ risk yönetimini "Govern, Map, Measure, Manage" eksenleri etrafında ele almakta; risklerin teknoloji yaşam döngüsünün tamamında sürekli olarak değerlendirilmesini öngörmektedir. Bununla birlikte çerçevenin temel karakteri gönüllülüğe dayanmaktadır. NIST, 2026 itibarıyla kritik altyapılarda güvenilir yapay zekâ kullanımına ilişkin özel bir profil de geliştirmektedir. Böyle gönüllü standartlar teknik kültür oluşturabilir; fakat toplumun temel hakları söz konusu olduğunda yalnız gönüllülüğe dayalı sistem yeterli değildir. Etik tavsiye ile hukukî yükümlülük arasındaki sınır, şirketin uyup uymama serbestisinin başladığı yerde ortaya çıkar.

Üçüncü ilke **sorumluluğun algoritmaya kaçırılamaması**dır. Yapay zekâ yanlış teşhis verdiğinde, banka hukuka aykırı biçimde kredi reddettiğinde, işveren algoritmik sistemle çalışanı ayrımcılığa uğrattığında veya otonom sistem fizikî zarar oluşturduğunda "AI böyle karar verdi" hukukî savunma olmamalıdır. Yapay zekânın hukuk kişiliğine sahip olmaması, onun kullanıldığı sistemdeki insan ve kurumların sorumluluğunu ortadan kaldırmaz. Aksine yapay zekâ ne kadar özerkleşirse sorumluluk zincirinin daha ayrıntılı biçimde belirlenmesi gerekir. Model geliştiricisi, sistemi piyasaya süren şirket, entegrasyonu yapan kurum ve sistemi kullanan nihai aktörün hangi durumda ne ölçüde sorumlu olacağı önceden tanımlanmalıdır.

Dördüncü ilke **izlenebilirliktir**. Kritik yapay zekâ kararlarında hangi modelin kullanıldığı, hangi sürümün çalıştığı, hangi verilerin karar üzerinde etkili olduğu, sistemin hangi aracı ne zaman kullandığı ve hangi insanın hangi aşamada onay verdiği kayıt altına alınmalıdır. Uçak kazasından sonra kara kutu aranıyorsa, milyonlarca insanın hayatını etkileyen algoritmik sistemlerin de kurumsal kara kutusu bulunmalıdır. İz bırakmayan iktidar denetlenemez.

Beşinci mesele güç yoğunlaşmasıdır. Yapay zekâyı yalnız güvenlik açısından düzenleyip mülkiyet meselesini görmezden gelmek büyük hata olur. Dünyanın en güçlü modellerinin, veri merkezlerinin, çiplerin ve bulut altyapısının birkaç şirketin elinde yoğunlaşması, yapay zekâ yönetişimini ekonomik rekabet meselesinin ötesine taşımaktadır. Bu nedenle rekabet hukuku, veri taşınabilirliği, birlikte çalışabilirlik, akademik araştırmaya makûl erişim ve kamu araştırma kapasitesi yapay zekâ rejiminin temel unsurları olmalıdır. Teknolojik kudret tekelleştikçe siyasî denetim zorlaşır. **Yapay zekânın demokratikleştirilmesi, yalnız herkesin chatbot kullanabilmesi değil, yapay zekâ üzerindeki iktidarın dağılmasıdır.**

Altıncı sınır çocuklardır. Çocuklar yapay zekâ ekonomisinin sıradan tüketicileri olarak değerlendirilemez. Henüz kimliği, değer dünyası, eleştirel muhakemesi ve duygusal sınırları oluşmakta olan çocuklara yönelik yapay zekâ sistemlerinin tasarımında çok daha ağır standartlar uygulanmalıdır. Çocuğun dikkatini mümkün olduğunca uzun süre sistemde tutmaya çalışan bağımlılık tasarımları, psikolojik zayıflıklardan yararlanan manipülasyon teknikleri, mahrem veri çıkarımı ve çocuğu insan ilişkisinin yerine yapay ilişkiye bağımlı hâle getiren uygulamalar özel koruma alanına alınmalıdır. Çocuk güvenliği, teknolojik yenilik karşısında pazarlığa açık bir maliyet kalemi değildir. Bir teknolojinin topluma faydalı olup olmadığını anlamanın ahlâkî ölçülerinden biri, **en savunmasız çocuğun karşısında nasıl davrandığıdır**.

Yedinci ve belki de en sert kırmızı çizgi askerî alandır. İnsan öldürme kararının bütünüyle makineye bırakılması, yapay zekâ çağının aşılmaması gereken sınırlarından biri olmalıdır. Avrupa Birliği AI Act'in yalnızca askerî, savunma veya ulusal güvenlik amacıyla kullanılan yapay zekâ sistemlerini kendi kapsamı dışında bırakması da küresel düzeyde ayrı bir hukukî düzenleme ihtiyacını göstermektedir. Ağustos 2026'da Birleşmiş Milletler ve Uluslararası Kızılhaç Komitesi cephesinden yapılan yeni çağrılarda da insanların makineler tarafından özerk biçimde hedef alınması "ahlâkî kırmızı çizgi" olarak nitelendirilmiş ve devletlere bağlayıcı sınırlamalar oluşturma çağrısı yenilenmiştir. İnsan yaşamına son verme kararında anlamlı insan kontrolü yalnız teknik güvenlik meselesi değil, insan haysiyetinin hukukî sınırıdır.

Fakat ulusal düzenleme tek başına yeterli olmayacaktır. En güçlü yapay zekâ modelleri ulusal sınır tanımadan kullanılmaktadır. Bir ülkede geliştirilen model birkaç saat içinde dünyanın başka taraflarında etkili olabilir. Bu nedenle yapay zekâ güvenliği zamanla nükleer güvenlik, havacılık veya küresel salgınlar gibi uluslararası koordinasyon gerektiren bir alana dönüşecektir. Avrupa Konseyi'nin 5 Eylül 2024'te imzaya açtığı Yapay Zekâ, İnsan Hakları, Demokrasi ve Hukukun Üstünlüğü Çerçeve Sözleşmesi bu yönde önemli bir adımdır ve yapay zekâ alanındaki ilk bağlayıcı uluslararası antlaşma çerçevesini oluşturmuştur. Sözleşme risk ve etki değerlendirmeleri, insan hakları, demokrasi ve hukuk devleti bakımından koruyucu tedbirler ve gerektiğinde bazı uygulamalara yasak veya moratoryum getirilmesi imkânını öngörmektedir.

Bunun daha ilerisine gidilmesi gerekecektir. Çok yüksek kapasiteye sahip modeller konusunda uluslararası olay bildirim sistemi, ortak güvenlik testleri, minimum siber güvenlik standartları, kritik kapasite eşikleri ve bağımsız bilimsel değerlendirme ağları oluşturulmalıdır. Bu sistem doğrudan nükleer denetimin kopyası olamaz; yapay zekâ teknolojisinin dağıtılabilirliği ve özel şirketlerin ağırlığı çok farklıdır. Fakat temel mantık benzerdir: sonucu bütün insanlığı etkileyebilecek teknolojiler yalnız ulusal egemenlik veya şirket sırrı gerekçesiyle bütünüyle kapalı alan hâline getirilemez.

Burada önemli bir başka ilke daha ortaya çıkar: Düzenleme yeniliği öldürmemelidir, fakat yenilik söylemi de denetimsizliği meşrulaştırmamalıdır. "Çin bizi geçer", "rakip şirket önce davranır", "teknoloji zaten durdurulamaz" gibi ifadeler güvenlik standartlarını düşürmenin mazereti hâline gelirse yapay zekâ sektörü kolektif bir yarış trajedisine sürüklenebilir. Her şirket güvenlik için yavaşlamanın kendi rekabet gücünü azaltacağını düşünür; bütün şirketler hızlandığında ise toplumun toplam riski büyür. Tam da bu nedenle hukuk vardır: bireysel aktörlerin rekabet mantığıyla kendi başlarına üretemeyeceği ortak sınırları kurmak için.

Yapay zekâ yönetişiminin yalnız teknokratlara bırakılması da yanlış olacaktır. Mühendislerin sistemlerin nasıl çalıştığını bilmesi gereklidir; fakat sistemlerin nerede kullanılmasına izin verileceği yalnız mühendislik sorusu değildir. Hekimler, hukukçular, sosyologlar, psikologlar, eğitimciler, çalışan örgütleri, çocuk gelişimi uzmanları, felsefeciler ve yurttaşlar da bu kararların tarafıdır. Çünkü yapay zekâ teknik bir nesne olarak laboratuvarda doğabilir fakat toplumsal sonuçları laboratuvarın çok ötesinde gerçekleşir. Teknik uzmanlık, siyasî meşruiyet değildir.

Nihayet yapay zekâyla kurulacak yeni toplum sözleşmesinin merkezine insanın vazgeçilmez bazı hakları yerleştirilmelidir: İnsan olduğunu sandığı şeyin makine olduğunu bilme hakkı; kendisi hakkında kullanılan algoritmik karar sisteminden haberdar olma hakkı; ağır sonuç doğuran otomatik kararlara itiraz hakkı; kişisel verisinin sınırsız biçimde sömürülmemesi hakkı; yapay zekâ tarafından manipüle edilmemek hakkı; insanî muhataba erişme hakkı; kritik kararın gerekçesini öğrenme hakkı ve nihayet bazı alanlarda makine tarafından yönetilmeme hakkı.

Bu son hak geleceğin en önemli özgürlüklerinden biri olabilir. Nasıl modern hukuk işkence edilmeme, keyfî tutuklanmama ve özel hayatın korunması haklarını zaman içinde oluşturduysa, yapay zekâ çağında da bireyin hayatının tamamının algoritmik optimizasyona teslim edilmemesini güvence altına alan yeni özgürlük kategorileri doğacaktır.

İnsanlığın yapay zekâyla gelecekte kurması gereken ilişki efendi-köle ilişkisi değildir. Yapay zekâyı "köle" olarak görmek de hatalıdır; çünkü teknik kapasite büyüdükçe böyle bir metafor gerçeği açıklamaz. Daha doğru ilişki, anayasal araç ilişkisidir. İnsan teknolojiye büyük yetenekler verebilir fakat onun kullanım alanlarını hukukla sınırlar; sistem güçlü olabilir fakat yetkisiz kalır; çok şey bilebilir fakat meşruiyet sahibi olmaz; karar önerebilir fakat insanlığın kaderi üzerinde kendiliğinden egemenlik iddia edemez.

Bütün otopsinin sonunda ölüm sebebi henüz yazılmamıştır. Yapay zekâ insanlığın katili olmak zorunda değildir. Aynı teknoloji insan ömrünü uzatabilir, hastalıkları teşhis edebilir, bilimi hızlandırabilir, angaryayı azaltabilir, eğitimi erişilebilir kılabilir ve insanın yaratıcı imkânlarını genişletebilir. Fakat bunun gerçekleşmesi teknolojinin doğasından otomatik olarak çıkmayacaktır. Sonucu belirleyecek olan, insanlığın kendi kudretine sınır koyabilme kudreti olacaktır.

Çünkü nihayetinde yapay zekâyı yapan da insandır, ona hedef veren de, elektriğini sağlayan da, veri merkezini kuran da, onu savaşa veya hastaneye götüren de. Eğer bir gün yapay zekâ insanın efendisine dönüşürse bunun hikâyesi yalnız makinenin yükselişinin değil, insanın kendi egemenlik yetkilerini hangi aşamalarda terk ettiğinin de hikâyesi olacaktır.

Yapay zekâ karşısında insanlığın temel siyasî programı bu nedenle korku değil, örgütlenmedir; yasakçılık değil, sınır koymadır; teknolojiden kaçmak değil, teknolojiyi hukuk ve ahlâk içinde tutmaktır. İnsan makinenin düşünmesini engellemek zorunda değildir. **İnsan, makinenin düşünme kapasitesinin insan üzerinde sınırsız iktidara dönüşmesini engellemek zorundadır.**

Filozof Kirpi: **"Yapay zekâyı zincire vurmak gerekmez; yetkisini hukuka, kudretini ahlâka, düğmesini insana bağlamak yeterlidir."**