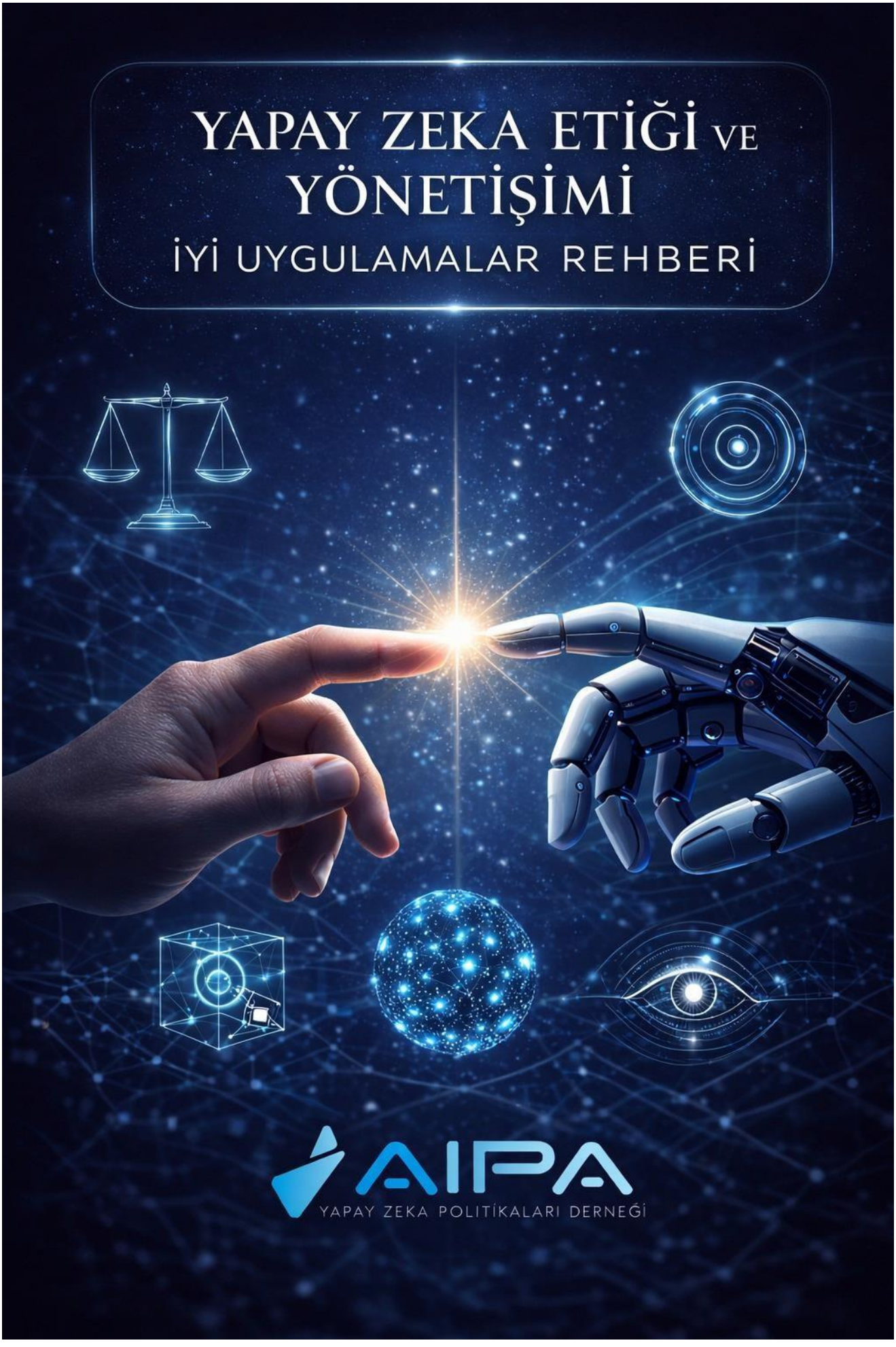
YAPAY ZEKA ETİĞİ VE YÖNETİŞİMİ

İYİ UYGULAMALAR REHBERİ



AIPA

YAPAY ZEKA POLİTİKALARI DERNEĞİ

YAPAY ZEKA ETİĞİ VE YÖNETİŞİMİ İYİ UYGULAMALAR REHBERİ

Mart 2026

Yapay Zekâ Politikaları Derneği

Adres: Ankara Teknoloji Geliştirme Bölgesi, Üniversiteler Mahallesi, 1606. Cadde, Kapı No: 4/B, Cyberpark Cyberplaza B Blok, Zemin Kat No: BZ11, 06800 Bilkent-Çankaya/Ankara

info@aipaturkey.org

AIPA

YAPAY ZEKÂ ETİĞİ VE YÖNETİŞİMİ İYİ UYGULAMALAR REHBERİ

Türkiye'de yapay zekâ sistemlerinin etik ve yönetişimsel olarak geliştirilmesi ve kullanımına yönelik bu iyi uygulamalar Rehberi, AIPA tarafından uluslararası ilke ve düzenlemeler ışığında hazırlanmıştır.

TEŞEKKÜR MESAJI

Bu çalışma Yapay Zekâ Politikaları Derneği (AIPA- Artificial Intelligence Policy Association) tarafından farklı disiplinlerde çalışan uzmanların görüşleri alınarak hazırlanmıştır. Türkiye'de bir sivil toplum örgütü tarafından ilk kez açıklanan yapay zekâ etiği ve yönetişimine ilişkin bu iyi uygulamalar rehberinin oluşturulmasına katkıda bulunan aşağıda belirtilen çok kıymetli isimlere teşekkürlerimizi sunarız.

- **Zafer Küçükşabanoğlu**

Yapay Zekâ Politikaları Derneği (AIPA) Kurucusu ve Başkanı, RecroTech Kurucusu ve Yönetim Kurulu Başkanı

- **Melih Alp**

AIPA Teknik Komisyon Üyesi, Akbank İç Kontrol Teknoloji Uygulamaları Müdürü, Yapay Zekâ Danışmanı

- **Çiğdem Şengül**

AIPA Eğitim Komisyonu Üyesi, Yapay Zekâ Yönetişim ve Çevik Dönüşüm Danışmanı

- **Şermin Eldek**

AIPA Yapay Zekâ Diplomasisi Başkan Yardımcısı ve Teknik Komisyon Üyesi, Yapay Zekâ ve Kalite Güvence Mühendisi

- **Av. Elif Selin Yılmaz Altay**

AIPA Yönetim Kurulu Üyesi, AIPA Hukuk ve İnsan Hakları Komisyonu Başkanı, Enerjisa Grup Uyum Müdürü

- **Defne Ağaoğlu**

AIPA Genel Koordinatörü

- **Özge Öztürk**

AIPA Dış İletişim Koordinatörü

- **Zeynep Koral**

AIPA İç İletişim Koordinatörü

ÖNSÖZ

‘’Gerçek bilgelik, gücü ölçülülükle kullanabilmektir’’

Seneca

Yapay zekâ, artık geleceğe ait bir ihtimal değil; ekonomik, toplumsal ve kurumsal yapıları bugünden dönüştüren güçlü bir gerçekliktir. İçinde bulunduğumuz dönemde yapay zekâ; yalnızca teknolojik yeniliklerin değil, aynı zamanda kalkınma modellerinin, rekabet anlayışının, karar alma süreçlerinin ve toplumsal organizasyon biçimlerinin yeniden tanımlandığı bir alan hâline gelmiştir. Bu nedenle yapay zekâ meselesi, sadece mühendislik ya da yazılım perspektifiyle ele alınabilecek dar bir konu olmaktan çıkmış; kamu politikalarından özel sektör stratejilerine, eğitimden hukuka kadar uzanan çok boyutlu bir yönetişim başlığına dönüşmüştür.

Bugün dünyada yapay zekâ yarışının merkezinde yalnızca daha güçlü sistemler geliştirmek yoktur. Asıl mesele, bu sistemlerin hangi ilkelerle geliştirileceği, hangi sınırlar içinde işletileceği ve toplum yararıyla nasıl uyumlu hâle getirileceğidir. Çünkü teknolojik kapasite ne kadar artarsa artsın, onu yönlendiren etik çerçeve ve kurumsal denge mekanizmaları yeterince güçlü değilse ortaya çıkan ilerleme, beraberinde ciddi riskler de doğurabilir. Ayrımcılık, hesap verebilirlik eksikliği, denetim zafiyetleri, veri güvenliği sorunları ve insan iradesinin giderek daha fazla geri plana itilmesi gibi meseleler, yapay zekâ çağının üzerinde en çok düşünülmesi gereken başlıkları arasındadır.

Tam da bu noktada etik ve yönetişim, yapay zekânın çevresine sonradan eklenen ikincil kavramlar değil; bu alanın sağlıklı gelişiminin asli şartlarıdır. Yapay zekâ sistemlerinin güven veren, insan onurunu gözeten, hukuki ve toplumsal sorumluluk taşıyan bir anlayışla geliştirilmesi; yalnızca kurumların itibarı açısından değil, toplumların geleceği açısından da belirleyicidir. Yapay zekânın sunduğu potansiyelin kalıcı ve kapsayıcı bir faydaya dönüşebilmesi, güçlü ilkelerle desteklenen bir yönetişim yaklaşımına bağlıdır. Yapay Zekâ Politikaları Derneği (AIPA) olarak biz,

Türkiye'de yapay zekâ alanındaki gelişimin yalnızca teknik başarı göstergeleriyle değerlendirilmesini yeterli görmüyoruz. Bizim için esas olan; bu gelişimin güven duygusunu besleyen, insanı merkeze alan, kurumsal sorumluluğu güçlendiren ve uzun vadeli toplumsal faydayı önceleyen bir zeminde ilerlemesidir. Yapay zekâya ilişkin etik ilkelerin, yönetişim modellerinin ve uygulama standartlarının güçlenmesi; ülkemizin bu alandaki kapasitesini daha sağlam, daha itibarlı ve daha sürdürülebilir bir yapıya kavuşturacaktır.

Bu çerçevede hazırlanan "Yapay Zekâ Etiği ve Yönetişimi İyi Uygulamalar Rehberi", yapay zekâ sistemleriyle çalışan ya da bu alanda kurumsal kapasite geliştirmek isteyen yapılar için yol gösterici bir kaynak niteliği taşımaktadır. Rehber; veri yönetiminden model süreçlerine, kurumsal karar mekanizmalarından izleme ve denetim yapılarına kadar uzanan geniş bir çerçevede, yapay zekâ alanında etik ve yönetişim temelli bir bakış açısının nasıl inşa edilebileceğine dair önemli bir perspektif sunmaktadır. Bu yönüyle çalışma, yalnızca bugünün ihtiyaçlarına değil, gelecekte daha da belirginleşecek kurumsal sorumluluk alanlarına da ışık tutmaktadır.

Bu kıymetli çalışmanın hazırlanmasına katkı sunan AIPA Teknik Komisyon Üyesi Melih Alp'e, AIPA Eğitim Komisyonu Üyesi Çiğdem Şengül'e, AIPA Yapay Zekâ Diplomasisi Başkan Yardımcısı ve Teknik Komisyon Üyesi Şermin Eldek'e ve AIPA Yönetim Kurulu Üyesi, Hukuk ve İnsan Hakları Komisyonu Başkanı Av. Elif Selin Yılmaz Altay'a içten teşekkürlerimi sunuyorum. Bu rehber, farklı uzmanlık alanlarının ortak emeğiyle ortaya çıkmış; çok boyutlu ve nitelikli bir çalışmanın ürünüdür.

İnanıyorum ki Türkiye'nin yapay zekâ alanındaki gücü, yalnızca ürettiği teknolojiyle değil; o teknolojiyi hangi değerler çerçevesinde konumlandırdığıyla da belirlenecektir. Önümüzdeki dönemde asıl fark yaratacak olan, yapay zekâyı ne kadar hızlı geliştirdiğimizden çok, onu ne kadar bilinçli, ne kadar sorumlu ve ne kadar isabetli bir anlayışla yönettiğimiz olacaktır. Bu rehberin, ülkemizde yapay zekâ alanında şekillenecek kurumsal yaklaşımlara katkı sunmasını; etik ilkelere dayalı, insan odaklı ve sağlam yönetişim anlayışının güçlenmesine vesile olmasını diliyorum.

Saygılarımla
Zafer KÜÇÜKŞABANOĞLU
Yapay Zekâ Politikaları Derneği (AIPA) Kurucusu ve Başkanı

İÇİNDEKİLER

Yapay Zekâ Etiği ve Yönetişimi İyi Uygulamalar Rehberi

Bu rehber, Türkiye'de yapay zekâ (YZ) sistemlerinin etik, güvenilir ve yönetişimsel olarak sorumlu biçimde geliştirilmesi ve kullanılmasına rehberlik etmek amacıyla, Yapay Zekâ Politikaları Derneği (AIPA) tarafından ulusal mevzuat ve uluslararası ilkeler ışığında hazırlanmıştır. Bu çalışma ile AIPA, özel sektör kuruluşları arasında yapay zekâ yönetişimine ilişkin ortak bir referans noktası oluşturmayı hedeflemektedir.

Bu bağlamda yapay zekâ yönetişimi (AI governance) kavramı, bu rehberin referans aldığı uluslararası çerçevelerden biridir. Yapay zeka yönetişimi; yalnızca etik ilkeler veya veri yönetimiyle sınırlı olmayan, yapay zekâ sistemlerinin tasarım, geliştirme, uygulama ve izleme aşamalarında hesap verebilirlik, insan gözetimi, güvenlik, açıklanabilirlik ve sürdürülebilirlik ilkelerini bütüncül biçimde ele alan bir anlayıştır. Bu yaklaşım, yapay zekâ sistemlerinin yalnızca mevzuata uyumlu değil, aynı zamanda sorumlu, güvenilir ve insan merkezli biçimde yönetilmesini amaçlar.

Rehberin temel hedeflerinden biri, yapay zekâ sistemlerinin iş dünyasında güvenilirliğini ve itibar algısını güçlendirmektir. Kullanıcıların ve paydaşların, kullanılan sistemlerin adil, şeffaf ve hesap verebilir şekilde tasarlanıp işletildiğine güven duyması esastır. Bu güven, yapay zekânın sürdürülebilir gelişimi ve kurumsal sorumluluk kültürünün yaygınlaşması için kritik bir dayanak oluşturur.

Dolayısıyla bu rehber, etik ve veri yönetişimini yapay zekâ yönetişiminin temel bileşenleri olarak ele almakta; ancak bu iki alanın ötesine geçerek kurumsal yapı, karar süreçleri ve sürekli izleme kültürü gibi unsurları da iyi uygulama perspektifinden değerlendirmektedir.

Kapsam

AIPA, *"Yapay Zekâ Etiği ve Yönetişimi İyi Uygulamalar Rehberi*", AIPA'nın bağımsız bir girişimi olarak, Türkiye'de özel sektör için yapay zekâ sistemlerinin etik, güvenilir ve sorumlu kullanımına ilişkin temel ilkeler ile iyi uygulama örneklerini bir arada sunar.

Rehber, kurumların kendi süreçlerine uyarlayabilecekleri asgari ilkeleri ve uygulayıcılara yol gösterecek kontrol listeleriyle iyi uygulama önerilerini içerir. Türkiye'nin dijital kapasitesini ve bağımsızlığını destekleyen bu yaklaşım, uluslararası standartlar ve etik çerçevelerle uyumlu biçimde geliştirilmiştir.

Bu rehberde yer alan tüm öneriler ve temel ilkeler; yapay zekâ etiği, şeffaflık, açıklanabilirlik, yönetişim, veri kalitesi, adalet, hesap verebilirlik ve insan odaklılık prensiplerine dayanır.

Uluslararası Uyum

Bu rehber, AB Yapay Zekâ Tüzüğü (EU AI Act) başta olmak üzere OECD Yapay Zekâ Prensipleri, G7 Hiroshima AI Süreci İlkeleri, CNIL Yapay Zekâ Denetim Kılavuzu (Fransa), ICO Yapay Zekâ Denetim Çerçevesi (Birleşik Krallık), ISO/IEC 42001:2023 Yapay Zekâ Yönetim Sistemi Standardı, Küresel Yapay Zekâ Ortaklığı (GPAI) ilkeleri, UNESCO Yapay Zekâ Etiği Tavsiyeleri, Kişisel Verilerin Korunması Kanunu (KVKK) ve Birleşmiş Milletler İnsan Hakları Sözleşmeleri gibi uluslararası ve ulusal referans çerçevelerle uyumlu biçimde hazırlanmıştır.

Bu çerçeveler doğrultusunda, Türkiye'de geliştirilen yapay zekâ sistemlerinin küresel ölçekte güvenilirlik, adalet, şeffaflık ve sorumlu yönetişim ilkeleriyle uyum göstermesi amaçlanmaktadır. Ayrıca rehber, Türkiye'nin 2021-2025 Ulusal Yapay Zekâ Stratejisi'nde yer alan orantılılık, güvenlik, tarafsızlık, mahremiyet, hesap verebilirlik, veri egemenliği ve yönetişim prensipleriyle de uyumlu bir referans noktası sunar.

Her bölüm aşağıda sıralanan yapıyı izlemektedir:

Temel Uyum Gereklilikleri	İlgili mevzuat, standart veya uluslararası ilkelere dayanan asgari gereklilikleri tanımlar.
Tavsiye Edilen İyi Uygulamalar	Kurumsal olgunluğu artıran, değer katan ve örnek alınabilecek uygulamaları açıklar.
Kontrol Listesi	Kurumların kendi yapay zekâ sistemlerini öz değerlendirme yöntemiyle analiz edebilmeleri için kontrol noktaları sunar.

Bu bütüncül yapı sayesinde, yapay zekâ sistemlerinin yaşam döngüsünün her aşamasında — tasarım, veri toplama, model geliştirme, devreye alma, izleme ve sonlandırma süreçlerinde — yalnızca etik ilkelere uyum değil, aynı zamanda iş dünyasında güvenilirlik, şeffaflık ve sürdürülebilir yönetişim açısından örnek alınabilecek yaklaşımlar ortaya konur.

AIPA, bağımsız bir girişim olarak bu rehberi düzenli aralıklarla güncelleyerek, Türkiye'de özel sektörün dijital olgunluğunun ve etik yapay zekâ farkındalığının gelişimine katkı sağlamayı hedeflemektedir.

Yapay Zekâ Yönetişimi

Yapay zekâ yönetişimi, bir kurumun yapay zekâ sistemlerini tasarımdan dağıtıma, izleme ve sürekli iyileştirmeye kadar tüm yaşam döngüsü boyunca güvenli, etik, şeffaf ve hesap verebilir şekilde yönetmesini sağlayan kurumsal çerçevedir. Bu çerçeve; veri kalitesi, model doğruluğu, güvenlik, insan gözetimi, açıklanabilirlik, hukuki uyum ve operasyonel denetim gibi alanlarda standartlar ve süreçler tanımlayarak, yapay zekâ uygulamalarının iş hedefleriyle tutarlı, risk açısından kontrol edilebilir ve sürdürülebilir biçimde ölçeklenmesini mümkün kılar. YZ yönetişimi, teknolojinin sadece teknik olarak değil; organizasyonel, kültürel ve stratejik açıdan da sorumlu bir şekilde yönetilmesi için kurumun temel omurgasını oluşturur.

Gelenekselden Agentic Yapay Zekâya Doğru

Yapay zekâ yönetişimi, teknolojinin evrimine paralel olarak kurumsal yapılar ve iş modelleri için yeni gereksinimlere uyum sağlayacak biçimde dönüşmektedir. Önceden belirli amaçlara hizmet eden, sınırlı etki alanına ve veri türüne sahip sistemler için yeterli olan yaklaşımlar, günümüzün üretken (generative) ve/veya ajan-temelli (agentic) yapay zekâ sistemlerinde aynı etkiyi gösterememektedir. Yeni nesil yapay zekâ sistemleri; araç çağırma (tool use), hafıza yönetimi, planlama ve hedef odaklı özerk karar alma yetkinlikleriyle birlikte çalışmaktadır. Bu durum, sürekli öğrenme döngüleri, dış veri kaynaklarıyla etkileşim, çoklu ajan koordinasyonu ve zincirleme karar etkileri gibi yeni risk alanlarını beraberinde getirmektedir. Bu nedenle yönetişim yaklaşımlarının da bu dönüşüme paralel olarak daha esnek, izlenebilir ve dinamik bir yapıya kavuşması önem kazanmaktadır. Statik uyum kontrolleri yerine; sürekli izleme, risk bazlı değerlendirme ve operasyonel gözetim süreçlerinin benimsenmesi, kurumların yapay zekâ sistemlerini güvenli ve sürdürülebilir biçimde yönetmelerine katkı sağlar.

Sistem yaşam döngüsünün her aşamasında *(tasarım → dağıtım → izleme → iyileştirme → sonlandırma)* kayıt tutma, açıklanabilirlik, insan gözetimi, eşik/uyarı mekanizmaları ve kriz müdahale planlarının uygulanması, agentic sistemlerin güvenli işletimi için önerilen iyi uygulama alanları arasında yer almaktadır. Bu yaklaşım, kurumların değişen teknolojik koşullar altında kontrol edilebilirlik, ölçeklenebilir güven ve hesap verebilirlik ilkelerini sürdürebilmelerini destekler.

1. Veri Yönetişimi

Yapay zekâ sistemlerinin temelini oluşturan veri, yaşam döngüsünün her aşamasında etik ilkelere ve ilgili mevzuata uygun biçimde yönetilmelidir. Bu, verinin toplanması, işlenmesi, depolanması ve imha edilmesi süreçlerinin şeffaf, güvenli, kaliteli ve adil bir şekilde yürütülmesini gerektirir.

Veri yönetişimi, yalnızca mahremiyet ve güvenlik açısından değil, aynı zamanda yapay zekâ sistemlerinin çıktılarının güvenilirliği, doğruluğu ve tarafsızlığı açısından da kritik önemdedir. Uluslararası standartlar ve iyi uygulama çerçeveleri, yapay zekâ yaşam döngüsü boyunca güçlü

bir veri yönetimi ve koruma yapısının oluşturulmasını, güvenilir ve insan odaklı sistemlerin temel koşulu olarak tanımlar.

Bu bölümde, veri toplama, etiketleme, işleme ve saklama süreçlerinde dikkate alınması gereken asgari ilkeler ile tavsiye edilen iyi uygulamalar tanımlanmakta; kurumlara veri kalitesi, gizlilik ve bütünlüğü sağlamaya yönelik bir rehberlik sunulmaktadır. Bu yaklaşım, algoritmik kararların adil, izlenebilir ve açıklanabilir olabilmesi için gerekli kontrollerin sürecin en başında, yani "veri" aşamasında uygulanmasına destek olur.

1.1. Temel Uyum Gereklilikleri

Veri Temsiliyet Analizi: Eğitim verisi içinde korunan grupların temsiliyeti analiz edilmeli, eksik veya aşırı temsil durumları tespit edilmelidir (EU AI Act Ek IV, md.2(a); OECD *Inclusive Growth, Sustainable Development Principle*). Bu analiz, modelde sistematik önyargıların önlenmesi açısından temel bir kontrol adımıdır. Cinsiyet, yaş, etnik köken, lehçeler, coğrafi dağılım, eğitim düzeyi, gelir seviyesi, engellilik durumu, inanç temelleri, göçmenlik veya meslek profilleri gibi değişkenler dikkatle incelenmelidir.

Önyargı (Bias) Tarama Testleri: Veride tarihsel önyargı olup olmadığı, farklı gruplara karşı haksız veya dengesiz sonuç üretip üretmediği istatistiksel fark testleriyle kontrol edilmelidir. Anlamlı bir önyargı tespit edilirse, durum raporlanmalı ve giderici önlemler planlanmalıdır (EU AI Act Ek IV; ISO/IEC 42001 Ek B). Yüksek riskli YZ uygulamalarında bu testlerin düzenli yapılması önerilmektedir.

Veri Kaynağının Yasallığı ve Etikliği: Kullanılan tüm verilerin yasal olarak edinilmiş, gerektiği noktada ilgili bireylerin rızasının alınmış ve etik ilkelere uygun biçimde kullanılmış olması gerekir. İzinsiz web kazıma, gizli gözetim veya telif hakkı ihlali yoluyla elde edilen verilerin kullanımından kaçınılmalıdır (KVKK Md.4; GDPR Md.5; OECD *Human-Centered Values and Fairness Principle*).

Veri Kaynağının Kontrolü: Verinin kaynağı ve bağlamı net biçimde tanımlanmalı ve kayıt altına alınmalıdır. Örneğin; dijital form veya online anketler (interneti olmayan katılamaz), tarihsel veriler (toplumsal önyargılar bugüne taşınabilir), kamu kurum verileri (algoritma sadece sistemde var olanlar için kayıt üretir), kullanıcı içerikleri (sosyal medya, forum yorumları toksik dil içerebilir), eğitim platformları (yalnızca özel okul veya platforma ödeme yapan kurumlar için veri gelebilir) ve kurumsal CRM verilerinin (sadece işlem yapmış müşteriler için veri gelir) her biri farklı risk profili taşır.

Veri Dokümantasyonu: Kuruluşlar, yapay zekâ sistemlerinde kullanılan veri döngüsünü uçtan uca yönetmekle yükümlüdür. Her veri setinin kullanım amacı, türü, kaynağı, toplama yöntemi, izin durumu, saklama süresi ve bilinen sınırlamaları açıklanarak veri gereksinimi tanımlama

dokümanı veya veri envanteri dokümanı hazırlanmalıdır (ISO/IEC *5338 Yapay Zekâ Yaşam Döngüsü*; OECD *Transparency Principle*; ISO/IEC 42001 Md.8). Bu dokümantasyonun, yüksek riskli sistemlerde teknik dosyanın bir parçası olarak tutulması önerilmektedir (EU AI Act Ek IV).

Kişisel Verilerin Korunması: Eğitim ve test verilerindeki kişisel veriler ilgili mevzuata uygun şekilde anonimleştirilmeli veya maskelenmeli; gerekliyse veri sahiplerinin açık rızası alınmalıdır (KVKK, GDPR – *Privacy and Data Governance*). Özellikle sağlık, biyometrik veya genetik veriler gibi hassas bilgilerin işlenmesinde güçlü şifreleme, erişim sınırlandırması vb. ek teknik ve organizasyonel önlemlerin alınması ve yasal dayanak oluşturulması gereklidir.

Veri Kalitesi ve Tutarlılığı: Veri setinde eksik, tutarsız veya yanlış etiketlenmiş kayıtların taranması ve temizlenmesi önerilir (ISO/IEC 42001 Md.10; EU AI Act Md.10). Kalite sorunları model performansını ve adaletini etkileyebileceğinden, kurumların asgari veri doğrulama prosedürleri tanımlaması önemlidir.

Veri Güncelliği ve Sapma (Drift) Kontrolü: Modelin zamanla bayatlamaması ve performans düşmemesi için düzenli veri güncelleme planı oluşturulmalıdır (ISO/IEC 42001 *Continual Improvement Principle*; NIST AI RMF *Monitor function*). Veri sapması tespit edildiğinde, güncelleme veya yeniden eğitim sürecinin başlatılması önem arz etmektedir.

Hassas Veri Kullanım Onayı: Eğitim sürecinde ırk, etnik köken, din veya sağlık gibi hassas veriler kullanılıyorsa, bu kullanımın gerekçesi kurum içi etik kurul veya üst yönetim tarafından onaylanmalı ve kayıt altına alınmalıdır (EU AI Act Md.10; OECD *Fairness Principle*).

Özel Nitelikli Veri Koruması: Sağlık, biyometrik, genetik veya benzeri özel nitelikli verilerin işlenmesinde ek teknik ve organizasyonel önlemler (örneğin güçlü şifreleme, erişim kısıtlaması) uygulanmalıdır (GDPR Md.9; EU AI Act Md.10). Bu veriler için kurum içi politika ve ek koruma katmanları tanımlanması önerilir.

Kırılgan Grupların Korunması: Çocuklar, yaşlılar veya diğer kırılgan gruplara ait veriler işleniyorsa, ek koruma önlemleri (örneğin işleme amaç sınırlandırması, ek rıza mekanizması) uygulanmalıdır (GDPR Recital 38; UNESCO 2021 *AI Recommendations* Md.23; EU AI Act Mad.5/1).

Veri Minimizasyonu: Toplanan veriler yalnızca amaçla doğrudan ilişkili ve gerekli olmalıdır. Amaçla ilgisi olmayan kişisel veriler toplanmamalı, gereksiz özellikler derhal veri setinden çıkarılmalıdır (GDPR Md.5/1(c) – *Data Minimization Principle*; EU AI Act Md.10 – *Data Necessity Requirement*).

Veri Soyağacı ve Bütünlüğü: Verinin yaşam döngüsü boyunca kaynağı, toplama yöntemi ve hak/lisans durumu veri soyağacı kayıtlarında tutulmalıdır (EU AI Act Ek IV; NIST AI RMF *Govern/Map*). Bu kayıtlar, gerektiğinde verinin kökeni ve bütünlüğü hakkında şeffaflık sağlar.

Veri Rol ve Sorumlulukları: Yapay zeka çözümlerinde kullanılan veri kümeleri için rol ve sorumluluk sahipliği tanımlanmalıdır. Her kritik veri varlığı için veri sahibi (data owner) ve veri sorumlusu (data steward) atanmalı; bu kişilerin veri doğruluğu, bütünlüğü, güncelliği ve imha süreçlerinden sorumlu olduğu iç düzenlemelerde açıkça belirtilmelidir (ISO/IEC 5338 *Yapay Zekâ Yaşam Döngüsü*).

Silme Taleplerine Uyum (Machine Unlearning): Bireylerin kendi verilerinin sistemden silinmesi taleplerine uyum için, yalnızca veritabanından silme değil, model üzerindeki etkilerin de kaldırılmasına yönelik bir prosedür bulunmalıdır (GDPR Md.17; ISO/IEC 42001 *Corrective Action*). Örneğin; bir kullanıcı verisinin silinmesi istendiğinde, model gerekiyorsa yeniden eğitilerek o verinin etkisi yok edilmelidir.

Telif ve Lisans Uyumu: Eğitim verileri toplanırken telif hakkı ve lisans kısıtlamalarına dikkat edilmelidir. İzinsiz veya lisanssız içerikler filtrelenmeli; telif korumalı içeriklerin kullanımı durumunda gerekli izinler alınmalıdır (EU AI Act Md.53(1)(c) – *obligation on copyright and data management*; GPAI ve EU *"Code of Practice on Disinformation").*

1.2. Tavsiye Edilen İyi Uygulamalar

Muhtelif ve Kapsayıcı Veri: Veri setlerinin mümkün olduğunca geniş ve çeşitli bir temsiliyete sahip olması önerilir. Azınlık grupların veya marjinal toplulukların yeterince temsil edilmesi, katılımcı adalet ilkesi açısından önemlidir. Gerektiğinde, az temsil edilen gruplar için sentetik veri üretimi gibi dengeleme yöntemleri kullanılabilir (ISO/IEC TR 24027 – *Bias Measurement Standards*).

Dış Denetim ve Doğrulanabilirlik: Kurumlar, veri yönetimi süreçlerini düzenli aralıklarla bağımsız denetimlere tabi tutmalıdır. Harici bir etik denetçi veya uzman kurul, veri kalitesi ve önyargı risklerini değerlendirerek olası kör noktaları tespit edebilir. Veri setlerine ilişkin doğrulama mekanizmaları (audit trail) üçüncü taraflarca denetlenebilir olmalı, böylece verinin kaynağı ve kalitesi gerektiğinde bağımsız doğrulamaya açık hale gelir (OECD – *Multi-Stakeholder Approach*).

Gizliliği Artırıcı Teknikler: Ham kişisel veriler gereksiz yere paylaşılmamalı; bunun yerine anonimleştirme, şifreleme, diferansiyel gizlilik veya federe öğrenme (federated learning) gibi yöntemlerle veri gizliliği güçlendirilmelidir. Bu uygulamalar, yalnızca yasal uyum sağlamakla kalmaz, aynı zamanda kurumun güvenilirlik ve etik sorumluluk algısını da güçlendirir (CNIL – *AI Data Privacy Recommendations*).

Veri Yaşam Döngüsü Yönetimi: Veri toplama, etiketleme, depolama, güncelleme, imha etme ve kullanılabilirlik adımları ile kalite kontrolleri için kurum içinde net prosedürler oluşturulmalıdır. Örneğin; yapay zekâ sistemlerinde kullanılan tüm veri akışlarını operasyonel seviyede gösteren görsel bir "Veri Yaşam Döngüsü Haritası" oluşturulabilir. Özellikle veri etiketleme aşamasında

tutarlılık yönergeleri ve hata-çelişki çözüm prosedürleri tanımlanmalıdır (ISO/IEC 5338 Md. 6.2.3; NIST AI RMF – *Measure / Manage*). Etiketleyicilere önyargı farkındalığı ve ayrımcılık eğitimi verilmesi önerilir.

Veri Uyum Senaryoları ve Kullanım Amacı Doğrulama Mekanizması: Kurumlar yapay zekâ sistemlerinde kullanılan her veri seti için uygun kullanım senaryoları ve yasak kullanım alanlarını tanımlayan bir doğrulama mekanizması oluşturmalıdır. Veri setinin, ilk belirlenen kullanım amacı dışında bir işlem için değerlendirilmesi durumunda otomatik veya manuel bir kontrol tetiklenmeli; amaç genişlemesi (function creep) riskini önlemek için kayıt altına alınan bir onay süreci çalışmalıdır. Bu yaklaşım, veri minimizasyonu ve amaç sınırlılığı ilkelerinin pratikte kurumsal düzeyde uygulanmasını pekiştirir (ISO/IEC 5338, Md. 5.4.2).

Versiyonlama ve Değişiklik Takibi: Veri setlerinde yapılan her değişiklik versiyonlanarak kayıt altına alınmalıdır. Yeni eklenen, çıkarılan veya temizlenen verilerin tarihçesi tutulduğunda, modelin hangi veriyle eğitildiği izlenebilir hale gelir. Bu yaklaşım şeffaflık ve hesap verebilirlik ilkelerini destekler (OECD – *Transparency Principle*).

Açık Veri ve Paylaşım: Kurumlar, yasal sınırlar dâhilinde ve gerekli anonimleştirmeyi sağlayarak açık veri paylaşımı kültürünü destekleyebilir. Bu tür paylaşımlar, sektörel ekosistemde tekrarlanabilirliği artırır ve yerli yapay zekâ geliştirme kapasitesinin güçlenmesine katkı sağlar. Paylaşım yapılırken lisans kısıtları ve etik ilkeler açıkça belirtilmelidir.

Sürekli Veri Yönetişimi: Veri yönetişimi yalnızca periyodik denetimlerle sınırlı olmamalı, kesintisiz izleme ve kanıt üretimi yaklaşımını benimsemelidir. Bu model, düzenleyici ve etik yükümlülüklerin sadece raporlama dönemlerinde değil, gerçek zamanlı olarak izlenmesini sağlar (NIST AI RMF – *Monitor*; ISO/IEC 42001 – *Continuous Improvement*). Bu dönüşüm; periyodik denetimlerin yerini anlık izleme sistemlerine, manuel dokümantasyonun yerini otomatik veri soyağacı takibine, reaktif yaklaşımların yerini öngörücü analitik uyarı mekanizmalarına bırakmaktadır. Böylece “görünmeyen veri hareketleri” gerçek zamanlı izlenebilir ve kayıt dışı veri kullanımı proaktif biçimde önlenebilir.

Multimodal Veri Yönetişimi: Modern yapay zekâ sistemleri artık metin, görüntü, ses, video ve sensör verilerini birlikte işlemektedir. Bu nedenle veri yönetişimi, her veri türünü ayrı ayrı kontrol etmek yerine modlar arası tutarlılık, adalet ve açıklanabilirlik sağlamaya odaklanmalıdır. Örneğin; Metin-görsel önyargı haritalamaları ile metindeki önyargıların görsel çıktılara yansıması analiz edilebilir, ses-görsel stereotip tespiti ile ses tonu ve görsel temsil arasındaki önyargı ilişkileri incelenebilir. Açıklanabilirlik yöntemleri de multimodal yapılara uyarlanmalı; her veri tipinin karara etkisi ayrı ayrı şeffaf biçimde gösterilmelidir (ör. *Cross-Modal Attribution* yöntemleriyle). Bu sayede çoklu veri kaynakları kullanan YZ sistemlerinde adalet, şeffaflık ve bütüncül yönetişim ilkeleri güvence altına alınır (EU AI Act Md.10; OECD *Principle* 1.2; ISO/IEC JTC 1/SC 42 *Technical Report*).

1.3. Veri Yönetişimi için Kontrol Listesi

Kurumlar, veri yönetimi süreçlerinde aşağıdaki kontrol sorularını kullanarak öz değerlendirme yapabilir. Her bir madde, veri yönetişiminin iyi uygulama alanlarına ilişkin bir kontrol noktası olarak tasarlanmıştır.

* **Veri Temsiliyet Analizi Yapıldı mı?** - Veri setinde hassas veya dezavantajlı grupların temsiliyeti analiz edildi mi? Eksik veya aşırı temsil durumları belirlendi ve raporlandı mı? (EU AI Act Ek IV-2(a); OECD *Inclusivity Principle*)
* **Önyargı Tarama Testi Uygulandı mı?** – Modeli eğitirken kullanılan verilerde geçmişten gelen önyargılar kontrol edildi mi? Belirli bir grubun sistematik olarak dezavantajlı çıktığı durumlar test edildi mi? Bunun için %80 kuralı (bir grubun seçim oranı diğer grubun %80'inden az olmamalı) veya istatistiksel karşılaştırmalar gibi yöntemler kullanıldı mı? Tespit edilen önyargılar raporlanıp düzeltici önlem planlandı mı? (EU AI Act – *Data Governance*; ISO/IEC 42001 Ek B)
* **Veri Kaynağı Yasal/Etik mi?** – Veriler yetkili sahibinden izinle mi elde edildi? İzinsiz web kazıma, gizli gözetim veya telif hakkı ihlali yoluyla toplanan veri bulunuyor mu? (KVKK Md.4; GDPR *Core Principles*, OECD *Human-centered Values Principle*)
* **Veri Belgeleri Hazırlandı mı?** – Her veri seti için türü, kaynağı, toplama yöntemi, izin durumu ve bilinen sınırları açıklayan bir veri yaşam döngüsü haritası veya veri envanteri dokümanı mevcut mu? Bu doküman veri işleme adımlarının tamamını kapsayacak şekilde düzenli olarak güncelleniyor mu? (ISO/IEC 5338 Md. 6.2.3.; OECD *Transparency Principle,* ISO/IEC 42001 Md.8)
* **Kişisel Veri Koruması Sağlandı mı?** – Eğitim ve test verilerinde kişisel veriler anonimleştirildi veya maskelendi mi? Gerekli durumlarda ilgili kişilerin açık rızası alındı mı? (KVKK; GDPR – *Privacy and Data Governance*)
* **Veri Kalitesi ve Tutarlılığı Doğrulandı mı?** – Veride eksik, tutarsız veya yanlış etiketlenmiş örnekler tespit edilip temizlendi mi? (ISO/IEC 42001 – *Data Quality Control*; EU AI Act – *Data Compliance Requirement*)
* **Veri kullanım amacı doğrulama işleyişi mevcut mu? –** Veri setlerinin yalnızca tanımlanmış kullanım amacı kapsamında işlendiğini doğrulayan bir amaç sınırlılığı kontrolü uygulanıyor mu? Amaç genişlemesini önleyen bir mekanizma mevcut mu? (ISO/IEC 5338 Md. 5.4.2.)
* **Veri Güncellik ve Sapma (Drift) Kontrolü Var mı?** – Modelin zamanla performans kaybı yaşamaması için düzenli veri güncelleme planı tanımlandı mı? Kavram sapması (concept drift) izleme mekanizması uygulanıyor mu? (ISO/IEC 42001 – *Continuous Improvement*; NIST AI RMF – *Monitor function*)

* **Hassas Özellik Kullanımı Onaylandı mı?** – Eğitim sürecinde ırk, sağlık, inanç gibi hassas verilerin kullanımı gerekçelendirilip üst yönetim veya etik kurul tarafından onaylandı mı? (EU AI Act Md.10 – *technic documentation*; OECD *Fairness Principle)*
* **Özel Nitelikli Veri Koruması Var mı?** – Sağlık, biyometrik veya dini içerikli özel nitelikli verilerin işlenmesi için yasal dayanak oluşturuldu mu? Ek koruma önlemleri (şifreleme, erişim kısıtlaması vb.) tanımlandı mı? (GDPR Md.9; EU AI Act Md.10)
* **Veri Etiketleme ve Hata Çözüm Prosedürü Mevcut mu?** – Etiketleme yapanlar için yönergeler tanımlandı, tutarsızlık durumlarında bir arabulucu/hata çözüm süreci kayıt altına alındı mı? (NIST AI RMF – *Measure/Manage*)
* **Veri Minimizasyonu ve Amaç-Eşleşme Kontrolü Yapıldı mı?** – Toplanan her veri alanının (örneğin yaş, cinsiyet, gelir) gerçekten gerekli olduğu doğrulandı mı? Amaçla ilişkisi olmayan veriler elendi mi? (GDPR Md.5(1)(c) – *Data Minimization*; EU AI Act Md.10)
* **Veri Soyağacı ve Bütünlük Zinciri Tutuluyor mu? –** Verinin kaynağı, toplama yöntemi, lisans ve hak bilgileri, değişiklik geçmişi kayıt altına alındı mı? (EU AI Act Ek IV – *data lineage*; NIST AI RMF – *Govern/Map*)
* **“Modelden Silme” (Unlearning) Planı Var mı?** – Bir kullanıcının verisinin sistemden silinmesi istendiğinde, sadece veri tabanından değil modelin öğreniminden de kaldırılması için bir prosedür mevcut mu? (GDPR Md.17 – *Right to be forgotten*; ISO/IEC 42001 – *Corrective Action Principle*)
* **Veri Telif & Lisans Uyum Kontrolü Yapıldı mı?** – Eğitim verilerinde telif hakkı koruması veya lisans kısıtına sahip içerikler belirlendi mi? Bu içerikler için izin alınarak veya filtreleme politikaları uygulanarak uyum sağlandı mı? (EU AI Act Md.53(1)(c); EU ”*Code of Practice on Disinformation”- Copyrights*)

2. Model Geliştirme ve Kontrolleri

Yapay zekâ modelinin geliştirilmesi, eğitilmesi ve doğrulanması süreçleri, bir kurumun etik duruşunu ve teknolojik olgunluğunu en doğrudan yansıtan aşamalardır. Bu bölüm, yalnızca teknik gereklilikleri değil, aynı zamanda adalet, açıklanabilirlik, güvenlik, insan gözetimi ve hesap verebilirlik ilkelerini de bütüncül biçimde ele alır. Model geliştirme süreci, veri temelli önyargıların azaltılması, performans metriklerinin adil şekilde ölçülmesi ve sonuçların şeffaf biçimde izlenmesi için sağlam kontrol mekanizmaları gerektirir.

Bu kapsamda, model yaşam döngüsünün her aşamasında (tasarım, geliştirme, test, devreye alma, bakım ve sonlandırma) versiyonlama, dokümantasyon, açıklanabilirlik ve risk yönetimi

süreçlerinin tanımlanması ve uygulanması önemli görülmektedir. Ayrıca, agentic (özerk) sistemlerin ve üretken yapay zekâ modellerinin yaygınlaştığı günümüzde, modellerin yalnızca doğru sonuç üretmesi değil, kontrol edilebilir, izlenebilir ve insan tarafından durdurulabilir olması da kritik öneme sahiptir.

Bu bölüm, model tasarımı, etik etki değerlendirmesi, performans doğrulaması ve insan gözetimi gibi başlıklarda asgari ilkeleri ve tavsiye edilen iyi uygulama prensiplerini ortaya koyar. Amaç, modelin yalnızca teknik olarak güçlü değil; aynı zamanda etik, adil, güvenilir ve sürdürülebilir biçimde çalışmasını desteklemektir.

2.1. Temel Uyum Gereklilikleri

Model Amaç ve Kapsam Tanımı: Geliştirilen her modelin ne amaçla kullanılacağı ve hangi sınırlar içinde çalışacağı yazılı olarak tanımlanmalıdır. Bu tanım, modelin görevini, kullanım senaryolarını ve kullanılmayacağı durumları da içermelidir. (ISO/IEC 42001 Md.5 – *Context of the organization and determination of objectives*; OECD *Transparency Principle*)

Model Kütüğü (Versiyon Kontrolü): Tüm modeller için bir Model Envanteri/Kütüğü oluşturulması önerilir. Her modelin versiyon numarası, eğitildiği tarih, kullanılan veri seti versiyonu, geliştiren sorumlular, model hiperparametreleri ve temel performans metrikleri kayıt altına alınmalıdır (ISO/IEC 42001 Md.7.5 – *documentation management*; NIST AI RMF *Govern*; OECD *Accountability Principle*).

Model Performansının Adil Ölçümü: Modelin doğruluk, hata oranı gibi genel performans metriklerinin yanı sıra korunan gruplar (örneğin cinsiyet, etnik köken) bazında performansı da ölçülmelidir. Farklı gruplar için hata oranları veya karışıklık matrisi karşılaştırmaları yapılmalı, adalet düzeyi analiz edilmelidir (EU AI Act Ek IV-4 – *Provisions on the documentation of accuracy, explainability and robustness metrics*).

Adalet Değerlendirmesi: Model kararlarının herhangi bir grup aleyhine sistematik bir sapma üretip üretmediğini kontrol etmek için adalet (fairness) değerlendirmesi yapılmalıdır. “Disparate impact analysis” (*Ayrık etki analizi: imtiyazsız grubun, imtiyazlı gruba göre daha az olumlu sonuç alması*), “equal opportunity” (E*şit fırsat değerlendirmesi: farklı gruplardan nitelikli kişilerin iyi bir sonuç için aynı şansa sahip olması*), “intersectional fairness” (*Kesişimsel adalet ölçütleri: bileşik ön yargıların kontrol edilmesi, modelin performansının tek bir faktörden değil; ırk, cinsiyet, yaş vb. diğer faktörlerin kombinasyonları üzerinden değerlendirilmesi*) gibi kriterler teknik dokümana eklenmeli ve gerekli durumlarda modelde veya veri ön işleme adımlarında gerekli düzeltmeler yapılmalıdır (OECD ”*Fairness” Principle*; ISO/IEC 42001 Ek A – *Ethical impact assessment recommendation*).

Açıklanabilirlik (XAI): Modelin her kararı için mümkün olduğunca anlaşılır açıklama üretebilen mekanizmalar kurulmalıdır. Kararların hangi girdilere dayandığı ve sonucu etkileyen temel faktörler kayıt altına alınmalı; gerektiğinde bu bilgiler ilgili taraflara sunulmalıdır. Yüksek riskli sistemlerde bu uygulama yasal düzenlemelerle zorunlu hale gelebilir (EU AI Act Md.13 – *Explanation obligation*; OECD *Transparency & Explainability Principle*).

Güvenlik ve Sağlamlık Testleri: Model, beklenmedik girdilere ve bilinen saldırı tekniklerine karşı test edilmelidir. Adversarial (düşman) örneklerle modelin dayanıklılığı test edilmeli; siber güvenlik açıkları belirlenip giderilmelidir. Bu testler, modelin veri enjeksiyonu ve model hırsızlığı gibi risklere karşı korunmasını da kapsamalıdır (EU AI Act Ek IV-5 – *security requirement*; ISO/IEC 42001 Md.6 & 8 – *risk management*).

Hiperparametre ve Özellik Seçimi Kayıtları: Model geliştirme sürecinde denenen hiperparametre kombinasyonları ve seçilen/çıkarılan özellikler (features) kayıt altına alınmalıdır. Bu kayıt, modelin yeniden üretilebilirliği (reproducibility) ve denetlenebilirliği açısından kritik öneme sahiptir (ISO/IEC 42001 Md.7.5 – *documentation*; ISO/IEC 5338 Md. 6.3.4.; OECD *Accountability Principle*).

İnsan Onayı ve Gözetimi: Modelin kararları insanları doğrudan etkiliyorsa (örneğin işe alım, kredi, sağlık vb. alanlarda), sistem tasarımında insan onayı veya müdahalesi mekanizması bulunmalıdır. Yüksek riskli kararlar insan onayı olmadan uygulanmamalı; insan operatörlerin sorumlulukları açıkça tanımlanmalıdır (EU AI Act Md.14 – *human oversight requirements*; OECD *Human-centered values principle*).

Yasal Kısıtlamalara Uyum: Modelin gerçekleştirdiği işlevin, yürürlükteki düzenlemeler ve etik çerçevelerle uyumlu olması sağlanmalıdır. AB Yapay Zekâ Tüzüğü'nde yasaklanan uygulamalar (örneğin sosyal puanlama, gerçek zamanlı biyometrik izleme gibi) geliştirilmeyip; riskli alanlarda faaliyet tespit edilirse süreç durdurulmalıdır. Türk mevzuatında yer alan özel hükümler de (örneğin biyometrik veri kullanımı gibi) ayrıca dikkate alınmalıdır (EU AI Act Bölüm II – *prohibited AI practices* ; KVKK ve ilgili mevzuatlar).

2.2. Tavsiye Edilen İyi Uygulamalar

Operasyonel İnsan Gözetimi: İnsan gözetimi mekanizmasının pratikte etkili olabilmesi için sistem tasarımında "kill-switch" (acil durdurma düğmesi) veya "manual override" (manuel geçersiz kılma) yetenekleri bulunmalıdır. İnsan operatör, sistemin hatalı karar verdiğini fark ettiğinde süreci durdurabilmeli veya kararı geçersiz kılabilmelidir. Bu yetkilerin işlevsel kalabilmesi için düzenli tatbikatlar yapılması iyi bir uygulamadır (EU AI Act Md.14).

Üyelik Çıkarım ve Model Sızıntı Testleri: Modelin eğitim verisinde yer alan gizli bilgileri sızdırma riskine karşı membership inference (üyelik çıkarım) testleri yapılması önerilir. Özellikle büyük dil modelleri (LLM) ve generative sistemler, eğitim verilerini hatırlayabilir veya yeniden üretebilir. Bu riskleri azaltmak için "differentially private training" gibi yöntemlerin uygulanması iyi bir uygulamadır (ISO/IEC 27001; NIST Guide– *privacy & security*).

Dağılım Dışı Veri (Out-of-Distribution) Tespiti: Model, kendi eğitim verisi dağılımının dışında kalan girdilerle karşılaştığında bunu tespit edip uyarı verebilmelidir. "Out-of-Distribution (OOD)" dedektörleri veya belirsizlik tahminleme yöntemleri kullanılarak modelin emin olmadığı durumlar belirlenebilir. Bu sayede, alışılmadık verilerde karar vermeden önce sistemin insan incelemesine yönlendirilmesi sağlanır (NIST AI RMF *"Measure" Function* – OOD *planning*).

Açıklama Araçlarının Etkin Kullanımı: Model kararlarını açıklamak için SHAP ve LIME (Modelin her bir tahmininde hangi girdinin ne kadar etkili olduğunun gösterilmesi. Örn; kararlar üzerinde modelin en önemli 5 özelliği listelenir ve listede beklenmedik, hassas bir özellik varsa vurgulanır.) gibi araçlar kullanılabilir. Ancak bu araçlar, kullanıcı profiline ve sistemin amacına uygun biçimde tasarlanmalıdır. Teknik detaylara boğulmak yerine, kararı etkileyen en önemli faktörleri sade, kullanıcı dostu bir arayüzle sunmak güveni artırır (OECD "*Explainability" Principle*).

Etik Etki Değerlendirmesi: Her yeni model için, sosyal ve çevresel etkileri analiz eden bir etik etki değerlendirmesi (Ethical Impact Assessment) yapılması önerilir. Bu değerlendirme, toplumsal değerler, kültürel farklılıklar ve çevresel sürdürülebilirlik gibi faktörleri kapsamalıdır (EU AI Act Md.29; ACM/IEEE *Ethics Guidelines*; ISO/IEC 42001:2023 Ek A).

Enerji Verimliliği ve Sürdürülebilirlik: Modelin eğitimi ve çalıştırılması sırasında enerji tüketimi ve karbon ayak izi izlenmelidir. Model mimarisi optimizasyonları, donanım verimliliği ve yeşil bilişim uygulamalarıyla çevresel etki azaltılabilir. Enerji tüketiminin düzenli raporlanması ve iyileştirme hedeflerinin tanımlanması önerilir (OECD *AI Principles – sustainable development*; UNESCO *AI Ethics Recommendations*).

Generative AI İçerik İşaretleme: Eğer model üretken bir yapay zekâ sistemi ise, oluşturduğu içeriklere filigran (watermark) veya benzeri işaretler eklenerek kaynağın doğrulanabilir olması sağlanmalıdır. C2PA (Content Provenance and Authenticity) standartlarının kullanımı, sahte içerikle mücadelede güçlü bir önlemdir. Bu uygulama henüz yasal olarak zorunlu olmasa da uluslararası iyi uygulama olarak kabul edilmektedir.

Rollback ve Sürüm İzolasyonu: Yeni bir model versiyonu devreye alınmadan önce eski sürümle paralel test (shadow mode) yapılmalıdır. Beklenmedik hatalar tespit edilirse hızlıca önceki sürüme dönülebilmesi için rollback planı hazırlanmalıdır. Bu yöntem, değişiklik yönetimi süreçlerinde kalite ve güvenilirliği artırır.

Agentic AI ve Multi-Agent Model Kontrolü: Otonom karar alma ve eylem yürütme yeteneğine sahip agentic AI sistemleri için klasik kontrol mekanizmalarına ek önlemler alınmalıdır. Birden fazla yapay zekâ ajanının birlikte çalıştığı sistemlerde güvenli iletişim ve koordinasyon sağlanmalıdır. Bunun için; her ajanın güvenilir dijital kimliğe sahip olması, mesajların kriptografik imzalarla doğrulanması, uçtan uca şifreleme ile iletişimin korunması önerilir. Ayrıca çoklu ajan yapılarında "Practical Byzantine Fault Tolerance" veya "Proof of Authority" gibi algoritmalar kullanılarak kötü niyetli ajanların sistemi bozması engellenmelidir (Zhao, Y., Li, H., & Chen, X. (2025). *Byzantine-resilient coordination mechanisms for agentic AI Systems*). Buna ek olarak, agentic sistemlerde kademeli karar onay (escalation) süreçleri tanımlanmalıdır. Ajanlar arası çapraz kontrol, üst seviye ajan denetimi, insan uzman müdahalesi ve gerekirse yönetim onayı gibi artan düzeylerdeki kontroller, yüksek riskli kararların güvenliğini güçlendirir. Bu sayede özerk sistemlerde hata olasılığı en aza indirilir.

2.3. Model için Kontrol Listesi

Aşağıdaki kontrol noktaları, model geliştirme ve değerlendirme süreçlerinde kurumlar tarafından öz denetim amacıyla uygulanabilir:

* **Model Amacı ve Kapsamı Tanımlandı mı? -** Modelin amacı, sınırları ve kullanım alanları yazılı biçimde belirlendi mi? (ISO/IEC 42001 *"Context of organization";* OECD *"Transparency")*

* **Model Versiyonları Kayıt Altında mı?-** Geliştirilen her model için versiyon numarası atandı mı? Model kayıt defterine eğitim tarihi, veri seti versiyonu, sorumlu kişiler, hiperparametreler ve performans metrikleri kaydedildi mi? (ISO/IEC 42001 – *documentation*; NIST AI RMF – *governance*; OECD – *accountability*)

* **Model Performansı Her Grup için Ölçüldü mü? -** Doğruluk veya hata oranları korunan gruplar (cinsiyet, yaş, etnik köken vb.) bazında raporlandı mı? (EU AI Act Ek IV-4 – *performance & robustness metrics*)

* **Adalet Değerlendirmesi Yapıldı mı? -** "Disparate impact", "equal opportunity", "intersectional fairness" gibi ölçütlerle modelin adil çalışıp çalışmadığı değerlendirildi ve sonuçlar teknik dokümana eklendi mi? (OECD *Fairness Principle*; ISO/IEC 42001 Ek A – *ethical impact assessment guide*)

* **Açıklanabilirlik Mekanizması Sağlandı mı? -** Modelin karar mantığını açıklayan bir yapı mevcut mu? Her kararın gerekçesi uygun biçimde loglanıyor mu? (EU AI Act Md.13 – *Explanation Requirement*; OECD "*Transparency & Explainability" Principles*)

* **Sağlamlık ve Güvenlik Testleri Yapıldı mı? -** Model, adversarial (hasım) örnekler ve olası saldırı senaryolarıyla test edilip güvenlik açıkları değerlendirildi mi? (EU AI Act Ek IV-5 – *security requirements*; ISO/IEC 42001 – *risk management*)

* **Hiperparametre ve Özellik Seçimi Belgeli mi? -** Modelin eğitiminde denenen hiperparametreler, seçilen veya hariç tutulan özellikler (features) kayıt altına alındı mı? (ISO/IEC 42001 – *documentation*; OECD "*Reproducibility" recommendation*)

* **İnsan Onayı / Gözetimi Mekanizması Var mı? -** Yüksek riskli kararlar için bir insan onayı veya izleme noktası tanımlandı mı? (EU AI Act Md.14 – *human oversight*; OECD "*Human-centered" principle*)

* **Model Yasal Kısıtlamalara Uygun mu? -** Modelin gerçekleştirdiği işlevlerin yasa dışı veya yasaklı bir uygulamaya girmediği doğrulandı mı? (EU AI Act Bölüm II – *prohibited AI practices*)

* **İnsan Gözetimi Operasyonel mi? -** İnsan gözetimi yalnızca belgelerde değil, pratikte uygulanabilir mi? Sistem kararlarını durdurmak veya geri almak için operatörlerin **"kill-switch"** yetkisi tanımlandı ve test edildi mi? (EU AI Act Md.14 – *human oversight*)

* **Model Gizlilik Sızıntı Testi Yapıldı mı? -** Modelin eğitim verisini ifşa etme (membership inference attack) veya tersine mühendislik yoluyla kopyalanma riski değerlendirildi mi? Bu riskler için saldırı simülasyonu ve azaltma (mitigation) teknikleri uygulandı mı? (EU AI Act Md.15 – *robustness & security requirements*)

* **Model Dağılım Dışı Veriye Karşı Duyarlı mı? -** Sistem, eğitim dağılımı dışında kalan bir girdiyle karşılaştığında bunu tespit edip uyarı üretebiliyor mu? Belirsizliği yüksek çıktılarda insan incelemesini tetikleyen bir OOD mekanizması mevcut mu? (NIST AI RMF – *Measure Function*)

* **Açıklama Araçları Amaç-Uygun mu? -** Modelin açıklanabilirliğinde kullanılan araçlar (LIME, SHAP vb.) hedef kullanıcı profiline uygun seçildi mi? Sonuçlar anlaşılır biçimde sunuluyor mu? (NIST AI RMF – *explainability assessment*)

* **Enerji / CO_2 ve Verimlilik İzleniyor mu? -** Modelin eğitimi ve çalıştırılması sırasında enerji tüketimi, işlem süresi ve donanım verimliliği izleniyor mu? Raporlanarak optimizasyon hedefleri belirleniyor mu? (OECD AI Principles – *sustainability*)

* **Generative AI İçerik İşaretlemesi Yapılıyor mu? -** Model üretken yapay zekâ içeriği oluşturuyorsa, içeriğe "YZ ile üretildi" ibaresi veya filigran ekleniyor mu? C2PA gibi standartlar uygulanıyor mu? (G7 Hiroshima AI Process; C2PA standard)

* **Rollback ve Sürüm İzolasyonu Planı Var mı? -** Model güncellemelerinde olası hatalara karşı eski sürüme dönüş (rollback) planı hazır mı? Yeni model devreye alınmadan önce gölge mod (shadow mode) testleri yapılıyor mu? (EU AI Act Ek IV – *change management*; NIST AI RMF)

3. İzleme, Denetim ve KPI Yönetimi

Yapay zekâ sistemlerinin güvenilirliğini ve etik bütünlüğünü sürdürebilmek, yalnızca geliştirme aşamasındaki uygulamalarla sınırlı değildir. Bu güvenin korunması, sistemin canlı kullanımı sonrasında da sürekli izleme, denetim ve iyileştirme mekanizmalarının etkin biçimde işletilmesine bağlıdır.

Bu bölüm, yapay zekâ sistemlerinin operasyonel döngüsü boyunca ortaya çıkabilecek risklerin erken tespiti, önleyici müdahale planlarının uygulanması ve performansın ölçülebilir metriklerle izlenmesi süreçlerine odaklanır. Amaç, teknik doğruluğun ötesinde, sistemin etik, adil, güvenli ve sürdürülebilir biçimde işlemesini sağlamaktır.

İzleme süreci yalnızca teknik parametreleri değil;

Doğruluk, adalet, sağlamlık, enerji verimliliği, kullanıcı güvenliği ve içerik bütünlüğü gibi performans göstergelerinin düzenli takibini,

İnsan gözetimi, kullanıcı şikâyet yönetimi, olay kayıtları, hata ve önyargı geri bildirim döngülerinin oluşturulmasını,

Düzenli etik denetim raporlarıyla şeffaf bir hesap verebilirlik mekanizmasının kurulmasını da kapsar.

Bu yaklaşım, kurumların yalnızca hataları sonradan tespit eden yapılar olmaktan çıkıp, riskleri önceden fark eden proaktif denetim kültürüne geçişini destekler.

Bölümün sonunda yer alan Kilit Performans Göstergeleri (Key Performance Indicators – KPI) önerileri, izleme faaliyetlerini ölçülebilir, denetlenebilir ve sürekli gelişmeye açık bir çerçeveye dönüştürmeyi hedefler. Böylece yapay zekâ sistemleri yalnızca “doğru çalışan” değil, aynı zamanda sorumlu, etik ve sürdürülebilir biçimde yönetilen sistemler haline gelir.

3.1. Temel Uyum Gereklilikleri

Gerçek Zamanlı ve Sürekli Performans İzleme: YZ sistemi operasyonel ortama geçtikten sonra modelin doğruluk, hata oranı ve çıktı kalitesi gibi performans göstergeleri düzenli aralıklarla izlenmelidir. İzleme altyapısı, olası performans düşüşlerini ve beklenmedik sonuçları erken tespit edecek şekilde tasarlanmalıdır (ISO/IEC 42001 Md.9 – *performance assessment*; NIST AI RMF ”Monitor*” function*).

Proaktif Önyargı ve Sapma Alarm Mekanizması: Sistem performansında adalet ve doğruluk için belirlenmiş eşik değerler tanımlanmalı; bu değerler aşıldığında otomatik uyarı üretecek mekanizmalar devreye girmelidir. Örneğin, farklı gruplar arasındaki başarı oranı farkı, belirlenen yüzdeden fazla olduğunda sistem alarm vermelidir. Bu sayede potansiyel önyargılar veya hatalar birikmeden önce fark edilip giderilebilir (OECD *“Accountability” Principle*; EU AI Act – *data & performance monitoring recommendations*).

Kullanıcı Hakları ve Geri Bildirim Mekanizması: Nihai kullanıcıların veya etkilenen bireylerin, sistemin verdiği kararlara itiraz edebilmesi veya hatalı buldukları sonuçları bildirebilmesi için kolay erişilebilir bir geri bildirim kanalı oluşturulmalıdır. Bu mekanizma, yalnızca kullanıcı deneyimini değil, sistemin etik şeffaflığını da güçlendirir (EU AI Act Md.52 – *user feedback requirement*; GDPR Md.22 ve IR 71 – *Right to contest automated decisions*; OECD “*Inclusive” principle*).

Kayıt ve Log Güvenliği: YZ sistemi çalışırken oluşan tüm kritik kararlar, olaylar ve işlem kayıtları güvenli log sistemleriyle saklanmalıdır. Kayıtların bütünlüğü korunmalı, yalnızca yetkili kişiler erişebilmelidir. Bu uygulama, özellikle yüksek riskli sistemlerde denetim ve sorumluluk izlenebilirliğini destekler (ISO/IEC 42001 Md.8 – *Information security*; GDPR/KVKK – *Data security*; OECD ”*Security” Principle*).

Karar ve Sorumluluk Zinciri İzlenebilirliği: Her model kararı için giriş verileri, kullanılan model versiyonu, uygulanan kural seti ve çıktılar ilişkilendirilebilir biçimde kayıt altına alınmalıdır. Bu amaçla, “Karar Log Kaydı” formatı oluşturulabilir: Karar ID’si, zaman damgası, model versiyonu, girdi özellikleri ve çıktı bilgileri tutulmalıdır (EU AI Act Ek IV – *record-keeping requirement*; ISO/IEC 42001 – *transparency*). Bu yaklaşım hem hesap verebilirliği hem de olası hata veya zarar durumlarında sorumluluk zincirinin izlenebilirliğini sağlar (EU AI Act Md.53; NIST AI RMF – *Govern/Manage*).

Sürekli İyileştirme Prosedürü: İzleme sırasında tespit edilen hata, performans düşüşü veya etik sapmalara yönelik bir sürekli iyileştirme süreci oluşturulmalıdır. Bu süreç, modelin yeniden eğitimi, veri seti güncellenmesi veya sistem parametrelerinin revizyonu gibi aksiyonları kapsamalıdır. Sürecin sorumluları belirlenmeli ve iyileştirmeler yazılı biçimde kaydedilmelidir (ISO/IEC 42001 Md.10 – *continuous improvement*).

Dokümantasyon Güncelliği: Sisteme ilişkin tüm dokümanlar (model kartları, teknik raporlar, kullanıcı rehberleri, veri sayfaları vb.) değişikliklerden sonra güncellenmelidir. Her güncelleme için “değişiklik günlüğü (changelog)” oluşturulmalı; kim, ne zaman, hangi değişikliği yaptı net biçimde izlenebilmelidir. Bu sayede kurumsal şeffaflık ve denetim süreçlerinde bütünlük sağlanır (ISO/IEC 42001 – *documentation*; EU AI Act Ek IV – *technical documentation*; OECD *“Transparency” principle*).

Acil Durum ve Olay Müdahale Planı: YZ sisteminin beklenmedik biçimde hatalı sonuçlar üretmesi veya bir ihlal durumunda uygulanacak adımlar yazılı biçimde tanımlanmalıdır. Plan, sistemin geçici olarak durdurulması, etkilenen bireylere ve ilgili mercilere bildirim yapılması, zararın sınırlandırılması gibi önlemleri içermelidir. Bu planın düzenli olarak test edilmesi, kriz anlarında hızlı ve sorumlu müdahale kapasitesi kazandırır (EU AI Act Md.62 – *incident reporting requirement*; ISO/IEC 42001 – *corrective action*; ACM/IEEE Etik İlkeleri – *harm reduction*).

3.2. Tavsiye Edilen İyi Uygulamalar

Proaktif Anomali ve Drift İzleme: YZ sisteminin normal davranış profili belirlenmeli ve bu profilin dışında kalan olağandışı durumlar gerçek zamanlı olarak tespit edilmelidir. Anomali ve davranışsal sapmaları izlemek için otomatik alarm üreten davranışsal anomali tespit mekanizmaları kurulmalı; istatistiksel süreç kontrol grafikleriyle sistem kararlılığı düzenli takip

edilmelidir. Veri dağılımlarında veya model performansında kayma meydana geldiğinde, sistem bu durumu erken evrede tespit edip otomatik olarak müdahale etmelidir. Bu müdahaleler, drift'in şiddetine göre güvenli modda durdurma, shadow model ile paralel izleme veya önceki başarılı sürüme rollback gibi kademeli aksiyonları içerebilir. Böylece model performansı korunur ve sürekli iyileştirme döngüsü desteklenir (ISO/IEC 42001 Md.9 – *performance assessment*; NIST AI RMF "*Monitor").*

Proaktif Aşamalı Uyarı Sistemleri: Sadece eşik aşımlarında alarm üretmek yerine, trend analizine dayalı erken uyarı mekanizmaları kurulmalıdır. Model hatalarında veya önyargı metriklerinde artış eğilimi tespit edildiğinde, eşik değer aşılmadan önce ilgili ekiplere uyarı gönderilmelidir. Bu yaklaşım, olası bozulmaları büyümeden yakalamayı sağlar ve önleyici denetimi güçlendirir (OECD "*Accountability" Principle*; EU AI Act Madde 61 – *operational monitoring recommendations*).

Dashboard ile Şeffaf İzleme: Kurum içinde ve mümkünse kamuya açık biçimde erişilebilen bir YZ Etik Dashboard'u oluşturulmalıdır. Bu pano, karar sayıları, kullanıcı itirazları, ortalama cevap süreleri, model güncellemeleri ve enerji verimliliği gibi KPI'ları görüntülemelidir. Şeffaf izleme panoları hem iç paydaşlar hem de düzenleyici otoriteler için hesap verebilirliği artırır (OECD "Transparency" Principle; ISO/IEC 42001 Md.7.5 – *documentation management*).

Düzenli İç Denetimler: Kurum içi bağımsız bir denetim birimi veya etik komite, YZ sisteminin loglarını, KPI sonuçlarını ve kullanıcı geri bildirimlerini periyodik olarak gözden geçirmelidir. Bu incelemelerin sonucunda iyileştirme önerileri hazırlanmalı ve düzeltici eylem planları üst yönetime raporlanmalıdır (ICO AI Audit Framework – *accountability and governance*).

Olay Sonrası Analiz ve Öğrenme Döngüsü: Ciddi hata veya olaylardan sonra kök neden analizi (root cause analysis) yapılmalı ve olay raporu oluşturulmalıdır. Bu rapor hem teknik hem de organizasyonel öğrenmeleri içermeli; aynı tür hataların tekrarını önlemek için süreçlere entegre edilmelidir. Öğrenilen dersler dokümante edilip ilgili tüm ekiplerle paylaşılmalıdır (NIST AI RMF "Improve"; ISO/IEC 42001 – *corrective action management*).

Kriz İletişimi ve Şeffaf Bilgilendirme: YZ sisteminin neden olduğu ciddi bir olay (örneğin ayrımcı sonuç, veri ihlali, yanlış karar) durumunda yalnızca teknik müdahale değil, paydaşlara ve kamuoyuna yönelik doğru, zamanında ve şeffaf iletişim sağlanmalıdır. Kriz iletişim planı; sorumlu kişiler, bildirim süreçleri ve yetkili mercilere yapılacak raporlamaları kapsamalıdır (EU AI Act Md.62 ve 73 – *incident notification*; GDPR Md.33 ve 34 – *data breach notification*).

Kullanıcı ve Topluluk Katılımı: İzleme süreçlerine, sistemi kullanan veya etkilenen son kullanıcılar ile sivil toplum temsilcilerinin dâhil edilmesi teşvik edilmelidir. Bu amaçla kullanıcı geri bildirim kanalları, etik danışma panelleri veya memnuniyet anketleri oluşturulabilir. Katılımcı

denetim yaklaşımı, gerçek dünyadaki etkilerin erken tespitini kolaylaştırır ve güven inşa eder (OECD “Inclusive stakeholder participation”).

Veri ve Model Güncellemelerinde Onay Mekanizması: İzleme sonuçlarına göre önerilen model veya veri seti güncellemeleri, bir Değişiklik Yönetim Kurulu (Change Review Board) tarafından değerlendirilmelidir. Kurul, değişikliğin olası etkilerini (performans, adalet, açıklanabilirlik, güvenlik) analiz ederek onay verir ve geçişi kademeli biçimde uygular (ISO/IEC 42001 – *change management control*; EU AI Act Ek IV – *technical documentation*).

Gerçek Zamanlı Uyarlanabilir Risk Yönetimi: Risk değerlendirmesi yalnızca geçmiş verilere dayanmamalı; anlık bağlamsal göstergeler (ortam koşulları, kullanıcı davranışı, veri kalitesi vb.) dikkate alınarak dinamik biçimde güncellenmelidir. Risk eşikleri statik olmamalı; modelin geçmiş performans trendlerine göre otomatik olarak yeniden hesaplanmalıdır. Altyapı güvenliği, enerji kullanımı ve model performansı gibi çok boyutlu göstergeler birlikte analiz edilerek bütüncül risk skorları oluşturulmalıdır. Bu skorlar, öngörücü analizlerle gelecekteki risk senaryolarını tahmin ederek kurumlara proaktif bir güvenlik kalkanı sağlar (NIST AI RMF “*Measure & Manage”*; ISO/IEC 42001 Md.10 – *continuous improvement*).

3.3. İzleme ve İyileştirme için Kontrol Listesi

İzleme ve sürekli iyileştirme süreçlerinin etkin biçimde yürütülmesi için aşağıdaki kontrol soruları kullanılabilir:

* **Gerçek Zamanlı İzleme Kuruldu mu? -** Modelin üretim ortamındaki performansı (doğruluk, hata oranı vb.) düzenli veya gerçek zamanlı olarak izleniyor mu? İzleme sonuçları performans değerlendirme raporlarına entegre ediliyor mu? *(ISO/IEC 42001 Md.9 – performance assessment; NIST AI RMF “Monitor”)*

* **Önyargı Alarm Mekanizması Aktif mi? -** Tanımlanmış adalet ve denge metrikleri için eşik değerler belirlendi mi? Bu eşikler aşıldığında sistem otomatik alarm üretebiliyor mu? *(OECD “Accountability” Principle; EU AI Act – high-risk system recommendations)*

* **Kullanıcı Geri Bildirim Mekanizması Mevcut mu? -** Son kullanıcılar, hatalı veya adaletsiz buldukları kararları kolayca bildirebiliyor mu? İtiraz veya insan incelemesi talep edilebiliyor mu? *(EU AI Act Md.52 – feedback requirement; GDPR Md.22; OECD AI Principles)*

* **Loglar Güvenli mi ve Erişim Yetkili mi?** - Sistem karar kayıtları güvenli biçimde saklanıyor mu? Logların bütünlüğü, gizliliği ve değiştirilemezliği sağlanıyor mu? *(ISO/IEC 42001 Md.8 – information security; GDPR/KVKK uyumu; OECD “Security” Principle)*

* **Karar Girdileri ve Çıktıları İzlenebilir mi?** - Her karar için giriş verileri, model versiyonu ve çıktılar güvenli bir log sisteminde tutuluyor mu? "Karar Log Kaydı" formatı tanımlı ve tüm modellerde uygulanıyor mu? *(EU AI Act Ek IV – record keeping; ISO/IEC 42001 – trasparency; OECD "Accountability" principle)*

* **Sürekli İyileştirme Süreci Tanımlı mı? -** İzleme bulguları, model veya veri seti güncellemelerine dönüştürülüyor mu? Düzeltici ve önleyici aksiyonlar için yazılı bir prosedür mevcut mu? *(ISO/IEC 42001 Md.10 – Continuous Improvement)*

* **Dokümantasyon Güncelleniyor mu?** - Her değişiklik sonrası model kartı, veri sayfası ve teknik raporlar güncelleniyor mu? Model değişiklikleri için "değişiklik günlüğü (changelog)" tutuluyor mu? *(ISO/IEC 42001 – documentation control; EU AI Act Ek IV – technical document requirements)*

* **Acil Durum Planı ve Müdahale Süreci Tanımlı mı? -** Sistemin beklenmedik bir hata veya ihlal durumunda nasıl durdurulacağına ilişkin planlama mevcut mu? Bu plan, olay bildirimi, zarar sınırlandırma ve yetkili mercilere bilgilendirme adımlarını içeriyor mu? *(NIST AI RMF "Manage"; ACM/IEEE Etik İlkeleri – harm reduction)*

3.4. Önerilen Kilit Performans Göstergeleri (KPI)

Aşağıdaki KPI örnekleri, yapay zekâ yönetişiminde kurumların performansı ölçmesine yardımcı olabilir. Her kurum, kendi ihtiyaç ve risk profiline göre ek metrikler tanımlamalı ve bunları düzenli olarak izlemelidir.

Ciddi Olay Bildirim KPI'ları: YZ sisteminin yol açtığı geri çağırma olaylarının veya ciddi ihlallerin sayısı ile bunların yetkili mercilere bildirim süreleri izlenmelidir. Amaç, EU AI Act Md.73 uyarınca "serious incident" bildirimlerinin zamanında ve düzenli biçimde yapılmasını sağlamaktır. Örneğin:

- Son çeyrekte bildirilen ciddi olay sayısı,
- Ortalama bildirim süresi (saat/gün),
- Kritik olayların yüzde kaçının yasal süre (örneğin 72 saat – GDPR Md.33; EU AI Act Md.62) içinde raporlandığı.

İçerik Şeffaflığı KPI'ları: Özellikle generative AI kullanılıyorsa, üretilen içeriğin yüzde kaçının kaynağının doğrulanabilir olduğu izlenmelidir. Amaç, üretken yapay zekâ içeriklerinde şeffaflık ve güvenilirlik düzeyini değerlendirmektir *(G7 Hiroshima AI Principles; C2PA Standard)*. Örneğin:

- “Tüm AI içeriklerinin %95’i uygun şekilde işaretlendi”.
- “Deepfake veya sahte içerik raporlarının oranı %X”.

Açıklanabilirlik KPI’ları: Sistem tarafından verilen kritik kararların ne kadarına anlaşılır açıklama eşlik ettiği ve bu açıklamaların kullanıcılar tarafından yararlı bulunma oranı takip edilmelidir. Amaç, açıklanabilirlik mekanizmalarının gerçekten anlaşılabilir ve faydalı olup olmadığını ölçmektir (NIST AI RMF). Örneğin:

- “Kritik kararların %85’inde açıklama sunuldu”.
- “Kullanıcı geri bildirimlerine göre açıklamaların %70’i tatmin edici bulundu”.
- “Farklı kullanıcı profillerinin %90’ı açıklamaları yeterli buldu”.

Adalet KPI’ları: Modelin farklı demografik gruplar arasında sonuç tutarlılığını ve tarafsızlığını ölçen göstergeler düzenli olarak izlenmelidir. Amaç, hiçbir grubun sistem çıktılarından sistematik olarak olumsuz etkilenmediğini doğrulamaktır *(EU AI Act Md.10–15; ISO/IEC 24027 – bias measurement standards).* Örneğin:

- *Disparate Impact* (%80 kuralı – korunan gruplar arasında performans farkı belirlenen sınırın altında mı?)
- *Equalized Odds* (gruplar arasındaki “False Positive”, “True Positive” oran farkları)
- *Treatment Equality* (“False Positive”, “False Negative” değerlerinin birbirine oranının tüm gruplar için aynı olması)

Şikâyet Çözüm Süresi KPI’ları: Etik, adalet veya şeffaflıkla ilgili kullanıcı şikâyetlerinin ortalama çözülme süresi ölçülmelidir. Amaç, kullanıcı geri bildirimlerinin etkin biçimde yönetildiğini ve kurum içi reaksiyon kabiliyetinin güçlü olduğunu göstermektir *(EU AI Act Md.62; OECD “Stakeholder participation”).* Örneğin:

- “Kullanıcı şikâyetlerinin %90’ı 7 iş günü içinde çözüme kavuşturuldu”.

Enerji ve Sürdürülebilirlik KPI’ları: Modelin eğitim ve çalıştırma süreçlerinde enerji tüketimi ve karbon ayak izi izlenmelidir. Amaç, yapay zekâ operasyonlarında çevresel sürdürülebilirlik hedeflerine katkı sağlamaktır *(OECD Sustainable Development; ISO/IEC 42001 Md.10).* Örneğin:

- “Model eğitimlerinde enerji tüketimi bir önceki döneme göre %X azaldı”.

Kriz İletişimi KPI’ları: Ciddi olaylardan sonra kamuya yapılan açıklamaların zamanlaması ölçülmelidir. Amaç, kriz durumlarında şeffaflık ve hızlı bilgilendirme kültürünün korunmasıdır. *(EU AI Act Md.73; UNESCO Ethics Md.27).* Örneğin:

- “Son 12 ayda yaşanan kritik olayların %95’inde kamu bilgilendirmesi 48 saat içinde yapıldı”.

Eğitim ve Farkındalık KPI’ları: Kurum içi çalışanların YZ etiği ve yönetişimi eğitimlerine katılım oranı izlenmelidir. Amaç, kurum genelinde etik farkındalık ve yönetişim kapasitesinin artırılmasıdır *(ISO/IEC 42001 Md.7; OECD Education).* Örneğin:

- “İlgili ekiplerin %100’ü yılda en az bir kez YZ etik eğitimi aldı”.

Model Güncelleme KPI’ları: Modelin yeniden eğitimi, drift kontrolü ve performans güncellemeleri ölçülmelidir. Amaç, sistemin güncel veri ve koşullara uygun biçimde çalıştığını güvence altına almaktır *(EU AI Act Md.61–62).* Örneğin:

- “Yüksek riskli modellerin %90’ı yılda en az bir kez yeniden eğitildi”.

Unlearning / Silme Talepleri KPI’ları: Kullanıcıların veri silme taleplerine verilen yanıt süresi ve başarı oranı takip edilmelidir. Amaç, kullanıcı haklarının (özellikle unutulma hakkının) etkin biçimde yerine getirildiğini göstermek *(GDPR Md.17; ISO/IEC 42001 – corrective action principle).* Örneğin:

- “Silme taleplerinin %95’i 30 gün içinde yerine getirildi”.

4. Yönetişim ve Etik Kurumsal Yapılanma

Yapay zekâ sistemlerinde etik, hesap verebilir ve sürdürülebilir bir uygulama kültürü oluşturmanın temeli, sağlam bir kurumsal yönetişim yapısı kurmaktan geçer. Bu bölüm, yapay zekâ uygulamalarında etik ilkelerin yalnızca doküman düzeyinde değil, kurumun işleyişine, karar mekanizmalarına ve stratejik hedeflerine entegre edilmesini hedefler.

Etkin bir YZ yönetişimi; açık ve izlenebilir bir sorumluluk zinciri, disiplinler arası etik komiteler, şeffaf karar alma süreçleri, bağımsız denetim ve sürekli farkındalık eğitimi gibi unsurların bütüncül biçimde işlemesiyle mümkün olur. Kurumlar, yapay zekâ sistemlerinin yaşam döngüsünün tüm aşamalarında (Tasarım, geliştirme, test, devreye alma, izleme ve sonlandırma süreçlerinde) etik gözetim ve iç kontrol mekanizmalarını kalıcı hale getirmelidir. Bu yaklaşım

sayesinde kurum, yalnızca düzenleyici uyum sağlayan değil, aynı zamanda kendi iç denetim ve etik kontrol kapasitesine sahip bir yapıya dönüşür.

Bu model, ulusal mevzuat (KVKK, Dijital Haklar Stratejisi vb.) ve uluslararası standartlarla (EU AI Act, ISO/IEC 42001, OECD AI Principles) uyumlu bir kurumsal yönetişim standardı kurulmasına yardımcı olur. Ayrıca yapay zekâ yönetişimi, veri koruma, insan gözetimi, açıklanabilirlik, güvenlik ve risk yönetimi ilkelerini tek çatı altında bütünleştiren bir yaklaşımı benimser.

Bu bölüm, aynı zamanda kurum içi rollerin net biçimde tanımlanmasını öngörür: AI Ethics Officer, Risk Owner, Compliance Lead, Data Protection Responsible gibi rollerin görev tanımları, yetki sınırları ve karar alma sorumlulukları açık biçimde belirlenmelidir. Üst yönetimin sahipliği ve desteği, yönetişimin yalnızca “teknik bir süreç” değil, kurumsal kültürün ayrılmaz bir parçası haline gelmesini sağlar.

AIPA’nın YZ etiği ve yönetişimi çerçevesi doğrultusunda tasarlanan bu yapı, yalnızca kurum içi süreçleri değil, aynı zamanda dış paydaşlarla olan iş birliklerini de etik, şeffaf ve hesap verebilir bir anlayışla yönlendirir.

Nihai hedef; kendi kendini denetleyebilen, etik farkındalığı yüksek, güven ve sorumluluk temelli, özel sektörün dinamizmiyle, ulusal egemenliğe ve toplumsal değerlere duyarlı bir yapay zekâ yönetişim ekosistemi inşa etmektir.

4.1. Temel Uyum Gereklilikleri

YZ Etik Kurulu Oluşturulması: Kurum bünyesinde disiplinler arası uzmanlardan oluşan bir YZ Etik Kurulu veya Komitesi kurulmalıdır. Bu kurul, yapay zekâ projelerini düzenli olarak gözden geçirir, etik riskleri değerlendirir ve yönetime tavsiyelerde bulunur. Kurulda, OECD’nin çok paydaşlı yaklaşımına uygun şekilde hukuk, teknik, sosyoloji, psikoloji gibi farklı uzmanlık alanlarından temsilciler bulunması önerilir. Birçok uluslararası kuruluşta standart hale gelen bu yapı, kurumun etik gözetim kapasitesini güçlendirir ve karar süreçlerine sistematik etik değerlendirme kazandırır.

Sorumluluk Atamaları: Her bir yapay zekâ sistemi için kurum içinde sorumlu bir kişi atanmalıdır. Bu rol, örneğin “etik ve uyum sorumlusu” veya benzeri bir unvanla tanımlanabilir. Görevi, ilgili sistemin tüm yaşam döngüsünde etik ve uyum süreçlerinin gözetimini sağlamaktır. ISO/IEC 42001 Liderlik maddesi, üst yönetimin sorumlulukları net şekilde atamasını öngörür; EU AI Act ise sağlayıcı ve kullanıcı rollerinde ayrı yükümlülükler tanımlar. Kurum içinde bu roller net ayrılmalı, görev ve yetkiler yazılı biçimde belirlenmelidir.

Eğitim ve Farkındalık Programı: Kurum çalışanlarının YZ etiği, veri gizliliği, güvenlik ve ilgili mevzuat konularında düzenli eğitimlerden geçmesi tavsiye edilir. Özellikle geliştiriciler, veri bilimcileri, ürün yöneticileri ve karar vericiler yılda en az bir kez bu konularda eğitim almalıdır. Eğitim katılım kayıtları tutulmalı, içerikler düzenli aralıklarla güncellenmelidir *(ISO/IEC 42001 Md.7 – Awareness and Education; OECD "Education" recommendations).*

Politika ve Rehber Dokümanları: Kurumun YZ Etiği ve Yönetişim Rehberi, üst yönetim tarafından onaylanmalı ve tüm çalışanların erişimine açık hale getirilmelidir. Geliştirici ekipler için ayrıca veri etiketi yönergeleri, model geliştirme standartları ve kodlama etik ilkelerini içeren iç rehberler hazırlanmalıdır. Bu belgeler, kurumun etik standartlarını operasyonel süreçlere entegre etmeyi amaçlar ve iyi uygulama çerçevesi oluşturur.

İç Raporlama ve Revizyon Mekanizması: Üst yönetim, yapay zekâ sistemlerinin performansı ve uyumu hakkında düzenli raporlama yapılmasını sağlamalıdır. En az yılda bir kez, belirlenen KPI'ların hedeflere ulaşıp ulaşmadığı, risk envanteri güncellemeleri, yaşanan olaylar, alınan dersler ve denetim sonuçları gözden geçirilmelidir. Bu incelemelerin sonucunda politika, hedef veya prosedürlerde iyileştirme gerekiyorsa revizyon süreci başlatılmalıdır *(ISO/IEC 42001 – Management Review).*

Olay Yönetimi ve Bildirim Prosedürü: Bir YZ sisteminin neden olduğu istenmeyen etki, hata veya kişisel veri ihlali durumunda uygulanacak prosedürler önceden tanımlanmalıdır. Bildirim zinciri, iletişim kanalları ve müdahale adımları açıkça belirlenmeli; ciddi olaylarda bildirim süreleri (örneğin 15 gün veya düzenleyici gerekliliklerde öngörülen 72 saat) takip edilmelidir. Bu prosedür, Acil Durum Planı (Bölüm 3.1) ile entegre biçimde çalışmalıdır *(EU AI Act Md.62 – Incident Notification; ISO/IEC 42001 – Corrective Action Process).*

Uyum ve Sertifikasyon Hedefleri: Kurum, kendi YZ sistemleri için uluslararası uyum standartlarına ve sertifikasyonlara ulaşmayı hedeflemelidir. Örneğin ISO/IEC 42001:2023 sertifikasyonu elde etmek veya AB YZ Tüzüğü kapsamında uygunluk beyanı hazırlamak gibi hedefler belirlenebilir. Bu hedefler, üst yönetim tarafından desteklenmeli ve gerekli kaynaklar tahsis edilmelidir *(ISO/IEC 42001 – Certification; EU AI Act – Conformity Assessment Process).*

Risk Envanteri ve Etki Değerlendirmesi: Tüm yapay zekâ projeleri için bir Risk Envanteri tutulmalı ve her proje başlamadan önce potansiyel etik, hukuki, güvenlik ve toplumsal etkiler değerlendirilmelidir. Yüksek riskli sistemler için ayrıca detaylı bir AI Etki Değerlendirme (AIA) raporu hazırlanmalı ve onaylanmalıdır. Bu değerlendirmeler, KVKK kapsamındaki **Veri Koruma Etki Değerlendirmesi (DPIA)** süreçleriyle entegre şekilde yürütülebilir *(EU AI Act Md.29 – High-risk systems; OECD "Robustness, Security, Safety" principles).*

Hukuki Uyumluluk Kontrolü: Her YZ projesi, yürürlükteki sektörel düzenlemeler ve ulusal mevzuat açısından hukuk birimi veya danışmanlar tarafından incelenmelidir. Finans, sağlık,

enerji gibi alanlarda yürütülen projelerde ilgili düzenleyici kurumların (ör. BDDK, TİTCK) yükümlülüklerine uyum sağlanmalıdır. Bu adım hem ulusal düzenlemelere saygıyı hem de proje güvenilirliğini güçlendirir. Uyum sağlanmadan hiçbir sistem canlı ortama alınmamalıdır.

Yaşam Döngüsü (Lifecycle) Yönetimi: Kurum, yapay zekâ sistemlerinin tüm yaşam döngüsünü (tasarım, geliştirme, dağıtım, izleme, bakım ve sonlandırma) yönetecek süreçleri tanımlamalıdır. Her aşamada kimlerin sorumlu olduğu, hangi onayların gerektiği ve geçiş şartları açıkça belirlenmelidir. Bu sayede sistemin ömrü boyunca kontrol ve izlenebilirlik sağlanır, sorumluluk zinciri kopmaz *(ISO/IEC 42001 Md.8 – Operation Planning; NIST AI RMF – Map/Measure/Manage).*

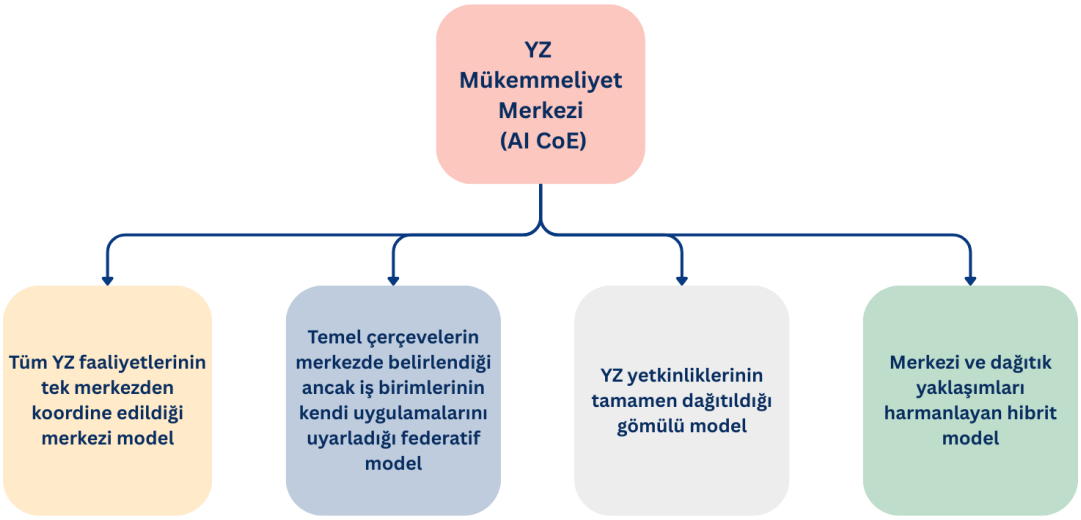
Risk Sınıfı Eşlemesi: Her YZ sistemi için, EU AI Act'teki yaklaşıma benzer biçimde kurum içi bir risk sınıflandırması yapılmalıdır. Sistemin yasaklı, yüksek riskli, sınırlı riskli veya asgari riskli kategorilerden hangisine girdiği netleştirilmelidir. Eğer sistem yüksek risk grubundaysa, bu kategoriye özel ek gereklilikler (ör. kayıt tutma, raporlama, insan gözetimi) uygulanmalıdır. Risk sınıflandırması, proje başlangıcında etik kurul veya risk komitesi tarafından onaylanmalıdır *(EU AI Act Md.9-15 – Risk Categories and Requirements).*

4.2. Tavsiye Edilen İyi Uygulamalar

Yönetişimde Çeşitlilik: YZ sistemlerinin etik ve yönetişim mekanizmalarında, etik kurullarda cinsiyet, yaş, engellilik ve kültürel çeşitliliklerin gözetilmesi; kararların daha kapsayıcı ve meşru olmasına, kör noktaların azaltılmasına katkı sağlar. Örneğin kurum içi AI etik kurulu yalnızca hukukçu ve mühendislerden değil, aynı zamanda sosyolog, psikolog ve iletişim uzmanlarından da oluşmalıdır *(EU AI Act Recital 44, G7 Hiroshima AI Process).*

Çok Paydaşlı Etik İnceleme: Kurum içindeki etik kurulun yanı sıra, önemli YZ projeleri için sivil toplum, akademi ve müşteri temsilcilerinden oluşan bir dış paydaş paneli oluşturmak faydalı olabilir. Bu panel, projeyi dışarıdan gözle değerlendirerek toplum nezdindeki algı ve riskler konusunda uyarılarda bulunabilir *(OECD "multi-stakeholder approach").* Bu bağımsız geri bildirimler, AIPA'nın ulusal düzeydeki etik değerlendirmelerine de girdi sağlayabilir.

YZ Mükemmeliyet Merkezi (AI CoE) Kurulumu: Kurum içinde yapay zekâ projelerinin koordinasyonu, standardizasyonu ve yaygınlaştırılmasından sorumlu bir YZ Mükemmeliyet Merkezi (AI Center of Excellence) kurulması önerilir. Bu birim, teknik projelerin yanı sıra stratejik karar alma süreçlerinde de yapay zekânın kurumsal düzeyde benimsenmesini destekler; veri yönetişimi, etik ve regülasyon gerekliliklerini takip eder. YZ Mükemmeliyet Merkezi, kurumun olgunluk düzeyine göre farklı modellerde yapılandırılabilir:



Resim: YZ Mükemmeliyet Merkezi (AI CoE) Seçenekleri

Kurum, birkaç başarılı YZ projesini hayata geçirdikten ve kurumsal ölçekte yaygınlaşma ihtiyacı belirginleştikten sonra "YZ Mükemmeliyet Merkezi" kurmayı değerlendirmelidir. Uygun zamanda ve doğru modelle kurulan bu yapı, yapay zekânın kurumsal dönüşümde kaldıraç etkisi yaratmasını sağlar.

Açık Politika ve Rehberler: Kurumun YZ politikası, kamuoyuyla paylaşılabilir biçimde şeffaf hazırlanmalıdır. Örneğin şirketin YZ Etik İlkeleri belgesi kurumsal web sitesinde yayımlanabilir. Ayrıca geliştiriciler için hazırlanan iç rehberler (ör. kodlama etik rehberi, veri kullanım kılavuzu vb.) herkese açık olmasa da kurum içinde yaygınlaştırılmalıdır. Bu şeffaflık hem çalışanların hem de kamuoyunun kuruma olan güvenini artırır.

Düzenli Yönetim Gözden Geçirmesi: Üst yönetim, YZ politika uyumunu yılda en az bir kez gözden geçirmelidir. Bu kapsamda belirlenen KPI'ların hedeflere ulaşıp ulaşmadığı, risk envanterindeki değişiklikler, gerçekleşen olaylar, alınan dersler ve sertifikasyon durumları değerlendirilir. Gerekirse politika ve hedeflerde revizyona gidilir *(ISO/IEC 42001 – Management Overview).*

Uluslararası Sertifikasyon ve İş Birliği: Kurum, uygun alanlarda uluslararası sertifika ve yarışmalara katılmalıdır. Örneğin YZ sistemleri için IEEE ECPAIS (Ethically Certified Artificial Intelligence Systems) etik sertifikası, ISO/IEC 42001:2023 standardına dayalı sertifikalar veya benzeri etik mühürler almak; GPAI veya benzeri forumlarda projelerini sunmak kurumun hem kendini test etmesini sağlar hem de küresel entegrasyonuna katkı sunar. Bu yaklaşım, Türkiye'nin yapay zekâ alanındaki bağımsız ve saygın konumunu güçlendirir.

Etik İnovasyon Teşviki: Kurumlar yalnızca uyum ve risk azaltmaya değil; toplumsal faydayı öncelemeye, insan merkezli ve etik ilkelere uygun yenilikçiliği teşvik etmeye de odaklanmalıdır.

Örneğin eğitim, sağlık, çevre veya engelli bireyler için erişilebilirlik gibi alanlarda sosyal fayda yaratan yapay zekâ projelerine öncelik verilmesi, kurumun etik olgunluğunu artırır *(EU AI Act Recital 1 & 5; UNESCO Recommendations on AI Ethics- Md.19 & 25).* Ayrıca kurum içinde etik kriterlere göre Ar-Ge bütçesi ayrılması önemlidir. Etik hackathonlar, yani üniversiteler, startuplar ve özel sektör arasında etik temelli inovasyon yarışmaları veya sandbox ortamları, bu girişimi destekleyebilir.

Açık Kaynak ve Üçüncü Taraf Yazılımlar: Kurum içinde kullanılan açık kaynak veya üçüncü taraf YZ modellerine ilişkin özel politikalar oluşturulmalıdır. Bu modellerde oluşabilecek hatalar için sorumluluk paylaşımı, güvenlik güncellemelerinin takibi ve gerektiğinde geçici sınırlama uygulanması gibi konular önceden planlanmalıdır. Bu, ileride ortaya çıkabilecek risklerde hızlı aksiyon almayı kolaylaştırır *(EU AI Act Md.53(2) – Third-party responsibilities).*

YZ Dönüşüm Programı: Kurumlar, yapay zekâyı birkaç izole proje yerine tüm organizasyona değer katan bir dönüşüm programı olarak ele almalıdır. Bu kapsamda sırasıyla;

Resim: YZ Dönüşüm Programı

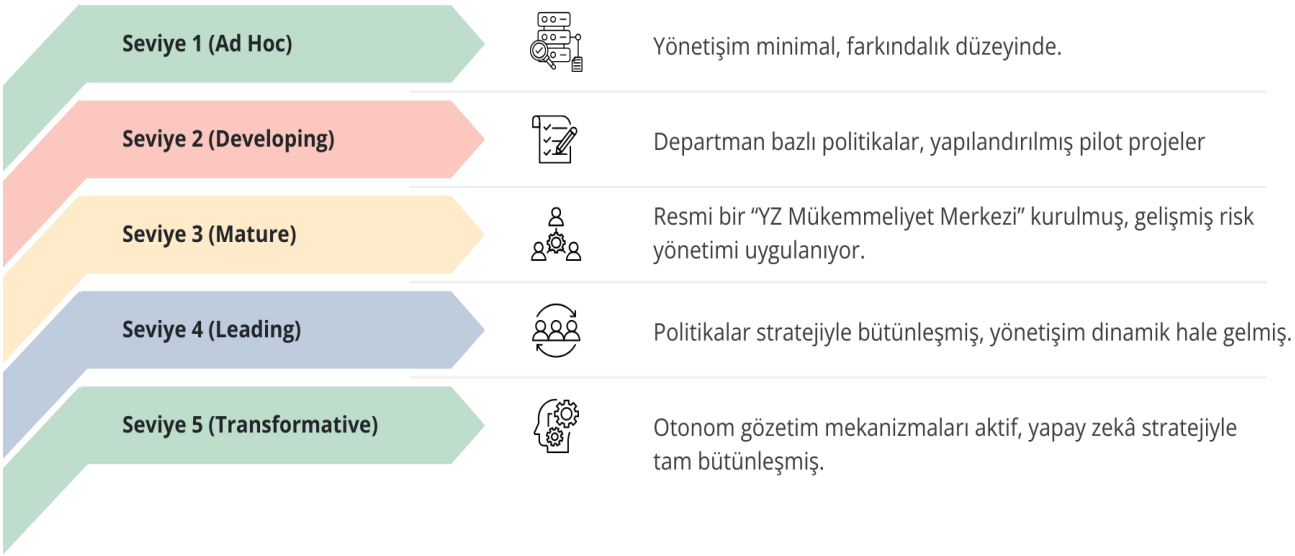
Bu yolculuk boyunca organizasyonel süreçler de dönüşüme uyarlanmalıdır; roller ve sorumluluklar yeniden tanımlanmalı, karar alma mekanizmaları güncellenmeli, bütçe ve portföy yönetimi stratejik hedeflerle uyumlu hale getirilmelidir. Böylelikle yapay zekâ, kurum kültürü ve stratejisinin ayrılmaz bir parçası olur.

Kurumsal YZ Dönüşüm Kasları: Kurumsal AI dönüşüm kasları, yapay zekânın proje seviyesinden çıkarak kurumun stratejik ritmine yerleşmesini sağlayan temel yapı taşlarını ifade eder. Bu yaklaşım, öncelikle yapay zekâ girişimlerinin iş hedefleriyle tutarlı ve güçlü biçimde hizalanmasını (stratejik uyum) gerektirir; bu sayede AI çözümleri, kurumun değer üretme amacını destekleyen bir çerçevede konumlanır. Ardından, teknolojinin yalnızca teknik bir

bileşen olarak kalmaması için süreçlere, iş akışlarına ve karar mekanizmalarına entegre edilmesi (operasyonel entegrasyon) kritik önem taşır. Dönüşümün sürdürülebilir olabilmesi için sağlam bir veri temeli oluşturmak, veri kalitesini korumak ve veri yönetişimini kurumsal bir standarda dönüştürmek (veri olgunluğu) esastır. Bunun yanında, yapay zekânın kurum içinde yarattığı iş değerini düzenli olarak ölçmek, izlemek ve optimize etmek (etki yönetimi) dönüşümün gerçek katkısını görünür kılar. Son olarak, çalışanların yetkinliklerini artıran, etik farkındalığı güçlendiren ve yeni çalışma biçimlerine uyum sağlayan bir kültürün desteklenmesi (kültürel ve yetkinlik gelişimi), teknolojinin güvenilir ve sorumlu kullanımını mümkün kılar. Tüm bu kaslar eş zamanlı olarak bir araya geldiğinde, kurumlar yapay zekâ uygulamalarını kontrollü, şeffaf, değer üreten ve sürdürülebilir bir şekilde ölçeklendirebilir.

Olgunluk Seviyeleri ve Kademeli Yönetişim: Kurumun yapay zekâ olgunluk seviyesi arttıkça, uygulaması gereken yönetişim mekanizmaları da giderek daha sofistike hale gelir. Başlangıç seviyesinde temel etik prensipler, kayıt mekanizmaları ve basit uyum kontrolleri yeterli olabilirken; ürünleştirme ve ölçekleme aşamalarında model kartları, kullanıcı geri bildirim sistemleri ve denetim izleri devreye girer. İleri olgunluk seviyelerinde ise proaktif risk izleme, şeffaf raporlama ve harici denetim mekanizmaları işletilir.

Bu kademeli yaklaşım sayesinde her kurum, bulunduğu olgunluk düzeyine uygun yönetişim araçlarını hayata geçirerek YZ kullanımını güvenilir ve sürdürülebilir kılabilir. Örneğin:



Resim: YZ Olgunluk Seviyeleri

Kurumlar kendi olgunluk seviyelerini değerlendirip bir üst seviyeye geçiş için gereken yönetişim yatırımlarını planlamalıdır *(OECD AI Policy Observatory – AI Governance Maturity Levels; World Economic Forum – AI Governance Maturity Model; ISO/IEC 42001).*

Çeviklik: Yapay zekâ sistemlerinin hızla değişen veri, kullanıcı davranışları ve iş ihtiyaçlarına uyum sağlayabilmesi için kurum içinde dinamik, öğrenen ve sürekli gelişen bir yapı

oluşturulmasını gerektirir. Bu yaklaşım, AI modellerinin yalnızca teknik olarak değil, operasyonel olarak da döngüsel biçimde yönetilmesini; risklerin erken fark edilmesini, iş birimleri, hukuk, risk, güvenlik ve veri ekipleri arasında güçlü bir işbirliği kültürü kurulmasını sağlar. Çeviklik, kontrol mekanizmalarını zayıflatmaz; aksine, gerçek zamanlı izleme, düzenli geri bildirim döngüleri, sürekli iyileştirme ve değer odaklı önceliklendirme gibi kaslarla yönetişimi daha etkin ve sürdürülebilir hale getirir. Böylece yapay zekâ uygulamaları, statik bir uyum yaklaşımından çıkarak kurumun ritmine entegre edilen, adaptif ve şeffaf bir yönetişim modeli içinde yönetilir.

İş Birlikleri ve Kolektif Öğrenme: Benzer alanlarda YZ geliştiren kurumlar (özellikle aynı sektördeki) bir araya gelerek sektörel etik konsorsiyumlar oluşturabilir. Bu konsorsiyumlar, ortak riskler ve çözümler konusunda bilgi paylaşımı yaparak sektör çapında standartlar geliştirebilir. Örneğin, bankacılık sektöründe yapay zekâ kullanımı için ortak etik kurallar belirlemek amacıyla BDDK ile koordinasyon içinde çalışılabilir. Bu yaklaşım, ulusal egemenlik çerçevesinde sektörün kendi kendini düzenlemesini destekler.

4.3. Etik ve Yönetişim için Kontrol Listesi

Kurumun YZ yönetişim yapısını değerlendirmek için aşağıdaki kontrol noktaları kullanılabilir:

* **AI Etik Kurulu / Komitesi Aktif mi?** – Kurum içinde farklı disiplinleri barındıran bir YZ etik kurulu oluşturuldu mu ve düzenli olarak toplanıyor mu? *(OECD - Multi-Stakeholder Approach)*

* **Sorumluluk Ataması Yapıldı mı?** – Her bir YZ sistemi için “sorumlu etik ve uyum denetçisi” atandı mı ve bu kişiler görev tanımlarıyla birlikte biliniyor mu? *(ISO/IEC 42001 – Leadership, Assigning Responsibility)*

* **Eğitim / Farkındalık Programı Var mı?** – Geliştirici ve yönetici ekiplerine düzenli olarak YZ etiği, güvenilir YZ ve yasal uyum eğitimleri veriliyor mu? Katılım oranları ve eğitim içerikleri yeterli düzeyde mi? *(ISO/IEC 42001 – Awareness and education requirement; OECD education recommendation)*

* **Olay Yönetim Prosedürü Tanımlı mı?** – Bir hata, ihlal veya istenmeyen etki durumunda izlenecek adımlar açık biçimde tanımlandı mı ve ilgili ekipler bu prosedürü biliyor mu? *(EU AI Act Md.62 – Serious incident notification; ISO/IEC 42001 – Corrective action)*

* **Uyum ve Sertifikasyon Hedefleri Belirlendi mi?** – Kurum, YZ sistemleri için ISO 42001 vb. sertifika hedefleri belirleyip bunlara yönelik bir takvim oluşturdu mu? *(ISO/IEC 42001 – Certification; EU AI Act – Declaration of Conformity)*

* **Risk Envanteri Tutuluyor mu?** – Tüm YZ projeleri için güncel bir risk envanteri oluşturuldu mu ve risk analizleri düzenli olarak dökümante ediliyor mu? *(EU AI Act Md.29 – Compliance assessment for high-risk systems; OECD "Robustness, Security, Safety" Principle)*

* **Hukuki Uyumluluk Kontrolü Yapılıyor mu?** – YZ sistemleri, ilgili sektör mevzuatı ile genel veri koruma ve tüketici koruma yasalarına uygunluk açısından hukuk birimi tarafından düzenli olarak kontrol ediliyor mu? *(Ulusal düzenlemeler – KVKK, sektörel AI rehberleri vb.)*

* **Ömür Döngüsü Yönetimi Mevcut mu?** – YZ sisteminin yaşam döngüsündeki tüm aşamalar (tasarım, geliştirme, devreye alma, bakım, kullanım, sonlandırma) için sorumlular ve süreçler tanımlandı mı? *(ISO/IEC 42001 – Scope and lifecycle management; NIST AI RMF Stages)*

* **Risk Sınıfı Belirlendi mi?** – Sistem için kurum içi politikada bir risk sınıfı (yasaklı / yüksek / orta / düşük) belirlendi mi ve buna uygun önlemler alındı mı? *(EU AI Act Md.9-15 – Risk categories and requirements)*

* **YZ Yönetim Sistemi Kuruldu mu?** – Kurum genelinde YZ projelerini yönetecek bir YZ yönetim sistemi (politika, hedefler, denetim planları vb.) oluşturuldu mu? Kapsam ve hedefler açık biçimde tanımlandı mı? *(ISO/IEC 42001:2023 – AIMS Standard Requirements)*

* **İtiraz ve Açıklama Hakkı Uygulanıyor mu?** – Sistem tamamen otomatik kararlar veriyor ve bireyleri önemli ölçüde etkiliyorsa, bireylere karara itiraz etme ve kararın açıklamasını talep etme hakkı tanındı mı? *(GDPR Md.22; GDPR İlam 71; EDPB WP251 Guide)*

* **Açık Kaynak / Üçüncü Taraf Model Politikası Var mı?** – Kurum, harici açık kaynaklı YZ modelleri veya üçüncü taraf API'lar kullandığında, sorumluluk dağılımını ve yükümlülükleri tanımlayan bir politika belirledi mi? *(EU AI Act Md.53(2)-Third-party exceptions)*

* **ROPA (Record of Processing Activities) ve AI Envanteri Entegrasyonu Yapıldı mı?** – Kurumun KVKK/GDPR kapsamındaki Veri İşleme Faaliyetleri Envanteri (ROPA) ile YZ sistemleri envanteri entegre mi? YZ sistemlerinin kullandığı veri kategorileri ROPA'da listeleniyor ve birbirine referanslı mı? *(GDPR Md.30 – record keeping; EU AI Act Ek IV – technical documentation requirement)*

* **Sorumluluk Zinciri Haritalandı mı?** – YZ sisteminin tasarımcısından son kullanıcısına kadar (Tasarımcı – Sağlayıcı – Dağıtıcı – Kullanıcı) olan sorumluluk zinciri kurum tarafından analiz edilip tanımlandı mı? Olası bir zarar durumunda hesap verebilirlik zinciri belgelendi mi? *(EU AI Act Md.53 - Supply chain responsibilities)*

* **Veri Koruma Görevlisi ve AI Sorumlusu İş Birliği Var mı?** – Kurumun KVKK/GDPR kapsamındaki Veri Koruma Görevlisi (Data Protection Officer) ile YZ uyum sorumlusu/ekibi tanımlandı mı ve aralarında düzenli iş birliği mekanizması kuruldu mu? *(EU AI Act Md.4, Md.72 – Cooperation requirement; GDPR Md.37–39 – Data Protection Officer job description)*

Sonuç ve Vizyon

AIPA *Yapay Zekâ Etiği ve Yönetişimi İyi Uygulamalar Rehberi*, Türkiye'nin yapay zekâ alanında bağımsız, rekabetçi ve güven veren bir yol haritası ile ilerlemesini hedeflemektedir. Bu çerçeve sayesinde şirketler ve özel kurumlar, YZ sistemlerinin tasarımından canlı ortama alınmasına kadar her aşamada uymaları gereken kuralları ve iyi uygulamaları açık biçimde bileceklerdir.

Temel uyum gereklilikleri, Türkiye'nin hem kendi yasaları hem de uluslararası taahhütleriyle uyumunu sağlarken; tavsiye edilen rehber ilkeler ve kontrol listeleri, işletmelerin iç denetim ve yönetişim süreçlerinde pratik, ölçeklenebilir ve sürdürülebilir modeller geliştirmelerine olanak tanır. Bu yaklaşım, özel sektörde etik uyumun yalnızca bir yükümlülük değil, aynı zamanda marka itibarı, yatırım güveni ve inovasyon kalitesi açısından rekabet avantajı sağlayan bir unsur olduğunu vurgular. Özellikle agentic sistemler, üretken yapay zekâ ve çoklu ajan mimarileri gibi hızla gelişen alanların getirdiği riskler de bu çerçeveye dahil edilmiştir. Sürekli veri yönetişimi, multimodal şeffaflık standartları, proaktif risk izleme ve kademeli olgunluk seviyeleri gibi yapı taşları sayesinde, yalnızca bugünün değil, yakın geleceğin ticari gereksinimlerine de yanıt veren dinamik bir yönetişim yaklaşımı oluşturulmuştur.

Uluslararası referans standartları (EU AI Act, ISO/IEC 42001, OECD AI Principles, NIST AI RMF) Türkiye'nin kurumsal ekosistemine uyarlanmış; böylece evrensel etik ve güvenlik normları, yerel iş kültürü, regülasyon ortamı ve rekabet ihtiyaçlarıyla harmanlanmıştır. Bu çerçeve, Türkiye'nin özel sektör tabanlı yapay zekâ uygulamalarında bölgesel lider ve küresel ölçekte saygın bir konuma yükselmesini hedefler.

Sonuç olarak, iş dünyası ve teknoloji ekosisteminin tüm bileşenleri için ortak bir dil, beklenti ve hesap verebilirlik çerçevesi oluşturan AIPA Yapay Zekâ Etiği ve Yönetişimi İyi Uygulamalar Rehberi, özel sektörün etik, şeffaf ve güvenilir yapay zekâ sistemleri geliştirmesine yön verir. Bu çerçeve yalnızca yasal uyum aracı değil, aynı zamanda Türkiye'nin uluslararası alanda "etik ve güvenilir yapay zekâ" markasını güçlendiren bir rekabet stratejisidir. AIPA, bu vizyonu küresel iş birlikleri ve özel sektör ortaklıklarıyla güçlendirerek, Türkiye'yi etik yapay zekâ uygulamalarının güvenilir merkezlerinden biri haline getirmeyi amaçlamaktadır. Türkiye'nin dijital egemenliği, ancak bu tür güçlü kurumsal ilkeler, etik inovasyon modelleri ve sürdürülebilir iş standartlarıyla korunabilir. Yapay zekâ teknolojileri böylece, toplumsal fayda ve ticari değer dengesini koruyarak güven içinde gelişecektir.

AIPA, bu dönüşümün bağımsız gözlemcisi, stratejik rehberi ve özel sektörün etik pusulası olmaya devam edecektir.

Referans Alınan Ulusal/Uluslararası Çerçeve, Rehber ve Standartlar

Adı	**Açıklama**
OECD AI Principles	Policy and Governance Framework
OECD G7 Hiroshima AI Process	AI Code of Conduct and Reporting Framework
EU AI Act	Code of Practice, AI Cards, AI Compliance Checker
ISO/IEC 42001:2023	AI Management System (AIMS)
ISO/IEC 5338:2023	AI System Life Cycle Processes
IEEE 7000 series	Institute of Electrical and Electronics Engineers - Ethically Aligned Design/Ethics in Autonomous & Intelligent Systems Standards Family (7001, 7002, 7003 etc.) - **USA**
NIST AI RMF	National Institute of Standards and Technology AI Risk Management Framework - **USA**
GPAI	Global Partnership for Artificial Intelligence Guidelines, Technical Reports – **CANADA + FRANCE + OECD**
BSI AI Audit Guidelines	British Standards Institution AI Audit Guidelines, AI Lifecycle Management, Assurance Frameworks - **UK**
CNIL AI Auditing Guidelines	“Commission Nationale de l'Informatique et des Libertés” AI Auditing Guidelines - **FRANCE**
ICO AI Auditing Framework	Information Commissioner’s Office AI Auditing Framework - Explainability Guidelines, Data Protection Impact Assess. - **UK**
KVKK	Kişisel Verilerin Korunması Kanunu - **TÜRKİYE**
GDPR	General Data Protection Regulation – **EU**
C2PA	Coalition for Content Provenance and Authenticity