Türkiye Bilişim Derneği

YAPAY ZEKÂ STANDARTLARI VE DÜZENLEMELERİ

Türkiye İçin Yol Haritası

EDİTÖRLER :

Soner KARATAŞ

Ülkü BAYER

İ. İlker TABAK

YAPAY ZEKÂ STANDARTLARI VE DÜZENLEMELERİ

Türkiye İçin Yol Haritası

Türkiye Bilişim Derneği

Yapay Zekâ Standartları ve Düzenlemeleri: Türkiye İçin Yol Haritası

Yayımcı Adı

TÜRKİYE BİLİŞİM DERNEĞİ
Ceyhun Atuf Kansu Cad., 1246 Sk. No: 4/17 Balgat – ANKARA
Tel: +90 (312) 473 8215 (pbx) Faks: +90 (312) 473 8216
tbd-merkez@tbd.org.tr

Etkinlik

KamuBİB'28 – BİMY'32

Yayın Tarihi

18 Kasım 2025, Ankara

Editörler

Soner Karataş TBD YZ Standartları ÇG Başkanı.
Ülkü Bayer TBD YZ Standartları ÇG Bşk. Yrd.
İ. İlker Tabak TBD YZ Standartları ÇG Editör Üye; TBK Bilişim AŞ, YK Bşk.

İç/Dış Kapak Tasarımları

Mert Çobanov, Refik Anadol Studio

TBD Yayın Numarası : 2025 / 11-1

ISBN :

Sorumluluk Beyanı:

Bu raporda yer alan her bölüm, ilgili bölümün altında belirtilen yazar(lar) tarafından hazırlanmıştır. Her yazar, kendi bölümündeki içerik, görüş, analiz ve yorumlardan münhasıran sorumludur. Raporun genel bütünlüğü editoryal olarak gözden geçirilmiş olmakla birlikte, bölümlerdeki bilimsel veya teknik değerlendirmeler yazarların kişisel görüşlerini yansıtır.

Türkiye Bilişim Derneği

YAPAY ZEKÂ STANDARTLARI VE DÜZENLEMELERİ

Türkiye İçin Yol Haritası

EDİTÖRLER

Soner KARATAŞ

Ülkü BAYER

İ. İlker TABAK

Kasım 2025, Ankara

YAZARLAR

Bölüm 1: Standartlar, Süreçler ve İşbirliği

Ülkü BAYER
TBD Ankara Şubesi, Sistekno Bilişim Ltd.
ubayer@koplan.com.tr

Dr. Merve Ayyüce KIZRAK
Kurucu & CEO, HUX AI
ayyucekizrak@gmail.com

Bölüm 2: Etik Yapay Zekâ ve Sosyal Bilimler

Dr. Duygu ÖZALP
Başkent Üniversitesi, Sosyoloji Bölümü
h.duygu.ozalp@gmail.com

Ulaş ÇAKIR
Savunma Sanayi Başkanlığı
ulas_cakir@yahoo.com

Nidanur CABBAR
TBD Genç Denetleme Koordinatörü,
kadriyenidanurc@gmail.com

Bölüm 3: Yapay Zekâ Süreçlerinde Hukuk

Dr. Burak GÖRENTAŞ
Van Yüzüncü Yıl Üniversitesi, Başkale MYO
burakgorentas@yyu.edu.tr

Av. Özge EVCİ ERALP
Eralp Avukatlık Bürosu
ozge@eralp.av.tr

Bölüm 4: Yapay Zekâ Süreçlerinde İnsan Faktörü

Ülkü BAYER
TBD Ankara Şubesi, Sistekno Bilişim Ltd.
ubayer@koplan.com.tr

Prof. Dr. Serkan GÜNEŞ
Gazi Üniversitesi, Endüstriyel Tasarım Bölümü
serkangunes@gazi.edu.tr

Dr. Serkan ALKAN
ODTÜ Tasarım Fabrikası Yardımcı Direktörü
serkanalkan@live.com

Bölüm 5: Yapay Zekâ Sistemleri ve Teknik Süreçler

Ulaş ÇAKIR (5.1)
Savunma Sanayi Başkanlığı
ulas_cakir@yahoo.com

Dr. Cemal GEMCİ (5.2)
TBD Yönetim Kurulu Üyesi (Sayman), OSTİM Teknik Üniversitesi
cgemci@gmail.com

Sinan SİLİM (5.3)
Şans Girişim Ortak Girişimi & CEO, Thalvera Cyber Security
silim.sinan@gmail.com

Doç. Dr. Meltem ERYILMAZ (5.4)
TBD Yönetim Kurulu Üyesi, OSTİM Teknik Üniv. Yazılım Müh. Bölümü Öğr. Üyesi
meltem.eryilmaz@ostimteknik.edu.tr

Prof. Dr. Gökçe Nur YILMAZ (5.4)
TBD Yönetim Kurulu Üyesi, TED Üniversitesi Bilgisayar Müh. Bölüm Bşk.
nur.gkc@gmail.com

Mert ÇOBANOV (5.5)
Refik Anadol Stüdyo
mertcobanov@gmail.com

Dr. Merve Ayyüce KIZRAK (5.6)
Kurucu & CEO, HUX AI
ayyucekizrak@gmail.com

Bölüm 6: Yapay Zekâ Sistemleri Yaşam Döngüsü Sürecinde Kalite Uzmanlık Alanının Değerlendirilmesi

Dr. Hatice Merve KARATAŞ
TBD Ankara Şubesi Yönetim Kurulu Üyesi, Türk Standartları Enstitüsü (TSE)
merveolmez@gmail.com

Soner KARATAŞ
TBD YZ Standartları ÇG Başkanı
sonerkaratas@gmail.com

Bölüm 7: Yönetişim

Mehmet Ali GÜNAY
Çevre Şehircilik ve İklim Değişikliği Bakanlığı
YZ ve Yenilikçi Teknolojiler Dairesi Başkanı

Dr. Merve Ayyüce KIZRAK
Kurucu & CEO, HUX AI
ayyucekizrak@gmail.com

Ersin Tufan YALVAÇ
TBD Ankara Şubesi Denetleme Kurulu Başkanı
ersinyalvac@gmail.com

ÖNSÖZ

Yapay Zekâ (YZ), insanlığın bilgi üretme, karar alma ve değer yaratma biçimlerini yeniden tanımlayan tarihsel bir dönüşümün merkezinde yer almaktadır. Bu dönüşüm, yalnızca teknolojik bir atılım değil; aynı zamanda toplumların etik pusulalarını, hukuki çerçevelerini, kurumsal süreçlerini ve insan odaklılık anlayışlarını yeniden şekillendiren çok boyutlu bir evrimdir. Türkiye olarak bu dönüşümün edilgen bir takipçisi değil, yön veren aktörlerinden biri olma sorumluluğumuz olduğuna inanıyoruz.

Türkiye Bilişim Derneği (TBD) çatısı altında hazırladığımız *"Yapay Zekâ Standartları ve Düzenlemeleri: Türkiye İçin Yol Haritası"* başlıklı rapor bu vizyonun bir ürünü olarak ortaya çıkmıştır. Bu çalışma, ISO/IEC 5338 başta olmak üzere uluslararası standartların değerlendirilmesi, AB düzenlemelerinin analizi, güvenilirlik ve risk yönetimi çerçeveleri, etik ilkeler, yönetişim modelleri, insan faktörü ve YZ yaşam döngüsünün tüm aşamalarına ilişkin kapsamlı bir rehber sunmaktadır. Amacımız, Türkiye'nin YZ ekosistemine yalnızca bugünün ihtiyaçları için değil, geleceğin bilgi toplumuna liderlik edecek bir çerçeve kazandırmaktır.

Bu raporun her bölümü, farklı disiplinlerden gelen uzmanların bilgi birikimi ve deneyimlerinin birleşimiyle şekillenmiştir. Kamu kurumlarımızdan özel sektöre, akademiden genç araştırmacılara uzanan geniş bir ekosistemin ortak üretimi, çalışmaya hem bilimsel derinlik hem de uygulama gerçekliği kazandırmıştır. Editörler olarak bizler, bu ortak aklın Türkiye'nin sayısal (dijital) geleceği için en güçlü sermaye olduğuna inanıyoruz.

Bu çalışma yedi ana bölümden oluşmaktadır. Birinci bölümde standartlar, süreçler ve işbirliği konuları ele alınmıştır. İkinci bölümde, etik YZ ve sosyal bilimler ilişkisine yer verilmiştir. Üçüncü bölümde YZ süreçlerinde hukuk konusu işlenmektedir. Dördüncü bölümde YZ süreçlerinde insan faktörünün rolü irdelenmiştir. Beşinci bölümde YZ sistemleri teknik yönden ele alınmıştır. Altıncı bölümde YZ sistemleri yaşam döngüsünde kalite uzmanlık alanı değerlendirilmiştir. Son olarak Türkiye için önerileri de içeren yönetişim konusu işlenmiştir.

Bu çalışmanın, Türkiye'nin yapay zekâ alanında kendi standartlarını geliştiren, güvenilir ve etik YZ ekosistemleri oluşturan ve küresel düzeyde söz sahibi bir ülke olmasına katkı sağlayacağına yürekten inanıyoruz. Gelecek, yalnızca teknolojiyi geliştirenlerin değil; onu doğru, sorumlu ve insanlık yararına kullananların olacaktır.

Saygılarımızla,

Yazarlar adına editörler:

Soner Karataş

Ülkü Bayer

İ. İlker Tabak

TEŞEKKÜR

Türkiye Bilişim Derneği Genel Başkanı Sayın **Kenan Nurhan ALTINSAAT** Beyefendiye bu çalışmanın yapılması için her türlü destek ve katkıyı sunmuş olduğu için;

"Yapay Zekâ Standartları ve Düzenlemeleri: Türkiye İçin Yol Haritası" raporunun iyi bir şekilde hazırlanmasına uygun ortam hazırladıkları, verdikleri çok değerli ve önemli görüşler, çalışma sırasında sabırlı ve içten destekleri için TBD Genel Başkan Yardımcısı Sayın **Gökhan ERZURUMLUOĞLU** ve TBD Ankara Şubesi Başkanı Sayın **Bülent BORAN** Beyefendilere ve tüm değerli bilgi ve destekleri için çalışmalara özveri ile katılan ve aktif katkı sağlayan Çalışma Grubu üyelerine ve Bölüm Yazarlarına editörler olarak içten ve sonsuz şükranlarımızı sunuyoruz. Ayrıca, yazarlarımız arasında olmasalar da, Çalışma Grubu çalışmalarına aktif olarak katılıp değerli görüş, öneri ve yönlendirmeleriyle yazarlarımız ve editörlere önemli katkılar sunan Sayın **Prof. Dr. Sedat IŞIK,** Sayın **Prof. Dr. Kürşat ÇAĞILTAY** ve Sayın **Dr. Öğr. Ü. Ziya KARAKAYA** Beyefendilere de teşekkürü borç biliriz.

Bu çalışmanın içinde bulunduğumuz 21. Yüzyılın standartlar, mevzuat ve düzenlemelerle desteklenen teknolojik dönüşümün önemli bir aracı olan Yapay Zekâ ile başlayan, "Y-Zamanı"[1] olarak adlandırılan yeni dönüşüm çağında ülkemizin önemli küresel oyuncularla rekabet edebilmesinde yararlı olacağını öngörmekteyiz.

Çalışma Grubu üyelerinden ve sektörümüzden gelen görüşlerin yanı sıra pek çok ulusal ve uluslararası belge ile istatistiksel bilgi ve araştırmalardan yararlanılarak hazırlanan bu raporun ülkemizin Yapay Zekâ teknolojilerini yalnızca kullanan ve tüketenler arasında değil, YZ'nin üretkenliği için onu geliştiren ve üretenler arasında yer almasını sağlayacağını umut ediyoruz.

Bir teşekkür de Yapay Zekâ'ya... Bölümler arasındaki tutarlık denetimleri başta olmak üzere, ana bölümler dışında kalan kısımların hazırlanmasında yararlanmış olduğumuz YZ teknolojilerine de teşekkür ederiz.

Üretken YZ teknolojilerini kullanarak iç ve dış kapak tasarımlarını yapan Sayın **Mert ÇOBANOV** Beyefendiye ve Çalışma Grubu toplantı tutanaklarının hazırlanmasında emeğini esirgemeyen Sayın **Behiye İlayda SELÇUK** Hanımefendiye de özel teşekkürlerimizi iletiriz.

[1] https://www.bilisimdergisi.org.tr/bilisim-dergisi-sayi-196#flipbook-df_5760/21/

TBD Ankara Şubesi Yapay Zekâ Standartları ve Düzenlemeleri Çalışma Grubu üyeleri olarak bilişim teknikbilimi ile kalkınma vizyonu kapsamında ülkemizin aydınlık geleceğine Yapay Zekâ çağında da katkı vermiş olmanın haklı gururu ile görevimizi yapmış olmanın mutluluğunu her zaman duyacağız…

Editörler

Soner KARATAŞ, Ülkü BAYER, İ. İlker TABAK

18 Kasım 2025, Ankara

Bilişim Teknikbilimini ulusal bir kalkınma aracı olarak kullanacağız…

Aydın KÖKSAL, 1968

TBD Kurucusu ve Onursal Başkanı

İçindekiler

Rapora Katkı Verenler ve Çalışma Grubu Üyeleri

Yapay Zekâ Standartları ve Düzenlemeleri konusundaki bu çalışmaya sivil toplum kuruluşlarının temsilcileri ile kamu, özel sektör ve üniversitelerin temsilcileri katkı vermişlerdir. Bu raporun içeriğinde ve hazırlanmasında katkılarını sunan katılımcılar aşağıda yer almaktadır.

Kenan Nurhan ALTINSAAT	*TBD Genel Başkanı*
Gökhan ERZURUMLUOĞLU	*TBD Genel Başkan Yardımcısı*
Bülent BORAN	*TBD Ankara Şubesi YK Başkanı*

Editörler

Soner KARATAŞ	TBD Ankara Şb. Üyesi / TBD YZ Standartları ÇG Başkanı
Ülkü BAYER	TBD Ankara Şb. Üyesi / TBD YZ Standartları ÇG Başkan Yrd.
İ. İlker TABAK	*TBD Eski YK Başkanı / Denetleme K., TBK Bilişim AŞ, YK Bşk.*

Yazarlar

Dr. Serkan ALKAN	ODTÜ Tasarım Fabrikası Yardımcı Direktörü
Ülkü BAYER	TBD Ankara Şb. Üyesi / TBD YZ Standartları ÇG Başkan Yrd.
Nidanur CABBAR	TBD Genç Denetleme Koordinatörü
Ulaş ÇAKIR	Savunma Sanayi Başkanlığı
Mert ÇOBANOV	Refik Anadol Stüdyo
Av. Özge EVCİ ERALP	Eralp Avukatlık Bürosu
Doç. Dr. Meltem ERYILMAZ	TBD YK Üyesi, OSTİM Teknik Üniv. Yazılım Müh. Bölümü
Dr. Cemal GEMCİ	TBD YK Üyesi (Sayman), OSTİM Teknik Üniv.
Dr. Burak GÖRENTAŞ	Van Yüzüncü Yıl Üniv., Bilişim Hukuku
Mehmet Ali GÜNAY	Çevre Şeh. ve İklim Değ. Bak., YZ ve Yenilikçi Tekn. Daire Bşk.
Prof. Dr. Serkan GÜNEŞ	Gazi Üniv. Endüstriyel Tasarım Bölümü
Dr. Merve Hatice KARATAŞ	TBD Ankara Şubesi YK Üyesi, TSE
Soner KARATAŞ	TBD Ankara Şb. Üyesi / TBD YZ Standartları ÇG Başkanı
Dr. Merve Ayyüce KIZRAK	Kurucu & CEO, HUX AI
Dr. Duygu ÖZALP	Başkent Üniversitesi, Sosyoloji Bölümü
Sinan SİLİM	Şans Girişim Ortak Girişimi & CEO, Thalvera Cyber Security
Ersin Tufan YALVAÇ	TBD Ankara Şubesi Denetleme Kurulu Başkanı
Prof. Dr. Gökçe Nur YILMAZ	TBD YK Üyesi, TED Üniversitesi Bilgisayar Müh. Bölümü Bşk.

YZ Standartları Çalışma Grubu Diğer Üyeleri

TBD yöneticileri, editörler ve yazarlar dışında çalışmalara katkı veren, görüş ve önerilerini paylaşan diğer çalışma grubu üyeleri aşağıda yer almaktadır.

Prof. Dr. Sedat IŞIK	Gazi Üniversitesi
Prof. Dr. Kürşat ÇAĞILTAY	TED Üniversitesi, Eğitim Fakültesi Dekanı
Dr. Öğr. Ü. Ziya KARAKAYA	THK Üniversitesi, Rektör Danışmanı, Bilgisayar Müh. Böl. Bşk.
Behiye İlayda SELÇUK	TBD Genç
Eda Nur BAYSAL	TBD Genç
Aylin Nur AKDORA	TBD Genç
Hasan Hüseyin AYGAR	TBD Genç
Kerem ÖZSÜ	TSE Yapay Zekâ Ayna Komite Üyesi, Event Gates
Ahmet Münir KOCAMAN	Çevre, Şehircilik ve İklim Değ. Bak.lığı YZ Yazılım Müh.
Erkan UÇANER	Tarım ve Orman Bakanlığı, Daire Başkanı
Gamze CİMİLLİ	Sağlık Bakanlığı, YZ ve Yenilikçi Teknolojiler D. Bşk.

Kısaltmalar

A/B Testleri	Karşılaştırmalı Deneysel Testler
AB	Avrupa Birliği
ABD / US	Amerika Birleşik Devletleri
ADVERSARYAL	Düşmanca – Karşıt Saldırı veya Karşıt Girdi Temelli, Çekişmeli
AIA	Yapay Zekâ Yasası
AIIA	Yapay Zekâ Etki Değerlendirmesi
AIM	Yapay Zekâ Yönetim Sistemleri
AIOPS	BT Operasyonlarında YZ Uygulamaları
AİD	Algoritmik Etki Değerlendirmesi
ALARP	Kabul Edilebilir Risk Tolerans Seviyelerinin Belirlenmesi
API	Uygulama Programlama Arayüzü
AR-GE	Araştırma Geliştirme
ART	Düşmanca Saldırılara Karşı (Adversary) Dayanıklılık Araç Takımı
ASIC	Uygulamaya Özel Tümleşik Devre
AWI	Onaylanmış Çalışma Önerisi
BentoML	Tek Paketle Model Üretim Dağıtımı Platformu
BT	Bilgi Teknolojileri
BTK	Bilgi Teknolojileri ve İletişim Kurumu
CAD	Bilgisayar Destekli Tasarım
CAPA	Düzeltici ve Önleyici Faaliyet
CAPTUM	PyTorch Açıklanabilirlik Kütüphanesi
CD	Komite Taslağı
CE	Avrupa Uygunluğu
CEN	Avrupa Standardizasyon Komitesi
CENELEC	Avrupa Elektroteknik Standardizasyon Komitesi
CHF	İsviçre Frangı
CI/CD	Sürekli Entegrasyon ve Sürekli Teslimat (ya da Sürekli Dağıtım)
COREMLTOOLS	Apple Core ML Dönüştürme Araçları
CRISP-DM	Veri Madenciliği için Endüstriler Arası Standart Süreç Modeli
ÇG	Çalışma Grubu
DAST	Dinamik Uygulama Güvenliği Testi

DDOS	Dağıtılmış Hizmet Engelleme Saldırısı
DevSecOps	Geliştirme, Güvenlik ve Operasyon Entegrasyonu Yaklaşımı
DIN	Alman Standartlar Enstitüsü
DIS	Uluslararası Standart Taslağı
DKE	Alman Elektroteknik, Elektronik ve Bilgi Teknolojileri Komisyonu
DO – RTCA	Havacılık Tasarımı Standardı Belgesi
DOD	ABD Savunma Bakanlığı
DPIA	Veri Koruma Etki Değerlendirmesi
DRIFT	Veri veya Kavramsal Kayma (Model Performans Bozulması)
DRM	Dijital (Sayısal) Haklar Yönetimi
EDPB	Avrupa Veri Koruma Kurulu
EMC	Elektromanyetik Uyumluluk
EU AI Act	AB YZ Yasası
EUIPO	Avrupa Birliği Fikri Mülkiyet Ofisi
Evidently	Veri ve Model Kalitesi İzleme Aracı
FastAPI	Hızlı Uygulama Programlama Arayüzü (Python web framework)
FDIS	Nihai Uluslararası Standart Taslağı
FPGA	Sahada Programlanabilir Kapı Dizisi
FSEK	Fikir ve Sanat Eserleri Kanunu
GAI	Üretken Yapay Zekâ
GAN	Üretici Çekişmeli Ağ
GDPR	Genel Veri Koruma Tüzüğü
GPAI	Genel Amaçlı YZ Sistemleri
GPU	Grafik İşleme Birimi
GRAFANA	Görselleştirme ve Gözetim Panosu Aracı
gRPC	Google Uzak Prosedür Çağrısı Protokolü
IASR-2025	Uluslararası YZ Güvenlik Raporu
IEC	Uluslararası Elektroteknik Komisyonu
IEEE	Elektrik ve Elektronik Mühendisleri Enstitüsü
INT8	8-bit Tamsayı Nicemleme (Model Sıkıştırma Tekniği)
IP	İnternet Protokolü
IS	Uluslararası Standart

ISO	Uluslararası Standartlar Enstitüsü
ISTIO	Ağ Servisleri Denetim Katmanı
JRC	Ortak Araştırma Merkezi
JTC	Ortak Teknik Komite
KOSGEB	Küçük ve Orta Ölçekli İşletmeleri Geliştirme ve Destekleme İdaresi Başkanlığı
KPI	Anahtar Performans Göstergesi
KQI	Anahtar Kalite Göstergesi
KRI	Anahtar Risk Göstergesi
KServe	Kubernetes Tabanlı Model Servisleme Çerçevesi
KVK	Kişisel Verilerin Korunması
KVKK	Kişisel Verilerin Korunması Kanunu
KYZ	Konumlandırma, Yönlendirme, Zamanlama
LIME	Yerel Olarak Yorumlanabilir Modelden Bağımsız Açıklamalar
LLM	Büyük Dil Modellerinde
MIL-STD	ABD Savunma Bakanlığı Askerî Standardı
ML	Makine Öğrenmesi
MLflow	Makine Öğrenimi Akış Yönetim Aracı
MLOps	Makine Öğrenimi Operasyonları
MTC	Ayna Teknik Komite
mTL	Karşılıklı Taşıma Katmanı Güvenliği
NCSC	Ulusal Siber Güvenlik Merkezi
NIST AI RMF	Ulusal Standartlar ve Teknoloji Enstitüsü Yapay Zekâ Risk Yönetim Çerçevesi
NLP	Doğal Dil İşleme
NP	Yeni Çalışma Maddesi Önerisi
ODTÜ	Orta Doğu Teknik Üniversitesi
OECD	Ekonomik İş birliği ve Kalkınma Örgütü
OMÜ	Ondokuz Mayıs Üniversitesi
ONNX	Açık Sinir Ağı Değişim Formatı
OPACUS	PyTorch Diferansiyel Mahremiyet Kütüphanesi
OpenVINO	Açık Görsel Çıkarım ve Sinir Ağı Optimizasyon Araç Seti (Intel)
OWASP	Açık Kaynaklı Dünya Çapında Uygulama Güvenliği Projesi

PII	Kişisel Olarak Tanımlanabilir Bilgi
PLD	Programlanabilir Mantık Aygıtı
P-U-K-Ö	Planla-Uygula-Kontrol Et-Önlem Al
QA	Kalite Güvence
RAI Toolkit	Sorumlu Yapay Zekâ Araç Takımı
RDF	Kaynak Tanımlama Çerçevesi
REST	Temsili Durum Aktarımı (Web servis mimarisi)
ROI	Yatırım Getirisi
SAST	Statik Uygulama Güvenliği Testi
SC	Alt Komite
SDK	Yazılım Geliştirme Kiti
SDLC	Yazılım Geliştirme Yaşam Döngüsü
SDO	Standart Geliştirme Kuruluşları
SKAIA	Güney Kore YZ'nin Geliştirilmesi ve Güvenin Sağlanmasına İlişkin Temel Kanun
SLA	Hizmet Düzeyi Anlaşması
SLO	Hizmet Düzeyi Hedefi
SOAR	Güvenlik Orkestrasyonu, Otomasyon ve Müdahale Sistemi
SOC	Güvenlik Operasyon Merkezi
SOP	Standart Çalışma Prosedürü
SQuaRE	Sistemler ve Yazılımların Kalite Gereksinimleri ve Değerlendirmesi
STK	Sivil Toplum Kuruluşu
STRIDE	Kimlik Sahteciliği, Veri Manipülasyonu, İnkâr Edilebilirlik, Bilgi İfşası, Hizmet Engelleme ve Yetki Yükseltme Tehdit Modeli
TBD	Türkiye Bilişim Derneği
TBK	Türk Borçlar Kanunu
TensorFlow	Makine Öğrenimi Çerçevesi
TensorRT	NVIDIA Tensör Çalışma Zamanı (Model Hızlandırıcı)
TGI	Metin Üretimi Çıkarım Sunucusu (Hugging Face)
TPU	Tensör İşlem Birimi
TR	Teknik Rapor
TS	Teknik Şartname
TSE	Türk Standartlar Enstitüsü

TUCBS	Türkiye Ulusal Coğrafi Bilgi Sistemi Altyapısı
TÜBİTAK	Türkiye Bilimsel ve Teknik Araştırma Kurumu
TÜBİTAK BİLGEM	TÜBİTAK Bilişim ve Bilgi Güvenliği İleri Teknolojiler Araştırma Merkezi
TÜİK	Türkiye İstatistik Kurumu Başkanlığı
UK IPO	Birleşik Krallık Fikri Mülkiyet Ofisi
UML	Birleşik Modelleme Dili
UNDP	Birleşmiş Milletler Kalkınma Programı
UNESCO	Birleşmiş Milletler Eğitim, Bilim ve Kültür Örgütü
UX	Kullanıcı Deneyimi
VBE	Değer Odaklı Mühendislik
vLLM	Büyük Dil Modelleri için Hızlı Çıkarım Çalışma Ortamı
WAF	Web Uygulama Güvenlik Duvarı
WIPO	Dünya Fikri Mülkiyet Örgütü
XAI	Açıklanabilir Yapay Zekâ
XML	Genişletilebilir İşaretleme Dili
YZ	Yapay Zekâ
YZYS	Yapay Zekâ Yönetim Sistemi

Bilişim Toplumu Yasalarıyla Yönetelim, Yönetilelim

Türkiye Bilişim Derneği

SUNUŞ

Çağdaş, refah düzeyi yüksek ve teknolojiyle güçlenen bir Türkiye idealine ulaşmak için Türkiye Bilişim Derneği (TBD) olarak 54 yıldır kararlılıkla çalışıyoruz. 1994 yılında Kamu Yararına Çalışan Dernek statüsünü almış olan TBD; ülkemizin sayısal dönüşümünü hızlandırmak, bilişim kültürünü yaygınlaştırmak ve teknoloji alanındaki ulusal kapasiteyi güçlendirmek amacıyla binlerce üyemiz, gönüllülerimiz ve gençlik yapılanmamızla birlikte büyük bir toplumsal sorumluluk üstlenmektedir.

Günümüzde bilişim ekosistemi hiç olmadığı kadar hızlı bir dönüşüm yaşamaktadır. Yapay Zekâ (YZ) teknolojilerinin kamu yönetiminden özel sektöre, üretimden sağlık ve eğitim hizmetlerine kadar tüm alanlara nüfuz etmesi, bu teknolojilerin **güvenli, etik, adil ve insan odaklı** biçimde geliştirilmesi gerekliliğini tartışmasız hale getirmiştir. YZ'nin hızlı yayılımı yalnızca teknik ilerlemeyi değil; mahremiyet, insan hakları, hukuki sorumluluk, toplumsal güven, açıklanabilirlik ve hesap verebilirlik gibi kritik alanları da doğrudan etkilemektedir. Bu nedenle YZ'nin hem ulusal hem uluslararası normlarla uyumlu biçimde standardize edilmesi artık stratejik bir zorunluluktur.

Bu gereksinimden hareketle TBD, KamuBİB'28 ve BİMY'32 etkinlikleri sırasında kamuoyu ile paylaşılmak üzere **YZ Standartları ve Düzenlemeleri Raporu**nu hazırlamak amacıyla alanında uzman akademisyenler, sektör profesyonelleri, hukukçular, mühendisler ve kamu yöneticilerinden oluşan bir **Çalışma Grubu** oluşturmuştur. Bu çalışma grubu tarafından hazırlanan rapor, Türkiye'de yapay zekâ alanında güvenilir, denetlenebilir ve uluslararası standartlarla uyumlu bir ekosistem oluşturmak için gerekli teknik, etik ve yönetişimsel çerçeveyi bütüncül bir bakış açısıyla ele almaktadır.

Çalışmada ISO/IEC 5338 gibi yapay zekâ yaşam döngüsü standartlarını, ISO/IEC 24028 ve 23894 gibi güvenilirlik ve risk yönetimi çerçevelerini, AB Yapay Zekâ Yasası (AI Act) ile uluslararası düzenleyici yaklaşımları, etik ilkeleri, veri koruma yükümlülüklerini ve insan faktörünü kapsamlı biçimde incelenmiştir. Bu çalışma, ülkemizin sadece teknolojiyi tüketen bir ülke değil, **kendi standartlarıyla küresel ölçekte söz sahibi olan, teknoloji üreten** bir yapıya kavuşması için önemli bir adım niteliğindedir.

Özellikle kamu kurumlarımızın YZ temelli sistemleri giderek daha fazla kullanmaya başlaması, standartlara dayalı bir yaklaşımı zorunlu kılmaktadır. YZ projelerinin tedariğinden geliştirilmesine, test edilmesinden işletilmesine ve imhasına kadar tüm yaşam döngüsünün **uluslararası tanımlı süreçlerle uyumlu hale getirilmesi**, hem kamu güveni hem de hizmet kalitesi açısından kritik önemdedir. Bunun yanında etik denetim mekanizmaları, açıklanabilirlik ilkeleri, veri mahremiyeti ve güvenlik gereksinimleri de raporda güçlü biçimde vurgulanmıştır.

TBD olarak, bu çalışmayla ortaya konan teknik ve politik önerilerin, hem kamu kurumlarımız hem düzenleyici otoritelerimiz hem de özel sektör kuruluşlarımız tarafından dikkate alınmasının ülkemiz adına stratejik bir gereklilik olduğuna inanıyoruz. Bu kapsamda, raporda yer alan değerlendirmeleri ve çözüm önerilerini ilgili tüm bakanlıklarımızla, kamu kurumlarımızla ve ekosistemin diğer paydaşlarıyla paylaşmayı, etkinliklerimizde yer vermeyi planlamaktayız. Amacımız, kurumlarımız arasında ortak akıl oluşturmak ve somut eylem planları geliştirmektir.

Bu önemli çalışmada katkı sunan tüm uzmanlarımıza, çalışma grubu üyelerimize, akademisyenlerimize ve genç bilişimcilerimize teşekkür ediyor; raporun Türkiye'nin yapay zekâ alanındaki dönüşüm yolculuğuna değerli bir katkı sağlayacağına yürekten inanıyorum.

Kenan Nurhan ALTINSAAT

Türkiye Bilişim Derneği
Genel Başkanı

Türkiye Bilişim Derneği

Yapay Zekâ Standartları ve Düzenlemeleri

Türkiye İçin Yol Haritası

YÖNETİCİ ÖZETİ

Türkiye Bilişim Derneği (TBD) tarafından hazırlanan *Yapay Zekâ Standartları ve Düzenlemeleri Raporu (2025)*, Türkiye'nin Yapay Zekâ (YZ) ekosisteminin güvenli, etik, hesap verebilir ve uluslararası standartlarla uyumlu şekilde gelişmesi için kapsamlı bir yol haritası sunmaktadır. Rapor; teknik standartlardan etik ilkelere, hukuki düzenlemelerden yönetişim modellerine, insan faktöründen güvenlik standartlarına kadar bütüncül bir çerçeve ortaya koymaktadır.

1. Yapay Zekâ Sistemlerinin Karmaşıklığı ve Standartlara Duyulan İhtiyaç

YZ sistemleri, teknik, sosyal, etik ve hukuki boyutlarıyla çok-disiplinli bir yapıya sahiptir. Rapor, bu nedenle **ortak bir yöntem, standart terminoloji ve süreç yönetimi** gerekliliğini vurgulamaktadır. Uluslararası standartlar, ülkeler ve sektörler arası uyumu sağlayarak hem güveni hem de sürdürülebilir yatırımı güçlendirmektedir.

Özellikle **ISO/IEC 5338** standardı, YZ sistem yaşam döngüsünü baştan sona tanımlayan ana çerçeve olarak öne çıkmaktadır.

2. ISO/IEC 5338 ve YZ Yaşam Döngüsünün Standartlaştırılması

Rapor, ISO/IEC 5338'i tüm YZ projelerinin temel çerçevesi olarak ele almakta ve dört ana süreç grubunu detaylandırmaktadır:

- **Anlaşma süreçleri:** YZ ürünlerinin tedarik/satın alma süreçleri için sözleşmesel ve kalite güvencesi gereklilikleri.
- **Kurumsal proje-olanak sağlayıcı süreçler:** Altyapı, insan kaynağı, bilgi ve kalite yönetimi.
- **Teknik yönetim süreçleri:** Planlama, kontrol, risk yönetimi, ölçüm ve kalite güvencesi.
- **Teknik süreçler:** Veri toplama, model geliştirme, doğrulama, test, geçerleme, bakım ve imha.

Bu yapı, kurumlara **tutarlı, denetlenebilir ve tekrarlanabilir** bir yönetişim olanağı sağlamaktadır.

3. Etik Çerçeveler ve Sorumlu YZ İlkelerinin Kurumsallaştırılması

Raporda etik boyutu teknik standartlardan bağımsız olarak ele alınmamakta, bütünleşik olarak değerlendirilmektedir.

Temel ilkeler:

- Adalet ve önyargıların azaltılması,
- Şeffaflık ve açıklanabilirlik,
- Hesap verebilirlik,
- İnsan kontrolü ve gözetimi,
- Güvenilirlik ve güvenlik,
- Mahremiyetin korunması.

Etik yönetişim için önerilen mekanizmalar:

- Etik denetim kurulları,
- Algoritmik etki değerlendirmeleri,
- Veri setlerinde önyargı analizleri,
- Açıklanabilirlik dokümantasyonu.

Bu başlıklar özellikle raporun 3.3 bölümünde ayrıntılı olarak ele alınmaktadır.

4. Hukuki Çerçeve: KVKK – GDPR – AB YZ Yasası Uyumu

Rapor, Türkiye'nin mevcut hukuk sisteminin YZ uygulamalarını karşılamakta **eksik kaldığını**, özellikle **yüksek riskli YZ uygulamalarının** daha sıkı denetime tabi tutulması gerektiğini ortaya koymaktadır.

Karşılaştırılan düzenlemeler:

- **KVKK** – daha genel çerçeve, YZ bağlamında sınırlı.
- **GDPR** – veri koruma açısından güçlü; YZ kullanımına doğrudan etkileri var.
- **AB YZ Yasası (EU AI Act)** – risk temelli yaklaşımıyla YZ için kapsamlı bir düzenleme.

Raporda ayrıca Türkiye için:

- KVKK–GDPR uyumlaştırması,
- AB YZ Yasası'yla **kademeli uyum**,
- Telif hakları, eğitim verisi temizliği ve açık veri mekanizmalarının yeniden düzenlenmesi

önerileri yer almaktadır.

5. İnsan Faktörü ve Kullanıcı Merkezli Tasarım

YZ başarısı, sadece teknik değil **insan odaklı tasarım** süreçlerinin uygulanmasına bağlıdır.

YZ yaşam döngüsünde insan faktörü için raporda aşağıdaki önerilere yer verilmiştir:

- Kullanım bağlamı analizleri,
- Kullanıcı gereksinimlerinin sistematik toplanması,
- Kullanıcı destekli iteratif tasarım,
- Kullanıcı deneyimi (UX) uygunluğu,
- İnsan-denetimli YZ (Human-in-the-loop) mekanizmaları.

6. Güvenlik Standartları: Güvenilir YZ için Ulusal Strateji İhtiyacı

Raporda ayrıca YZ sistemlerinin karşı karşıya olduğu güvenlik riskleri tanımlanmıştır:

- Çekişmeli (adversarial) saldırılar,
- Veri zehirleme,
- Model hırsızlığı,
- Kavramsal kayma (concept drift),
- Kritik altyapılarda operasyonel riskler.

Önerilen çerçeve:

- ISO/IEC 24028 ve ilgili güvenlik standartlarının ulusal düzeyde benimsenmesi,
- Sektörel güvenlik protokollerinin geliştirilmesi,
- Ulusal bir **YZ Güvenlik Standardı** oluşturulması gerekliliği.

7. Yönetişim Yapısı ve Kamu Kurumları İçin Politika Önerileri

Raporun önemli çıktılarından biri, Türkiye için önerilen yönetişim modelidir:

- **Ulusal Yapay Zekâ Komitesi**
- **YZ Politika ve Güvenlik Araştırma Merkezleri**
- YZ işletmeleri için **temsilci atama yükümlülüğü**
- Derin aldatma (Deepfake) ve sayısal (dijital) istismar ile mücadele araçları
- Kamu alımlarında ISO/IEC 5338 tabanlı standart zorunluluğu
- Açık veri kaynaklarının genişletilmesi

Bu öneriler hem etik hem de güvenlik açısından kamu sisteminin güçlendirilmesini hedeflemektedir.

8. Genel Değerlendirme ve Sonuç

Rapor, Türkiye'nin YZ ekosisteminde **uluslararası standartlara uyum**, **şeffaflık**, **hesap verebilirlik**, **güvenlik** ve **insan odaklı yaklaşım** ilkelerini merkeze alan bütünleşik bir stratejiye ihtiyaç duyduğunu ortaya koymaktadır.

Stratejik odak noktaları:

1. Standartlara dayalı yaşam döngüsü yönetimi
2. Etik yönetişim ve risk temelli denetim
3. Hukuki çerçevenin güncellenmesi
4. İnsan merkezli tasarım
5. Güvenlik, dayanıklılık ve şeffaflık
6. Kurumlar arası işbirliği ve ulusal koordinasyon

öne çıkmaktadır.

Bu çerçeve yaşama geçirildiğinde, Türkiye'nin hem kamu hem de özel sektörde **güvenilir, rekabetçi ve uluslararası uyumlu** bir YZ ekosistemi oluşturması sağlanabilecektir.

TEMEL TANIMLAR VE KAVRAMLAR

Bu çalışma kapsamında yer alan Avrupa Yapay Zekâ Yasası (EU AI Act) başta olmak üzere standart kavramı ve akıllı standartlar konusundaki tanım ve kavramlar, rapordan yararlanacakların aynı bilgi ve tanımlarla okuma yapmaları için sunulmaktadır.

1. Avrupa Yapay Zekâ Yasası

Avrupa Birliği'nin ilk kapsamlı yapay zekâ yasası olan **EU AI Act (Regulation (EU) 2024/1689)**, kademeli olarak yürürlüğe girmekte ve dünya genelinde referans kabul edilecek bir yasal çerçeve oluşturmaktadır. *DLA Piper*'ın 2025 tarihli raporuna[2] göre, yasa Şubat 2025'te yasaklı uygulamalara ilişkin hükümlerle başlamış, **2 Ağustos 2025'te** temel yönetişim ve denetim altyapısı devreye alınmıştır. AB bünyesinde "**AI Office"**, "**AI Board"** ve "**Bilimsel Uzmanlar Paneli"** gibi yeni kurumlar oluşturulmuş; her üye ülke, ulusal denetim ve bildirim otoritelerini belirlemekle yükümlü kılınmıştır. Bu kurumların birlikte çalışmasıyla yasa kapsamında tutarlı uygulama, risk takibi ve şeffaf denetim sağlanması amaçlanmaktadır.

Genel Amaçlı Yapay Zekâ (GPAI) modellerine odaklanmakta olan yasaya göre bu modelleri geliştiren veya piyasaya süren kuruluşlar 2 Ağustos 2025 tarihinden itibaren teknik dokümantasyon hazırlamak, telif haklarına uyum politikaları oluşturmak ve eğitim verisi özetlerini kamuya açık şekilde paylaşmak zorundadır. "Sistemik risk" taşıyan büyük modeller için (örneğin 10^{25} FLOP ve üzeri işlem gücüyle eğitilenler) ayrıca **risk yönetimi, siber güvenlik ve olay bildirim** yükümlülükleri uygulanmaktadır. Avrupa Komisyonu, şirketlerin uyum süreçlerini kolaylaştırmak için "**GPAI Code of Practice"**[3] adlı kılavuzu yayınlamış olsa da bu belge hukuken bağlayıcı değildir.

Ağustos 2025'te yürürlüğe giren **ceza rejimi**, ihlalin türüne göre 35 milyon avroya veya küresel cironun %7'sine kadar para cezası öngörmektedir. Ancak GPAI modellerine özel cezalar 2026'ya ertelenmiştir. *DLA Piper Raporu*, bu dönemi "altyapı evresi" olarak tanımlamaktadır: yasalar yürürlüktedir ancak denetim kapasitesi ve ulusal uygulama mekanizmaları hâlâ oluşum sürecindedir. Raporun genel değerlendirmesine göre AB YZ Yasası'nın *etik*, *güvenli* ve *şeffaf* YZ sistemleri için küresel ölçekte bir referans çerçeve haline

[2] https://www.dlapiper.com/en-us/insights/publications/2025/08/latest-wave-of-obligations-under-the-eu-ai-act-take-effect
[3] https://digital-strategy.ec.europa.eu/en/policies/contents-code-gpai

geldiği, fakat **asıl sınavın 2026'da** yüksek riskli sistemler için tam uyumun zorunlu hale gelmesiyle yaşanacağı öngörülmektedir.

2. Standart Nedir?

Uluslararası standartlar, yalnızca teknik süreçleri düzenleyen belgeler olmayıp aynı zamanda kamu politikalarının uygulanmasında **stratejik araçlar** haline gelmiştir. ISO'nun *Public Policy Toolkit* raporunda[4] belirtildiği gibi, standartlar; kamu politikalarının hazırlanması, uygulanması ve izlenmesinde rehberlik sağlayarak düzenlemelerin kalitesini artırmaktadır. Bir standart, belirli bir işin "en iyi şekilde nasıl yapılacağını" tanımlayan uzlaşıya dayalı bir yöntemdir. Bu yönüyle, üretimden hizmet sunumuna kadar birçok alanda kalite, güvenlik ve sürdürülebilirlik için ortak dil oluşturmaktadır.

Standartların ekonomik büyümeye katkısı da güçlü olmaktadır. Birleşmiş Milletler'in Sürdürülebilir Kalkınma Amaçları[5] doğrultusunda, uluslararası standartlar sürdürülebilir ve kapsayıcı büyümenin temel altyapısını oluşturmaktadır. Çeşitli araştırmalar, standartların bilgi paylaşımını kolaylaştırarak üretkenliği ve verimliliği artırdığını göstermiştir. Örneğin İngiltere'de yapılan bir çalışma, son 100 yıldaki GSYH büyümesinin %23'ünün ve verimlilik artışının %38'inin doğrudan standartlara bağlı olduğunu ortaya koymuştur. Böylece standartlar, yalnızca ekonomik değil, aynı zamanda sosyal refahın da taşıyıcısı haline gelmiştir.

Kamu politikaları açısından bakıldığında, standartlar hem **bilimsel temele dayalı** hem de **uluslararası uyumu kolaylaştıran** araçlardır. ISO ve IEC tarafından geliştirilen standartlar, hükümetlerin teknik düzenlemelerini sadeleştirir, uygulamada tutarlılığı sağlar ve gereksiz ticaret engellerini azaltır. Örneğin, Almanya ISO 50001 enerji yönetim sistemini uygulayan işletmelere vergi teşviki verirken, Kanada ISO 9001 kalite yönetim sistemlerini kamu ihalelerinde zorunlu hale getirmiştir. Bu tür örnekler, standartların kamu yönetimi kalitesini yükseltme, ekonomik rekabeti artırma ve sürdürülebilirlik politikalarını destekleme gücünü göstermektedir.

Uluslararası standartların kamu politikalarına entegrasyonu; **regülasyonların sadeleşmesini**, **kamu hizmetlerinde verimliliği**, **uluslararası ticarette uyumu** ve **sürdürülebilir kalkınmayı** aynı anda mümkün kılmaktadır. Bugün birçok ülke standartları yalnızca teknik gereklilik olarak değil, aynı zamanda **politik bir stratejik unsur** olarak

[4] https://www.iso.org/files/live/sites/isoorg/files/publications/en/ISO_Public-Policy-Toolkit.pdf
[5] https://turkiye.un.org/tr/sdgs

kullanmakta olup toplumda güven, kalite ve kapsayıcı kalkınma hedeflerine ulaşmada temel bir araç haline getirmektedir.

3. Akıllı Standartlar

Akıllı Standartlar (SMART Standards) kavramı, klasik standardizasyonun sayısal (dijital) dönüşümünü temsil etmektedir.[6] Artık standartlar yalnızca insanlar tarafından okunan belgeler değil, **makineler, yazılımlar ve yapay zekâ sistemleri tarafından da okunabilir, yorumlanabilir ve uygulanabilir** hâle getirilmektedir. Bu yaklaşım, standartların PDF doküman formatından çıkıp yapılandırılmış veri modeli, bilgi ağı ve makine tarafından yürütülebilir kod bileşenine dönüşmesini sağlamaktadır. Böylece firmalar, standartlardan doğrudan veri çekebilir (örneğin test yöntemleri, 3B modeller, yazılım gereksinimleri) ve bu bilgileri üretim, kalite kontrol veya doğrulama süreçlerine otomatik biçimde entegre edebilmektedir.

2022 itibarıyla Alman Standartlar Enstitüsü (DIN) ve Alman Elektroteknik, Elektronik ve Bilgi Teknolojileri Komisyonu (DKE), Avrupa standardizasyon kurumlarıyla birlikte **CEN-CENELEC çatısı altında "SMART Standards" iş akışlarını** başlatmıştır. Bu kapsamda bilgi modeli (Workstream 3), operasyonel süreç (Workstream 4) ve iş modeli (Workstream 5) alanlarında çalışmalar yürütülmektedir. Aynı zamanda ISO ve IEC düzeyinde de paralel girişimler ("ISO SMART" ve "IEC SG 12") geliştirilmiştir. Amaç, Avrupa'da geliştirilen bu sayısal (dijital) standardizasyon modelini küresel ölçekte benimsetmektir.

Bu yeni model, **IEC'nin beş seviyeli olgunluk yapısına** dayanmaktadır: düşük seviyelerde standartlar insan tarafından okunabilirken, **3. seviye** ile birlikte makine tarafından işlenebilir hâle gelmektedir. Üst seviyelerde (4–5) ise standartlar yapay zekâ tarafından **yorumlanabilir, hatta geliştirilebilir** hâle gelecektir. Almanya'daki **IDiS (Initiative Digital Standards)** projesi bu dönüşümü uygulamalı olarak test etmektedir; 40.000'den fazla DIN ve VDE standardına dayalı bir dil modeli kullanarak standart metinlerinden otomatik bilgi çıkarımı yapabilmektedir.

Bu durum yalnızca teknik bir yenilik değil, **insan faktörünü ve süreç yönetimini** de kökten dönüştürmektedir. Artık standart oluşturma, yönetim ve dağıtım süreçlerinde görev alan uzmanların sayısal (dijital) yetkinliklere sahip olması beklenmektedir. DIN bu nedenle yeni

[6] https://www.din.de/resource/blob/916798/ed09ae58b60f0d3a498fa90fa5085b7c/nrm-ki-engl-2023-final-web-250-neu-data.pdf

görev tanımları ve eğitim programları hazırlamaktadır. Ayrıca, **SAM (Standard Architecture Model)** ve **SAS (Standard Administration Shell)** adlı yeni veri modelleri, standardın sayısal (dijital) ikiz gibi yönetilmesini sağlamaktadır; yani standart artık yaşayan, güncellenebilen bir "e-nesneye" dönüşmektedir.

Ekonomik olarak, Almanya'da klasik standartların yıllık katkısı 17 milyar avro civarındadır; Akıllı (Smart) standartlarla bu katkının önemli ölçüde artması beklenmektedir. Çünkü bilgiye erişim hızlanmakta, uygulama maliyetleri düşmekte ve yapay zekâ sistemleri standart içeriklerini doğrudan işleyebilmektedir. Bu dönüşümün nihai hedefi, **2025 itibarıyla yapay zekânın yalnızca standartları kullanması değil, onların geliştirilmesine de katkı veren bir aktör hâline gelmesidir.** Böylece standartlar, insan–yapay zekâ iş birliğine dayalı, sürekli öğrenen sayısal (dijital) bilgi sistemlerine dönüşerek Endüstri 4.0'ın temel bilgi altyapısını oluşturacaktır.

GİRİŞ

Yapay Zekâ (YZ) teknolojileri, sayısal dönüşümün küresel ölçekte en stratejik bileşenlerinden biri haline gelmiş ve toplumların ekonomik, sosyal ve kültürel yapısında köklü değişimlere yol açmıştır. Günümüzde YZ sistemleri; kamu hizmetlerinden sağlık sektörüne, ulaşımdan enerjiye, finansal sistemlerden savunma ve güvenlik alanlarına kadar pek çok kritik sektörde karar verme süreçlerini destekleyen, hatta zaman zaman bu süreçleri şekillendiren temel bir unsur haline gelmiştir. Bu hızlı yaygınlaşma, beraberinde önemli fırsatlar getirirken aynı zamanda ciddi teknik, etik, hukuki ve yönetişimsel sorunların da gündeme taşınmasına neden olmaktadır.

YZ'nin geniş ölçekli etkileri, sadece teknolojik performansı değil; insan haklarını, toplumsal eşitliği, kurumların hesap verebilirliğini, güvenliğini ve hatta ulusal egemenliği ilgilendiren çok katmanlı riskler yaratmaktadır. Özellikle makine öğrenmesi tabanlı sistemlerin karmaşık ve öngörülemez davranışlar sergileyebilmesi, "algoritmik önyargı", "şeffaflık eksikliği", "beklenmeyen kullanım senaryoları" ve "güvenlik açıkları" gibi sorunların giderek artmasına yol açmaktadır. Bu nedenlerle YZ'nin yaşam döngüsü boyunca tanımlı, denetlenebilir, şeffaf süreçlerle yönetilmesi; teknik, etik ve hukuki risklerin kontrol altına alınması artık bir tercih değil zorunluluktur.

Bu rapor, Türkiye Bilişim Derneği'nin (TBD) uzun yıllara dayanan deneyimi ve kamu yararına çalışan bir sivil toplum kuruluşu kimliğinden aldığı güçle, YZ ekosisteminin ihtiyaç duyduğu **standartlar, düzenlemeler ve yönetişim yapıları** için kapsamlı bir çerçeve sunmaktadır. Rapor; kamu, özel sektör, üniversiteler, bağımsız uzmanlar ve STK'ların katkısıyla hazırlanmış olup, Türkiye'nin hem ulusal hem de uluslararası düzeyde güvenilir ve rekabetçi bir YZ altyapısı geliştirmesine destek olmayı amaçlamaktadır.

YZ'nin toplumsal etkilerinin ve risklerinin tartışıldığı çalışmaların önemine vurgu yapan TBD, bu alanda standartların ve düzenleyici çerçevelerin geliştirilmesini; kamu politikalarının, etik ilkelerin ve teknik yönergelerin bütünsel bir yaklaşım ile tanımlanmasını öncelikli bir görev olarak görmektedir. Bu kapsamda TBD, KamuBİB ve BİMY etkinlikleri çerçevesinde uzmanlardan oluşturduğu çalışma grupları aracılığıyla YZ'nin etkilerini sürekli gündemde tutmuş, politika yapıcılarla sektörün ihtiyaçlarını buluşturan bir köprü görevi üstlenmiştir. Bu rapor da söz konusu çalışmaların derinleştirilmesiyle işbirliği içinde ortaya çıkan kolektif bir değerlendirmenin ürünüdür.

YZ sistemlerinin toplumla etkileşimini belirleyen teknik standartların küresel ölçekte hızla gelişmesi, ülkelerin bu alanda proaktif davranma gerekliliğini artırmıştır. Özellikle **ISO/IEC 5338 Yapay Zekâ Yaşam Döngüsü Standardı**, kurumların YZ projelerini nasıl tasarlayacağı, yöneteceği, değerlendireceği ve sonlandıracağına yönelik kapsamlı bir süreç çerçevesi sağlamaktadır. Bu standart, geleneksel yazılım geliştirme yaşam döngüsü modellerinin ötesine geçerek YZ sistemlerine özgü gereksinimleri, belirsizlikleri ve riskleri ortaya koymaktadır. Böylelikle kurumların; veri toplama, veri mühendisliği, model eğitimi, doğrulama, validasyon, operasyon, bakım ve imha gibi kritik süreçleri uluslararası tanımlarla uyumlu şekilde yönetmesi mümkün hale gelmektedir.

Raporun temel odak noktalarından biri, YZ'nin sadece teknik bir konu olmadığı; etik, sosyolojik, hukuki ve yönetişim boyutlarıyla birlikte ele alınması gerektiği gerçeğidir. **Etik YZ ilkeleri**, açıklanabilirlik, adalet, insan denetimi, mahremiyetin korunması ve hesap verebilirlik gibi kritik başlıkları kapsamaktadır. Bu ilkelerin YZ yaşam döngüsünün tüm aşamalarına yerleştirilmesi, sistemlerin toplumsal güveni zedelemeden ve birey haklarını ihlal etmeden kullanılabilmesi açısından zorunludur. Raporun etik bölümleri, uluslararası normlar ve iyi uygulama örneklerine dayanarak gerek kamu kurumları gerek özel sektör için uygulanabilir öneriler sunmaktadır.

Bir diğer önemli başlık ise **hukuki uyum ve düzenleyici çerçeveler**dir. Avrupa Birliği Genel Veri Koruma Tüzüğü (GDPR), Türkiye'nin KVKK yapısı ve AB YZ Yasası (AI Act), YZ uygulamalarının hukuki risklerinin yönetilmesinde önemli referans çerçeveler olarak değerlendirilmiştir. Özellikle AB YZ Yasası'nın risk temelli yaklaşımının Türkiye'deki mevzuata uyarlanması, yüksek riskli YZ uygulamalarının denetlenmesi ve YZ geliştiricilerinin yükümlülüklerinin belirlenmesi raporda ayrıntılı şekilde ele alınmıştır. Eğitim verilerinde telif hakları, açık veri mekanizmaları ve sayısal içeriklerin kimlik doğrulaması gibi konular ise YZ'nin giderek karmaşıklaşan hukuki boyutlarını daha görünür hale getirmektedir.

Raporun güçlü yönlerinden biri, **insan faktörünü** YZ süreçlerinden bağımsız düşünmeyen yaklaşımıdır. İnsan merkezli tasarım ilkeleri, kullanıcı gereksinimlerinin analiz edilmesi, erişilebilirlik ilkeleri, kullanıcı deneyimi ve insan–makine etkileşiminin etik sınırları gibi konular detaylı olarak ele alınmıştır. Bu yaklaşım, özellikle kamu hizmetleri gibi kritik alanlarda YZ sistemlerinin güvenli, anlaşılır ve kapsayıcı bir şekilde tasarlanmasını desteklemektedir.

YZ'nin ulusal güvenlik ve siber güvenlik açısından doğurduğu riskler de raporda ayrı bir başlık olarak ele alınmıştır. Çekişmeli (adversarial) saldırılar, veri manipülasyonu, veri modeli zehirleme, bilgi sızıntısı ve kritik altyapılara yönelik tehditler dikkate alınarak; ulusal düzeyde YZ güvenlik standartlarının geliştirilmesi, siber güvenlik protokollerinin güncellenmesi ve kurumlar arası işbirliğinin güçlendirilmesi önerilmiştir.

Son olarak, rapor Türkiye için kapsamlı bir **YZ yönetişim modeli** önermektedir. Ulusal Yapay Zekâ Komitesi'nin kurulması, YZ Politika ve Güvenlik Araştırma Merkezlerinin oluşturulması, kamu kurumlarında YZ sorumlularının atanması, açık veri ve şeffaflık mekanizmalarının geliştirilmesi gibi öneriler; Türkiye'nin YZ teknolojilerini etik, güvenli ve toplumsal faydaya uygun biçimde yönetecek kurumsal kapasiteyi oluşturmasına katkı sağlayacaktır.

Raporda yer verilen teknik, etik, hukuki ve yönetişim boyutlu değerlendirmeler kapsamlı ve bütüncül bir çerçeveyle ortaya konmakta; yapay zekâ çağında Türkiye'nin nasıl bir standart yaklaşımına ihtiyaç duyduğu özetlenmektedir.

YZ'nin geleceğini belirleyecek temel unsur, sadece teknolojinin gelişimi değil, bu teknolojinin hangi ilkelerle, hangi standartlarla ve hangi toplumsal değerler üzerine inşa edileceğidir. Bu bağlamda rapor Türkiye'ye stratejik bir rehber sunmaktadır.


Standartlar
Süreçler
İşbirliği

1. STANDARTLAR, SÜREÇLER ve İŞBİRLİĞİ

Ülkü Bayer[a], Dr. Merve Ayyüce Kızrak[b]

[a] *TBD Ankara Şubesi, Sistekno Bilişim Ltd., ubayer@koplan.com.tr*
[b] *Kurucu & CEO, HUX AI, ayyucekizrak@gmail.com*

Bir YZ sisteminin nasıl kullanıldığı, hangi aktörler tarafından işletildiği, diğer sistemlerle nasıl etkileşimde bulunduğu ve hangi toplumsal bağlamda uygulandığı; teknik, etik, sosyal ve operasyonel boyutlarda ortaya çıkan risk etkenleridir. Bu nedenle, YZ risklerini ve potansiyel etkilerini – olumlu ya da olumsuz – belirlemek ve yönetmek, yaşam döngüsünün tüm aşamalarında farklı bakış açıları ve aktörlerin katkısını gerektirmektedir. Çeşitli deneyimlere, uzmanlıklara sahip; disiplinler açısından çeşitlilik barındıran ekiplerin sürece dahil edilmesi önerilmektedir.[7]

YZ projelerinde gerekli olan disiplinler arası işbirliği, farklı uzmanlık alanlarındaki bilgi ve beceri eksiklikleri nedeniyle sorumlulukların anlaşılmasını ve etkili iletişim kurulmasını güçleştirmektedir. Teknik terim ve tanımlardaki tutarsızlıkların yol açtığı yanlış anlamalarla birlikte, ekiplerin uzmanlık alanlarına göre gruplara ayrılması ve gruplar arası iletişimin zayıf olması da entegrasyon sorunlarını derinleştirmektedir.[8]

ISO/IEC 5338 standardı ise bu çok aktörlü ve disiplinler arası işbirliği ihtiyacına yanıt olarak, YZ yaşam döngüsü boyunca kullanılabilecek ortak bir dil ve yöntemler sunmaktadır. Standart, farklı uzmanlık alanlarından gelen aktörlerin aynı çerçevede çalışmasını kolaylaştırmak için süreçleri, kavramları, YZ'ye özgü özellikleri ve durumları tanımlamaktadır. Böyle bir yapı, disiplinler arası katılımı kurumsallaştırmaktadır.

Bu çalışmada, **Sorumlu Yapay Zekâ Araç Takımı**'nda **(RAI Toolkit)** sunulan *persona listesi* temel alınmış ve burada tanımlanan uzmanlık alanları referans noktası olarak değerlendirilmiştir. Persona listesi, bir projede veya çerçevede yer alabilecek farklı rollerin (aktörlerin) tanımlandığı sistematik bir listedir. Sorumlu YZ Araç Takımı, ABD Savunma

---

[7] National Institute of Standards and Technology. (2023) Artificial Intelligence Risk Management Framework (AI RMF 1.0)(NIST AI 100-1). Retrieved from https://nvlpubs.nist.gov/nistpubs/ai/NIST.AI.100-1.pdf

[8] Nahar, N., Zhou, S., Zhang, H., Kästner, C., & Lewis, G. (2023). *A Meta-Summary of Challenges in Building Products with ML Components: Collecting Experiences from 4758+ Practitioners. Project Report. Carnegie Mellon University. Retrieved from osf.io/y5edu

Bakanlığı bünyesindeki “Dijital ve YZ Ofisi” tarafından geliştirilmiş bir araç takımıdır. Araç kılavuzu, YZ Etik İlkeleri ile uyumlu olacak şekilde YZ projelerinin süreç boyunca sorumlu ve etik bir biçimde yürütülmesini sağlamak üzere rehberlik, değerlendirme yöntemleri, kontrol listeleri ve modüler bileşenler sunmaktadır.

Söz konusu listede yer alan farklı rollerin YZ sistemlerinin yaşam döngüsünde üstlendikleri sorumluluklar dikkate alınarak, bu aktörlerin nasıl etkili biçimde işbirliği yapabileceği incelenmektedir. Bu kapsamda, **ISO/IEC 5338 standardının süreç odaklı yapısı** kullanılarak, disiplinler arası katılımın hangi aşamalarda kritik olduğu, ortak dil ve yöntemlerin nasıl geliştirilebileceği ve işbirliği mekanizmalarının nasıl kurumsallaştırılabileceği değerlendirilmektedir.

1.1. ISO/IEC 5338: 2023 YZ Yaşam Döngüsü Standardı İncelemesi

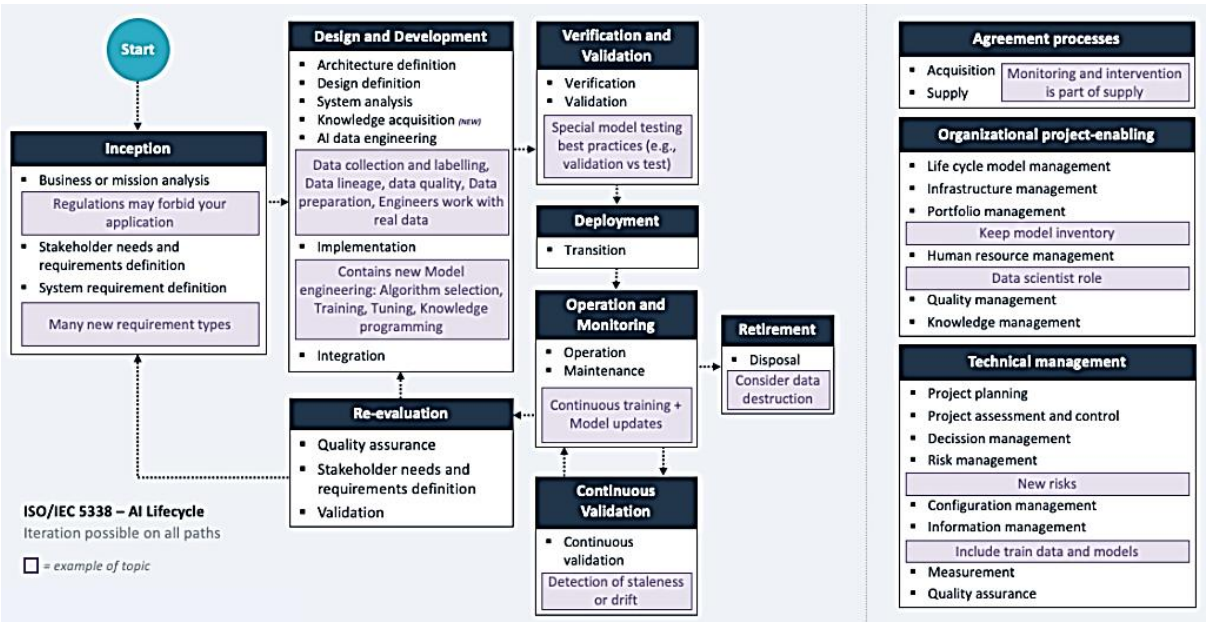
ISO/IEC 5338, YZ sistemlerinin yaşam döngüsündeki süreçleri tanımlayan ve ISO/IEC/IEEE 15288 yaşam döngüsü süreçlerini temel alan, bu süreçlere YZ’ye özgü gereksinim ve örnekleri ekleyen tamamlayıcı bir uluslararası standarttır. Toplam **33 süreç,** standardın 6. bölümünde dört ana kategori altında toplanmıştır:

- 6.1 Anlaşma Süreçleri (Agreement Processes),
- 6.2 Kurumsal Proje-Olanak Sağlayıcı Süreçler (Organizational Project-Enabling Processes),
- 6.3 Teknik Yönetim Süreçleri (Technical Management Processes)
- 6.4 Teknik Süreçler (Technical Processes).

Bu yapı, YZ sistemlerinin geliştirilmesi ve yönetilmesi sırasında takip edilmesi gereken adımları ve kontrol noktalarını tanımlayarak kuruluşlara standart bir çerçeve sunmaktadır. ISO/IEC 5338 mevcut yazılım yaşam döngüsü standartlarına (ör. ISO/IEC 12207, ISO/IEC 15288) dayanmakla birlikte, makine öğrenmesi ve sezgisel sistemler gibi YZ’ye özgü özellikleri kapsayacak şekilde genişletilmiştir. Bu sayede kuruluşlar, bir YZ projesini başlangıçtan (inception) sonlandırmaya (retirement) kadar olan tüm aşamalarda net tanımlar, sıkı kontrol, etkin yönetim, sorunsuz uygulama ve sürekli iyileştirme olanağı bulmaktadır.

Standardın 6.1–6.4 numaralı bölümlerinde tanımlanan süreç kategorileri, kurumsal uygulama ve risk yönetimi (ör. kalite, güvenlik, şeffaflık, sürdürülebilirlik) bağlamında detaylı ve anlaşılır bir şekilde aşağıda açıklanmaktadır.

Şekil 1 - ISO/IEC 5338 standardında tanımlanan YZ sistemi yaşam döngüsü süreçlerinin genel görünümü



Şekil 1: ISO/IEC 5338 standardında tanımlanan YZ sistem yaşam döngüsü süreçlerinin genel görünümü. Diyagram, bir YZ projesinin başlangıç aşamasından sonlandırma aşamasına kadar yaşam döngüsü evrelerini ve bu evrelerle ilişkili süreçleri göstermektedir. Sağ tarafta Anlaşma, Kurumsal Hazırlık, Teknik Yönetim ve Teknik olmak üzere dört ana süreç kategorisi ve bu kategorilere ait süreç başlıkları listelenmiştir. Sol tarafta ise yaşam döngüsü aşamaları (ör. analiz, tasarım, geliştirme, doğrulama, operasyon) ve ilgili teknik süreçler ardışık bir akış olarak sunulmuştur. Bu yapı, YZ sistemleri geliştirilirken izlenmesi önerilen adımları ve her adımda dikkate alınması gereken noktaları görselleştirerek kuruluşlara rehberlik etmektedir.

1.1.1. Anlaşma Süreçleri (Bölüm 6.1)

Anlaşma Süreçleri (Agreement Processes), bir YZ sisteminin tedarik edilmesi veya sağlanması sırasında alıcı ile tedarikçi arasındaki anlaşmaların yönetilmesine odaklanan süreçlerdir.

Bu kategori altında iki temel süreç tanımlanmaktadır:

- Satın Alma/Edinme Süreci (Acquisition Process)
- Tedarik/Sağlama Süreci (Supply Process).

Her iki süreç de bir kurumun dışarıdan YZ sistemi temin etmesi ya da başka bir kuruma YZ sistemi sunması durumlarında devreye girmekte ve bu ilişkilerin başarılı bir şekilde yürütülmesini amaçlamaktadır.

1. **Satın Alma (Edinme) Süreci (Acquisition Process):**

Bir kuruluşun kendi kullanımına yönelik bir YZ sistemini dış kaynaklardan temin etmesi sürecini kapsamaktadır. Bu süreçte, kurumun ihtiyaçlarının tanımlanması, gereksinimlerin tespit edilmesi, uygun tedarikçi veya çözümün seçilmesi ve sözleşme yönetimi gibi adımlar bulunmaktadır.

Örneğin, bir kamu kurumu YZ tabanlı bir yazılım satın alırken bu süreci izleyerek ihale dokümanlarını hazırlar, teknik ve etik gereksinimleri (ör. veri gizliliği, güvenlik kriterleri, algoritmik tarafsızlık) netleştirir ve tedarikçi ile bu koşulları içeren bir sözleşme yapar.

Amaç, satın alınan YZ sisteminin beklenen işlevselliği, kalitesi ve uyumluluğu sağlamasını güvence altına almaktır. Bu sayede kurum, ileride doğabilecek riskleri (ör. ürünün söz verileni yapmaması, güvenlik açıkları veya yasal uyumsuzluklar) en baştan azaltmış olur.

2. **Tedarik (Sağlama) Süreci (Supply Process):**

Bir tedarikçi kuruluşun, bir müşteri için YZ sistemini geliştirmesi, sağlaması ve teslim etmesi sürecidir. Tedarikçi perspektifinden bu süreç, müşteri gereksinimlerini anlama, teklif hazırlama, proje planlama, YZ sistemini geliştirme veya uyarlama, test etme ve müşteriye teslim etme gibi adımları içermektedir. Tedarikçi ayrıca, teslim edilen sistemin anlaşılan gereksinimleri karşıladığını ve kalite standartlarına uygun olduğunu garanti etmelidir. Bu süreç boyunca tedarikçi, sözleşmesel yükümlülüklerini yerine getirirken aynı zamanda ortaya çıkabilecek değişiklik taleplerini yönetir ve müşteriye düzenli bilgi akışı sağlamaktadır. Böylece hem müşteri hem tedarikçi açısından beklentiler yönetilmekte, anlaşmazlıklar en aza indirilmektedir.

Anlaşma süreçlerinin mantığı, YZ sistemi alım-satım ilişkilerini düzenleyerek her iki tarafın da hak ve sorumluluklarını netleştirmektir. Kurumsal açıdan bu süreçler son derece önemlidir; çünkü, hukuki uyum, kalite güvencesi ve risk paylaşımı bu aşamada belirlenmektedir.

Örneğin, bir YZ sisteminin geliştirilmesi sırasında ortaya çıkabilecek gizlilik ihlali riskleri veya fikri mülkiyet sorunları, anlaşma metinlerinde önceden ele alınmalıdır. İyi tanımlanmış bir satın alma/tedarik süreci, YZ sisteminin istenen performans kriterlerini (ör. doğruluk oranı, güvenlik sertifikaları) ve etik/çevresel standartları (ör. veri koruma, enerji verimliliği) karşılamasını sözleşme yoluyla güvence altına alır. Bu sayede kurumlar, YZ teknolojisini şeffaflık ve hesap verebilirlik ilkelerine uygun biçimde edinerek olası itibar kayıpları, hukuki yaptırımlar veya başarısız yatırımlardan kaçınabilir.

Kısacası, 6.1 kapsamındaki anlaşma süreçleri, YZ sistemi yaşam döngüsünün daha başlamadan güvence altına alınması ve tüm taraflar için sağlam bir temel oluşturulması amacına hizmet etmektedir.

1.1.2. Kurumsal Proje-Olanak Sağlayıcı Süreçler (Bölüm 6.2)

Kurumsal Proje-Olanak Sağlayıcı Süreçler (Organizational Project-Enabling Processes), bir kurumun YZ projelerini başarıyla gerçekleştirebilmesi için gerekli kurumsal altyapıyı, kaynakları ve yönergeleri oluşturup sürdürmeye odaklanan süreçler bütünüdür.

Bu süreçler, tek tek proje seviyesinden ziyade organizasyon seviyesinde faaliyet göstermekte ve YZ projelerinin etkin bir şekilde yürütülmesi için uygun ortamın sağlanmasını hedeflemektedir.

ISO/IEC 5338, bu kategori altında altı temel süreç tanımlamıştır:

1. **Yaşam Döngüsü Modeli Yönetimi Süreci (Life Cycle Model Management Process)**

   Kurumun, projelerinde kullanacağı yaşam döngüsü modellerini (ör. geleneksel yazılım geliştirme modeli, çevik metodolojiler veya CRISP-DM gibi veri bilimi süreç modelleri) tanımlayıp yönetmesini içerir. YZ projeleri için, klasik yazılım süreçlerine ek olarak veri hazırlama ve model eğitimi gibi adımların dahil edildiği özel yaşam döngüsü yaklaşımları gerekebilir. Bu süreç, organizasyonun tüm projelerde ortak bir metodoloji kullanmasını ve böylece tutarlılık ve kaliteyi sağlamasını hedeflemektedir.

Örneğin, bir kurum YZ projelerinde daima bir keşif/aşamalandırma (proof-of-concept) süreci öngörüyorsa, bu model kurum tarafından resmi olarak tanımlanır ve proje ekiplerine rehberlik eder.

2. Altyapı Yönetimi Süreci (Infrastructure Management Process)

YZ sistemlerinin geliştirilmesi, test edilmesi ve işletilmesi için gerekli teknik altyapının sağlanması ve sürdürülmesi sürecidir. Donanım kaynakları (sunucular, GPU çiftlikleri), bulut hizmetleri, veri depolama üniteleri, yazılım araçları (ör. makine öğrenmesi kütüphaneleri, MLOps araçları) ve bilişim güvenliği altyapıları bu sürecin kapsamına girmektedir.

Kurum, YZ projelerinin ihtiyaç duyacağı yüksek işlem gücü, büyük veri kümeleri için depolama, yedekleme ve siber güvenlik önlemleri gibi unsurları önceden planlayıp temin etmelidir. Altyapı yönetimi, sürdürülebilirlik açısından da kritik önemde olmaktadır.

Örneğin; enerji verimli veri merkezleri kullanmak veya hesaplama kaynaklarını ölçeklendirmek, YZ projelerinin çevresel ayak izini azaltırken maliyetleri de kontrol altında tutar.

3. Portföy Yönetimi Süreci (Portfolio Management Process)

Kurumun yürüttüğü veya planladığı tüm YZ projelerinin stratejik düzeyde yönetimini ifade eder.

Bu süreç, hangi YZ projelerinin önceliklendirileceği, kaynakların (bütçe, personel) nasıl dağıtılacağı ve projelerin kurumun genel hedeflerine hizalanıp hizalanmadığı konularını kapsamaktadır. Portföy yönetimi sayesinde, bir kuruluş birbirinden bağımsız pek çok YZ girişimini koordine edebilmekte, tekrarlayan çalışmalardan kaçınmakta ve yatırımlarının beklenen getirisini en üst düzeye çıkarabilmektedir.

Örneğin, bir STK / NGO düşünelim: Hem tarım verimliliğini artırmak için bir makine öğrenmesi projesi yürütüyor, hem de finansal süreçlerini iyileştirmek için bir YZ aracı almayı planlıyor olsun. Portföy yönetimi, bu iki girişimin de örgütün misyonuna uygun olup olmadığını değerlendirir, risk ve fayda analizlerini yapar ve sınırlı kaynakları en doğru şekilde tahsis eder.

4. İnsan Kaynakları Yönetimi Süreci (Human Resource Management Process)

YZ sistemleriyle çalışacak personelin yeterliliklerinin sağlanması, geliştirilmesi ve idamesi bu sürecin konusudur. Kurum, veri bilimciler, makine öğrenmesi mühendisleri, etik uzmanları, siber güvenlik analistleri gibi çeşitli uzmanlıklara ihtiyaç duyabilmektedir. İnsan kaynakları yönetimi süreci, doğru yetenekleri istihdam etmeyi, mevcut çalışanlara YZ alanında eğitimler vermeyi (ör. etik YZ eğitimi, yeni bir ML çerçevesinin kullanımı) ve ekip içinde sağlıklı bir görev dağılımı kurmayı içermektedir. YZ projelerinin başarılı olabilmesi için farklı beceri setlerine ihtiyaç duyulduğu standardın vurguladığı noktalardan biridir. Ayrıca bu süreç

kapsamında, ekip üyelerinin motivasyonunun korunması, görev tanımlarının netleştirilmesi ve gerektiğinde dış uzman desteğinin (danışmanlık, akademik işbirlikleri) alınması da planlanmaktadır. Böylece kurum, YZ projelerini yürütebilecek yeterli ve yetkin insan kaynağına sahip olmaktadır.

5. Kalite Yönetimi Süreci (Quality Management Process):

Kurum genelinde YZ projelerinin ve ürünlerinin kalite standartlarını tanımlayan ve bu standartlara uyumu sağlayan süreçtir.

Kalite yönetimi, hem süreç kalitesini (yani proje geliştirme süreçlerinin kurallara uygunluğunu) hem de ürün kalitesini (ortaya çıkan YZ yazılımının güvenilirliği, doğruluğu, emniyeti vb.) kapsar. Bu süreç çerçevesinde, kurumun YZ projeleri için geçerli olacak kalite politikaları belirlenmektedir.

Örneğin, her YZ modelinin geçmesi gereken en az doğruluk/eğitim başarımı eşiği, kodlama standartları, dokümantasyon gereklilikleri, güvenlik testleri veya etik değerlendirme kriterleri tanımlanabilir. Kalite yönetimi, bu standartların tüm projelerde uygulandığını denetler ve iyileştirmeler için geri bildirim mekanizmaları kurar. Kurumsal düzeyde güçlü bir kalite yönetimi, YZ sistemlerinin güvenilir ve güvenli olmasını, aynı zamanda şeffaflık ve hesap verebilirlik kriterlerini karşılamasını güvence altına alır.

6. Bilgi (Know-How) Yönetimi Süreci (Knowledge Management Process)

YZ projelerinden elde edilen bilgi ve deneyimin kurumsal bellekte tutulması, paylaşılması ve gelecek projelerde yeniden kullanılabilir hale getirilmesi sürecidir.

YZ sistemleri genellikle karmaşık ve uzmanlık gerektiren projelerdir; bu projelerden çıkan dersler, en iyi uygulamalar, kullanılan veri kümeleri ve model parametreleri gibi bilgiler kurum için değerlidir. Bilgi yönetimi süreci, bu tür bilgilerin belgelenmesini (ör. proje sonunda bir "öğrenilen dersler" raporu hazırlanması), merkezi bir depoda saklanmasını (iç wiki, bilgi tabanı vb.) ve ilgili çalışanlar arasında yayılmasını sağlamaktadır.

Örneğin, bir projede belirli bir makine öğrenmesi algoritmasının başarısız olma nedenleri veya başarılı bir veri ön işleme tekniği, başka bir ekip için yol gösterici olabilir. Bilgi yönetimi ayrıca, süreklilik ve sürdürülebilirlik açısından da kritik rol oynar: kilit bir YZ uzmanının kurumdan ayrılması durumunda, onun bilgisi dokümante edilmişse projede süreklilik sağlanabilir. Bu süreç, kurum içindeki kurumsal belleği güçlendirerek YZ

projelerinde tekrarlanan hataları önlemeye ve yenilikçi çözümlerin yeniden kullanılmasına olanak tanır.

Kurumsal proje-olanak sağlayıcı süreçlerin genel amacı, organizasyonun YZ projelerini başarıyla yaşama geçirebilmesi için gereken tüm kurumsal düzenlemeleri yapmasıdır. Bu süreçler sayesinde kurum, YZ projelerine başlamadan önce gerekli altyapıyı kurar, insan kaynağını hazırlar, kalite ve etik standartlarını belirler ve stratejik bir yol haritası çizer.

Sonuç olarak, proje ekipleri işe koyulduklarında belirsizlikler azalır, çünkü net bir çerçeve ve destekleyici bir ortam zaten oluşturulmuştur. Risk yönetimi açısından bakıldığında, bu süreçler önetkin (proaktif) bir yaklaşımla riskleri kaynağında azaltmaktadır.

Örneğin, güçlü bir bilgi yönetimi olmazsa kritik bir modelle ilgili bilgi kaybedilebilir (risk); ancak bu süreç bunu engeller. Benzer şekilde, eğer kalite yönetimi kurumsal düzeyde tanımlanmamışsa, projeler arasında tutarsızlıklar olabilir ve güvenlik veya doğruluk açıkları ortaya çıkabilir; bu süreç sayesinde bu riskler en aza indirilir. Şeffaflık da bu kategoriyle ilgilidir: Kurum genelinde standart süreçler ve politikalar olması, paydaşlara karşı hesap verebilirlik sağlar ve YZ projelerinin toplum gözünde uygunluğunu artırır.

Kısacası, 6.2'deki süreçler, bir kurumun YZ girişimlerinde kurumsal kapasitesini ve yönetişimini güçlendiren, böylece projelerin başarı şansını ve güvenilirliğini yükselten kritik unsurlardır.

1.1.3. Teknik Yönetim Süreçleri (Bölüm 6.3)

Teknik Yönetim Süreçleri (Technical Management Processes), bir YZ projesinin yürütülmesi sırasında projenin planlanması, izlenmesi, kontrolü ve desteklenmesi için gereken yönetimsel süreçleri kapsamaktadır.

Bu süreçler, proje yöneticileri ve teknik liderler tarafından uygulanmakta olup projenin zamanında, bütçesinde, hedeflenen kalite ve kapsamda ilerlemesini sağlamayı amaçlamaktadır. ISO/IEC 5338 standardı bu kategori altında sekiz temel süreç tanımlamıştır ve bunların her biri, özellikle YZ projelerinin kendine özgü belirsizlik ve risklerini yönetmeye yardımcı olmaktadır.

1. **Proje Planlama Süreci (Project Planning Process)**

YZ projesinin başında, yapılacak işlerin planlanması bu sürecin konusudur. Kapsamın tanımlanması, takvim ve kilometre taşlarının oluşturulması, kaynak planlaması (ekip üyeleri,

veri, donanım, bütçe) ve proje yönetim prosedürlerinin belirlenmesi bu aşamada gerçekleştirilmektedir. YZ projelerinde planlama, klasik yazılım projelerine göre daha fazla belirsizlik barındırabilir; örneğin bir makine öğrenmesi modelinin istenen performansa ne zaman ulaşacağı önceden kestirilemeyebilir. Bu nedenle proje planlama süreci, iteratif yaklaşımları (ör. ardışık prototipleme, sprintler) ve belirsizlik yönetimini içerecek şekilde adapte edilir. İyi bir proje planı, riskleri erken dönemde öngörerek bunlara yönelik aksiyonları da içerir (ör. "veri bulunamazsa ne yapılacak?" gibi senaryolar). Böylece planlama süreci, projenin yol haritasını çizerken belirsizlikleri asgariye indirir ve tüm paydaşların beklentilerini yönetir.

2. **Proje Değerlendirme ve Kontrol Süreci (Project Assessment and Control Process)**

Proje ilerledikçe, planlanan hedeflere göre düzenli değerlendirme yapılması ve gerektiğinde düzeltici/önleyici tedbirlerin alınması bu sürecin odak noktasıdır.

Proje yöneticileri belirli aralıklarla projenin durumunu inceler: Takvimde gecikme var mı? Bütçe aşılıyor mu? Ara kilometre taşları (ör. bir model prototipinin tamamlanması) gerçekleşmiş mi? Bu değerlendirmeler sonucunda eğer sapmalar varsa kontrol mekanizmaları devreye girer.

Örneğin, ek kaynak atamak, hedefleri revize etmek veya kapsamı daraltmak gibi kararlar alınabilir. YZ projelerinde özellikle, başlangıçta öngörülmeyen teknik zorluklar (ör. model eğitiminin beklenenden uzun sürmesi, performansın yeterli olmaması) çıkabilir. Bu süreç, erken uyarı sistemi gibi çalışarak sorunlar büyümeden müdahale edilmesini sağlar. Düzenli durum raporları, paydaş iletişimleri ve gerektiğinde kriz planlarının uygulanması bu sürecin parçalarıdır.

Sonuç olarak, proje değerlendirme ve kontrol süreci, projenin rotada kalmasına destek olurken kalite ve hedeflerden taviz verilmemesini güvence altına alır.

3. **Karar Yönetimi Süreci (Decision Management Process)**

Proje sırasında alınması gereken önemli teknik veya yönetsel kararların sistematik bir yöntemle ele alınmasını sağlamaktadır. YZ projelerinde, hangi makine öğrenmesi algoritmasının kullanılacağından, modelin mimarisine veya hangi veri kaynaklarının tercih edileceğine kadar pek çok kritik karar alınmaktadır. Karar yönetimi süreci, bu kararlar için

kriterlerin belirlenmesini, alternatiflerin değerlendirilmesini ve nesnel verilerle en iyi seçeneğin seçilmesini içermektedir.

Örneğin, bir karar noktası "Model A'yı mı kullanmalı yoksa Model B'yi mi?" ise, kriterler doğruluk, hesaplama maliyeti, açıklanabilirlik ve etik uygunluk olabilir; her iki model bu kriterlere göre değerlendirilip dokümante edilir. Bu süreç ayrıca, alınan kararların kaydını tutarak ileride hesap verebilirlik sağlar (hangi karar neden alındı, kim onayladı gibi). Kurumlar açısından bu önemlidir, çünkü özellikle YZ etiği veya güvenliğiyle ilgili kararlar (örneğin, daha yüksek doğruluk sağlayan ama daha az açıklanabilir bir model seçimi) daha sonra sorgulanabilir; karar yönetimi süreci sayesinde şeffaflık ve izlenebilirlik sağlanır.

4. Risk Yönetimi Süreci (Risk Management Process)

Bir YZ projesine özgü olası risklerin tanımlanması, analiz edilmesi, izlenmesi ve azaltılması faaliyetlerini kapsamaktadır. Risk yönetimi, proje başlangıcından itibaren sürekli yürütülen bir süreçtir.

YZ projelerinde risk kavramı çok boyutludur: **teknik riskler** (ör. model istenen başarıya ulaşamaz, veri kalitesi düşüktür), **proje riskleri** (bütçe/takvim aşımları), **etik riskler** (algoritmik önyargı; bias, ayrımcılık yapma ihtimali), **güvenlik riskleri** (modelin adversarial saldırılara açık olması, veri sızıntısı) ve hatta **düzenleyici riskler** (yasal mevzuata uyumsuzluk). Bu süreçte ekip, olası riskleri listeler, her bir riskin olasılık ve etkisini değerlendirir ve en kritik riskler için önleyici stratejiler geliştirir.

Örneğin, "eğitim verisinde istenmeyen önyargı tespit edilirse" bir risk ise, buna karşı proaktif olarak veri denetimleri ve iyileştirme yöntemleri planlanır. ISO/IEC 5338, YZ projelerinde ortaya çıkan yeni risk konularına özel dikkat çekmektedir. Örneğin şeffaflık eksikliği, istenmeyen önyargılar, amaç dışı kullanım gibi risk başlıkları geleneksel yazılımlara kıyasla YZ'de daha ön plandadır. Risk yönetimi süreci, bu gibi konuları da risk evrenine dahil ederek kurumun kalite, güvenlik ve etik değerlerden ödün vermemesini sağlar. Etkili bir risk yönetimi, ileride ortaya çıkabilecek skandalların, yasal sorunların veya teknik başarısızlıkların önüne geçerek YZ projesinin güvenilir bir şekilde ilerlemesine katkıda bulunur.

5. Yapılandırma Yönetimi Süreci (Configuration Management Process)

Proje çıktılarının (yazılım kodu, veri kümeleri, model sürümleri, dokümantasyon gibi) konfigürasyonunun tanımlanması ve kontrol altında tutulması sürecidir. YZ projelerinde konfigürasyon yönetimi özellikle önemlidir; zira sadece kaynak kodu değil, kullanılan veri

kümelerinin ve eğitilen modellerin sürümleri de ürünün bir parçasıdır. Bu süreç, hangi sürümde hangi verinin kullanıldığı, model parametrelerinin ne olduğu, değişikliklerin kim tarafından ne zaman yapıldığı gibi bilgileri kayıt altına alır.

Örneğin, bir makine öğrenmesi modelini v1.0'dan v1.1'e güncellediğinizde, v1.1'de hangi eğitim verisi kullanıldı, hangi üst parametreler ayarlandı, değişiklik ne getirdi. Tüm bunlar sürüm kontrol sistemlerinde izlenmelidir. Yapılandırma yönetimi sayesinde ekip, geriye dönüp baktığında herhangi bir sonucun tekrar üretilebilirliğini (reproducibility) sağlar; bu, bilimsel geçerlik ve güvenilirlik açısından kritiktir.

Ayrıca bir sorun çıktığında, hangi sürümün hataya sebep olduğunu bulmak ve gerektiğinde eski sürüme dönmek mümkün olur. Bu süreç, YZ sistemlerinde şeffaflık ve izlenebilirlik için de bir altyapı sunar, çünkü bir modele dair tam bir değişiklik tarihi (tarihçe) tutulmaktadır.

6. Bilgi Yönetimi Süreci (Information Management Process)

Proje boyunca oluşturulan ve kullanılan tüm bilgi varlıklarının (dokümanlar, veri tanımları, raporlar, modellerin teknik detayları) düzenli bir şekilde yönetilmesi sürecidir.

Bilgi yönetimi, bir anlamda 6.2'deki kurumsal bilgi yönetiminin proje seviyesine uygulanmasıdır. Her YZ projesinin başlangıcında bir dokümantasyon planı oluşturulur: gereksinim dokümanları, tasarım özellikleri, veri sözlükleri, test raporları, kullanıcı kılavuzları vb. hangi belgelerin üretileceği tanımlanır. Bu süreç, bu belgelerin güncel tutulmasını, doğru paydaşlarla paylaşılmasını ve güvenli biçimde saklanmasını sağlar.

Örneğin, hassas bir sağlık verisiyle eğitilen bir modelin geliştirildiği projede, veri yönetimi prosedürleri, onay formları, etik kurul izin belgeleri gibi önemli dokümanlar oluşur; bilgi yönetimi süreci bunların tamamının kayıt altında olmasını ve erişilebilir olmasını temin eder. Ayrıca proje sonunda bilgi birikiminin kurumsal hafızaya aktarılmasına da katkı sunar. Etkili bir bilgi yönetimi, şeffaflık ilkesini destekler; zira projenin her adımı belgelenmiş durumdadır ve dış denetim veya değerlendirme durumlarında kuruma güvenilirlik kazandırır.

7. Ölçüm Süreci (Measurement Process)

Proje performansını ve ürün kalitesini nicel metriklerle izlemeyi ve değerlendirmeyi kapsamaktadır. Bu süreçte, projenin başarısını ölçmek için uygun göstergeler tanımlanır; hem yönetimsel (ör. proje zaman planına uyum, harcanan efor vs.) hem de teknik (ör. modelin

doğruluk oranı, hata oranı, sistem yanıt süresi) ölçütler belirlenir. YZ projelerinde ölçüm, kalite güvencesi ve sürekli iyileştirme için kritik bir araçtır.

Örneğin, bir derin öğrenme modelinin performansını izlemek için doğruluk (accuracy), hatırlama (recall), kesinlik (precision) gibi metrikler proje boyunca takip edilir. Eğer bu metriklerde istenmeyen sapmalar varsa (beklenen kaliteyi karşılamıyorsa), ekip bunun nedenlerini araştırır ve iyileştirme yapar. Ölçüm süreci aynı zamanda proje yönetiminde de kullanılır: "%80'i tamamlandı" gibi öznel beyanlar yerine, ölçülebilir verilerle (tamamlanan iş öğelerinin yüzdesi, kalan hataların sayısı vb.) projenin durumu ifade edilir. Bu, paydaşlara karşı hesap verebilirliği artırır ve karar almayı kolaylaştırır. Ayrıca, proje sonunda toplanan metrikler gelecekteki projeler için referans teşkil ederek kurumsal öğrenmeye katkı sağlar.

8. Kalite Güvence Süreci (Quality Assurance Process)

Proje ve ürün kalitesinin, belirlenen standartlara uygun olmasını bağımsız değerlendirmelerle teyit eden süreçtir. Kalite güvence ekibi veya sorumluları, proje ekibinden ayrı bir bakış açısıyla süreçlerin düzgün işlediğini ve projenin kalite kriterlerini karşıladığını denetler.

Bu denetim; süreç denetimleri (ör. gereksinim belirleme adımları uygun yapılmış mı, kod gözden geçirmeleri gerçekleştiriliyor mu?) ve ürün denetimleri (yazılım testleri yeterli mi, model doğruluğu kabul kriterlerini karşılıyor mu, güvenlik taramaları yapıldı mı?) şeklinde olabilir. YZ projelerinde kalite güvence, özellikle belgelendirme, test ve doğrulama alanlarında eksiklikleri tespit edip giderilmesini sağlayarak projenin sağlıklı ilerlemesine katkı sunmaktadır.

Örneğin, bir kalite güvence kontrolünde, projenin başlangıcında tanımlanan "model kararlarının açıklanabilirlik dokümanı oluşturulacak" maddesinin atlandığı fark edilirse, QA bunu raporlar ve ekip gerekli düzeltmeyi yapar. Bu süreç, kurum içi denetim mekanizması gibi çalışarak, hataların ve ihmallerin erken aşamada yakalanmasına olanak tanır. Kalite güvence sayesinde, proje sonunda teslim edilen YZ sisteminin sadece çalışması değil, aynı zamanda kurumsal ve etik beklentilere uygun olması sağlanır.

Genel olarak Teknik Yönetim Süreçleri, bir YZ projesinin başarılı bir şekilde yönetilmesi için gereken tüm araçları kurumun kullanımına sunar. Bu süreçler, proje ekibini uygun rotada tutarken ortaya çıkabilecek teknik ve yönetimsel riskleri sürekli olarak gözlem altında tutar. Özellikle YZ projelerinde vurgulanan risk yönetimi ve kalite güvencesi, projenin

hem yasal/düzenleyici gerekliliklerle uyumunu hem de güvenilir bir sonucun elde edilmesini güvence altına alır. Teknik yönetim süreçleri sayesinde, kurumlar YZ projelerinde yönetişim (governance) ilkesini hayata geçirmiş olur. Yani proje boyunca hesap verebilirlik, şeffaflık ve kontrol mekanizmaları tesis edilmiş olur. Bu da uzun vadede, YZ uygulamalarının başarısını artırırken beklenmedik maliyetler, skandallar veya yasalarla çatışma gibi olumsuz sonuçların önüne geçilmesine yardımcı olur.

1.1.4. Teknik Süreçler (Bölüm 6.4)

Teknik Süreçler (Technical Processes), YZ sisteminin fikir aşamasından operasyonuna ve en sonunda devre dışı bırakılmasına (sonlandırılmasına) kadar geçen yaşam döngüsü boyunca yapılan tüm teknik faaliyetleri tanımlar. Bu süreçler doğrudan ürünün (YZ sisteminin) kendisiyle ilgilidir ve bir bakıma projenin somut çıktılarını üretirler.

ISO/IEC 5338'e göre teknik süreçler, klasik sistem ve yazılım mühendisliği faaliyetlerini YZ'ye özgü eklemelerle birlikte kapsayacak şekilde geniş bir yelpazeye yayılır. Standardın 6.4 bölümünde 17 adet teknik süreç tanımlanmıştır. Aşağıda bu süreçlerin her biri, yaşam döngüsündeki sıralarına yakın bir düzende, kurumsal uygulama ve risk yönetimi bakış açısıyla açıklanmaktadır:

1. **İş veya Misyon Analizi Süreci (Business or Mission Analysis Process)**

Bu en baştaki teknik süreç, kurumun veya projenin **i**ş amacını, misyon hedeflerini ve problem tanımını netleştirmeyi hedefler.

Bir YZ sistemine başlanmadan önce, "Bu sistem ne amaçla gereklidir? Hangi iş ihtiyacını veya toplumsal ihtiyacı karşılayacak?" sorularına yanıt aranır.

Örneğin, bir sağlık STK'sı (sivil toplum kuruluşu) hastalık teşhisinde yardımcı olacak bir YZ aracı geliştirmeyi düşünüyorsa, iş analizi sürecinde mevcut teşhis sürecinin zayıf yönleri, YZ'nin nasıl bir değer katacağı, hedeflenen iyileşme oranları veya hız kazanımı gibi konular ele alınır. Bu süreçte paydaşların iş beklentileri de genel hatlarıyla anlaşılır ancak ayrıntıya girilmez; daha çok yüksek seviye hedefler ve kısıtlar belirlenir. Aynı zamanda, uyulması gereken yasal çerçeve veya etik ilkeler varsa (ör. sağlık verilerinin gizliliği gibi), bunlar da bu aşamada not edilir.

İş/misyon analizi, tüm projenin yönünü belirleyen kritik bir adımdır; zira bu analizden çıkan sonuç, sonraki tüm gereksinim ve tasarım çalışmalarının referans noktası olacaktır. Kurumsal açıdan, iş analizi süreci sayesinde YZ projesinin gerçekten stratejik bir amaca hizmet ettiğinden ve elde edilecek çıktının kuruma/değer zincirine anlamlı bir katkı sağlayacağından

emin olunur. Ayrıca bu süreç, daha ilk adımdan itibaren kalite (doğru problemi çözme) ve sürdürülebilirlik (iş hedeflerine uyum) adına önemli bir kontrol noktasıdır.

2. **Paydaş İhtiyaçları ve Gereksinimleri Tanımlama Süreci (Stakeholder Needs and Requirements Definition Process)**

Bu süreçte, YZ sisteminden etkilenecek veya onu kullanacak tüm **p**aydaşların (stakeholder) beklenti, ihtiyaç ve kısıtları toplanmakta ve tanımlanmaktadır. Paydaşlar; son kullanıcılar, sistemi yöneten operatörler, sistemi satın alan müşteri, düzenleyici otoriteler veya toplumun geneli (özellikle kararları YZ'den etkilenecekse) olabilmektedir.

Örneğin, bir yüz tanıma sistemi geliştiriliyorsa, paydaş gereksinimleri arasında kullanıcıların hızlı ve doğru tanıma isteği, veri koruma otoritelerinin mahremiyet talepleri, kurumun BT departmanının sistemin mevcut altyapıya entegrasyon koşulları sayılabilir. Bu süreçte ihtiyaçlar olabildiğince kullanıcı dostu ve anlaşılır ifadelerle tanımlanır; "sistem adil olmalı" gibi geniş ifadeler dahi not edilir, sonrasında netleştirilip ölçülebilir gereksinimlere dönüştürülmek üzere. Teknik olmayan gereksinimler (ör. güvenlik, gizlilik, açıklanabilirlik, kullanım kolaylığı) bu aşamada belirgin hale getirilir, böylece YZ sisteminin başarısı sadece teknik performansla değil, aynı zamanda etik ve kullanıcı kabulü kriterleriyle ilişkilendirilir.

Bu sürecin kurumlar için önemi büyüktür: Tüm paydaşların sesinin baştan duyulmasını sağlayarak, ileride oluşabilecek uyumsuzluk risklerini azaltır. Özellikle YZ projelerinde, sonradan ortaya çıkan "Bu sistem bizim ihtiyacımızı karşılamıyor" gibi durumlar, paydaş gereksinimlerinin yeterince alınmamasından kaynaklanabilir. Ayrıca şeffaflık ilkesi gereği, paydaşlarla iletişim kurulması ve onların beklentilerinin dikkate alınması, projenin sosyal kabulü açısından da kritiktir.

3. **Sistem Gereksinimlerinin Tanımlanması Süreci (System Requirements Definition Process)**

Paydaş ihtiyaçları belirlendikten sonra, bu talepler teknik ve ölçülebilir sistem gereksinimlerine dönüştürülür. Yani "ne isteniyor?" sorusu, mühendislik dilinde "sistem ne yapmalı?" sorusuna dönüştürülür. Bu gereksinimler işlevsel olabilir (sistemin yapacağı spesifik işler, ör. "sistem girdi olarak X verisini alıp Y çıktısını üretebilmeli") veya kaliteyle ilgili olabilir (performans, güvenlik, güvenilirlik, sürdürülebilirlik gibi). YZ sistemlerinde ayrıca veri ile ilgili gereksinimler de tanımlanır.

Örneğin, "eğitim verisi en az %20 farklı kaynaktan gelmeli" veya "model kararları 100 ms altında üretmeli" gibi teknik kriterler bu aşamada netleştirilir. Bu süreçte, her paydaş ihtiyacının izini sürecek en az bir sistem gereksinimi olmasına dikkat edilir (izlenebilirlik matrisi oluşturulabilir). Aynı zamanda, düzenleyici gereksinimler de (ör. YZ sisteminin CE sertifikası alması için gereken koşullar, GDPR uyumu vb.) sistem gereksinimleri setine dahil edilir.

Bu süreç, projenin sonraki tasarım ve doğrulama aşamalarının temelini oluşturur; çünkü tasarımcılar ve geliştiriciler neyi başarmaları gerektiğini bu gereksinim listesinden öğrenir. Kurum açısından bakıldığında, iyi tanımlanmış gereksinimler, projenin başarı kriterlerini somutlaştırdığı için çok değerlidir. Hem proje ekibi odaklanması gereken noktaları bilir, hem de proje sonunda başarının ölçülmesi mümkün hale gelir (gereksinimler karşılandı mı diye bakılır). Eksik veya hatalı gereksinimler, doğrudan projenin yanlış ürün geliştirmesine yol açabileceği için bu süreçte yapılan titiz çalışma, kalite güvencesinin ilk adımı olarak görülebilir.

4. Sistem Mimari Tanımlama Süreci (System Architecture Definition Process):

Bu süreç, YZ sisteminin yüksek seviyeli tasarımının yapılmasını içerir. Sistem mimarisi, gereksinimleri karşılamak için sistemin hangi bileşenlerden oluşacağını ve bu bileşenlerin birbiriyle nasıl etkileşeceğini belirler.

Örneğin, bir YZ destekli karar sistemi geliştiriliyorsa, mimaride şu bileşenler olabilir: kullanıcı arayüzü, veri toplama modülü, makine öğrenmesi modeli, karar destek modülü, sonuç raporlama modülü vb. Mimari tasarım aynı zamanda altyapısal kararları da kapsar: sistem bulutta mı koşacak, uç bir cihazda mı; gerçek zamanlı veri akışı mı var, yoksa toplu işlem mi; hangi veritabanı veya veri gölü kullanılacak gibi.

YZ sistemlerinin mimarisinde dikkat edilmesi gereken özel hususlar da vardır: Örneğin, makine öğrenmesi modelinin güncellenebilir olması için mimaride bir model yönetim bileşeni öngörülebilir ya da şeffaflık için bir açıklanabilirlik modülü dâhil edilebilir. Sistem mimari tanımlama süreci, aynı zamanda güvenlik ve gizlilik tasarım ilkelerini de barındırır ("en baştan güvenlik" prensibi). Kurumlar için bu süreç kritik önem taşır çünkü mimari, sistemin hem işlevselliğini hem de kalitesini (ör. performans, güvenlik, ölçeklenebilirlik) büyük ölçüde belirler. İyi bir mimari, gelecekte değişen ihtiyaçlara uyum sağlamayı (sürdürülebilirlik) kolaylaştırır ve geliştirme sürecini hızlandırır; kötü tasarlanmış bir mimari ise projeyi çıkmaza

sokabilir. Bu nedenle genellikle deneyimli yazılım mimarları ve YZ uzmanları bir arada çalışarak mimariyi oluşturur, ve bu tasarım kilit paydaşlarla gözden geçirilerek doğrulanır.

5. Tasarım Tanımlama Süreci (Design Definition Process)

Mimari seviyede belirlenen bileşenlerin iç yapılarının daha ayrıntılı bir şekilde tasarlanması bu sürecin konusudur. Her bir bileşen veya modül için nasıl çalışacağı, hangi algoritmaları veya yöntemleri kullanacağı, arayüzlerinin ne olacağı tasarım adımında netleştirilir.

Örneğin, mimarideki "makine öğrenmesi modeli" bileşeni için tasarım aşamasında karar verilecekler şunlar olabilir: Hangi algoritma veya mimari (karar ağaçları mı, sinir ağları mı, derin öğrenme mi), girdi özellikleri nelerdir, çıktı nasıl temsil edilecek, modelin üst (hiper) parametre ayarları neler olmalı vb.

Yine benzer şekilde, "veri toplama modülü" için tasarım, hangi API'ların kullanılacağı, veri temizleme kuralları, hata durumunda sistemin davranışı gibi detayları belirler. Tasarım tanımlama sürecinde modelleme dilleri veya araçları kullanılabilir (UML diagramları, akış şemaları, Python pseudo-code, vb.) ve her tasarım kararı, ilgili gereksinimlere izlenebilir şekilde bağlanır (örneğin, yüksek güvenlik gereksinimi olan bir bileşen için tasarımda ek şifreleme adımları eklenir).

YZ sistemlerinde tasarım aşamasında ayrıca deneysel prototipleme de olabilir; özellikle birden fazla algoritma denenecekse, tasarım süreci bu deneyleri planlar. Kurum açısından tasarım süreci, soyut gereksinimlerin somut bir plana dönüştüğü kritik geçiş noktasıdır. İyi yapılmış bir tasarım, geliştirme ekibinin neyi nasıl uygulayacağını netleştirir ve kalite ile verimliliği artırır. Ayrıca tasarım dokümanları, ileride bakım ve güncelleme yapacak ekipler için yol gösterici olacağından, sistemin sürdürülebilirliği (maintainability) ve bakımı için de önemlidir.

6. Sistem Analizi Süreci (System Analysis Process)

Bu süreç bazen diğer teknik adımlarla paralel ilerleyen, özellikle karmaşık problemleri analiz etmek, alternatif çözümleri değerlendirmek veya sistem performansını modelleyip anlamak için yürütülen çalışmaları kapsar.

Bir YZ projesinde, sistem analizi süreci, örneğin kritik teknik kararlar öncesinde yapılacak fizibilite çalışmaları, trade-off analizleri (ödünleşim analizleri) ve simülasyonları

içerir. Örneğin, "Veri kümemiz büyürse sistemin performansı ne olacak?" veya "Daha karmaşık bir model daha yüksek doğruluk verir mi, yoksa getirisi az mı olur?" gibi sorular sistem analizi kapsamında ele alınır.

Bu süreçte matematiksel analizler, istatistiksel modeller veya benzetimler kullanılabilir. Sistem analizi, hem tasarım hem de doğrulama aşamalarına girdi sağlayabilir; belirsiz noktaları netleştirir, riskli görünen teknik yaklaşımların önceden test edilmesini sağlar. Kurumsal bağlamda, bu süreç karar vericilere projeyle ilgili önemli içgörüler sunar:

Örneğin, analiz sonucunda bir yaklaşımın teknik olarak sürdürülebilir olmadığı görülürse, erken aşamada yön değiştirilebilir. Bu da zaman ve kaynak israfını önler. Ayrıca sistem analizi, inovasyon ve iyileştirme fırsatlarını da ortaya çıkarabilir. Belki de analiz sırasında, eldeki problem için daha uygun yeni bir YZ yöntemi keşfedilir. Kısacası, sistem analizi süreci YZ projesinin "derin düşünme" adımıdır; riski azaltmak ve en iyi teknik kararları vermek için veriye dayalı öngörüler üretir.

7. Bilgi Edinimi Süreci (Knowledge Acquisition Process)

ISO/IEC 5338'in teknik süreçlerine özgü bir adım olarak, bilgi edinimi süreci YZ sisteminin gerektirdiği uzmanlık bilgisinin toplanmasını ifade eder.

Bu süreç, özellikle uzman sistemler veya bilgi tabanlı YZ uygulamaları için kritik olmakla birlikte, makine öğrenmesi projelerinde de dolaylı olarak önemlidir. Bilgi edinimi sürecinde, alan uzmanlarıyla görüşmeler yapılması, literatür araştırması, mevcut veri ve bilgi kaynaklarının derlenmesi gibi faaliyetler yer alır.

Örneğin, tıbbi teşhis koyan bir YZ modeli geliştirilecekse, doktorların karar verme süreçlerindeki kritik kriterlerin anlaşılması (belki kuralların çıkarılması), tıbbi rehberlerin incelenmesi, önceki vaka verilerinin toplanması bu bilgi edinimi aşamasının bir parçasıdır. Makine öğrenmesi projelerinde de bilgi edinimi, veri söz konusu olmadığında veya veriyle desteklenmesi gereken ön bilgiler olduğunda devreye girer; mesela bir fabrikadaki kalite kontrol YZ sistemi için mühendislerden üretim süreciyle ilgili bilgi alınması, hangi hataların kritik olduğunun öğrenilmesi gibi. Bu süreçte elde edilen bilgi, YZ sisteminin tasarım ve geliştirme kararlarına rehberlik eder.

Örneğin, bir uzman kuralı olarak "X şartı gerçekleştiğinde alarm ver" bilgisi, sistemde bir kural veya ek özellik olarak yer alabilir. Kurumlar açısından bilgi edinimi süreci, sessiz bilgiyi (tacit knowledge) yakalamanın yoludur. Yani çalışanların deneyimleri veya literatürde dağınık halde bulunan bilgiler sistematik biçimde toplanıp proje ekibinin kullanımına sunulur. Bu, hem sistemin performansını artırır (çünkü modele dışarıdan akıl katılmış olur) hem de bilginin kurumsallaşmasını sağlar. Risk yönetimi açısından da değerlidir: Projede kritik bilgiler göz ardı edilirse, sonuçta ortaya çıkan YZ sistemi gerçek dünyada beklenmedik hatalar yapabilir. Bilgi edinimi süreci bu riski azaltmaya yarar.

8. YZ Veri Mühendisliği Süreci (AI Data Engineering Process)

YZ sistemlerinin geleneksel yazılımlardan en büyük farkı, başarısının büyük ölçüde veri kalitesine ve veri yönetimine bağlı olmasıdır. Bu süreç, YZ projelerinde kullanılacak verinin edinilmesi, hazırlanması, işlenmesi ve yönetilmesi adımlarını kapsar. Verinin toplanması (iç veya dış kaynaklardan), temizlenmesi (hatalı/veri dışı kayıtların ayıklanması, tutarsızlıkların giderilmesi), dönüştürülmesi (gerekli formatlara sokulması), etiketlenmesi (eğer süpervizyonlu öğrenme ise doğru etiketlerin insan ya da başka yollarla sağlanması) ve uygun veri kümelerine bölünmesi (eğitim, doğrulama, test) bu sürecin kritik alt faaliyetleridir.

Örneğin, yüz tanıma sistemi örneğimize dönersek, veri mühendisliği sürecinde farklı demografik gruplardan dengeli bir yüz görüntüsü veri seti toplamak, görüntüleri standart bir boyuta getirmek, aydınlatma farklarını normalleştirmek ve her görüntüye doğru kimlik etiketini vermek gerekir. ISO/IEC 5338, YZ veri mühendisliği konusunda özellikle veri gizliliği, güvenliği ve önyargıları giderme konularına dikkat çeker.

Örneğin, eğitim verisinin hassas olabileceği ve korunması gerektiği vurgulanır; bu, geleneksel yazılım geliştirmede pek karşılaşılmayan bir durumdur çünkü klasik yazılımlarda genelde anonim test verileri kullanılır. Aynı şekilde veri kümesindeki dengesizlikler veya yanlı örnekler YZ modelinde sistematik hatalara (bias) yol açabilir, bu nedenle veri mühendisliği sürecinde bu riskleri azaltmak üzere çeşitli teknikler (veri artırma, yeniden örnekleme, bias tespiti ve düzeltme yöntemleri) uygulanır.

Kurumlar için YZ veri mühendisliği süreci, bir YZ projesinin temelini oluşturur: "Garbage in, garbage out" (çöp girerse çöp çıkar) prensibiyle, eğer veri aşamasında hata yapılırsa, en mükemmel algoritma bile kötü sonuç verecektir. Bu yüzden veri mühendisliği

sürecine yatırım yapmak, uzun vadede kaliteyi ve güvenilirliği garanti altına almanın ön koşuludur. Ayrıca, veriyle ilgili yasal sorumluluklar (kişisel verilerin korunması gibi) bu süreçte yönetildiği için, bu adım aynı zamanda yasal uyum ve etik sorumluluk açısından da hayati önem taşımaktadır.

9. Uygulama (Gerçekleştirme) Süreci (Implementation Process)

Tasarımı yapılan YZ sisteminin gerçek kod ve model olarak hayata geçirilmesi bu sürecin konusudur. Yazılım geliştiriciler ve veri bilimciler, belirlenen tasarıma uygun şekilde kodlamayı yaparlar, gerekli algoritmaları yazarlar veya mevcut YZ kütüphanelerini entegre ederler. Makine öğrenmesi modelinin eğitilmesi de bu sürecin bir parçasıdır: veri mühendisliği sürecinde hazırlanmış eğitim verisi kullanılarak model eğitilir (model engineering, model geliştirme faaliyetleri bu aşamaya dahildir, ancak ISO/IEC 5338 bunları ayrı bir süreç olarak değil, uygulama sürecinin bir parçası olarak ele almaktadır). Uygulama sürecinde aynı zamanda birim testleri, kod incelemeleri gibi kalite aktiviteleri de yürütülür; böylece her parçanın doğruluğu erkenden kontrol edilir.

Örneğin, projemizdeki yüz tanıma sistemi için geliştiriciler arayüz kodunu, veri akış kodunu yazar; veri bilimciler yüz tanıma modelini uygun bir derin öğrenme çerçevesinde eğitir; ve sistem entegre bir şekilde çalışacak hale getirilir. Bu süreçte takım genellikle yinelemeli (iteratif) bir yaklaşımla çalışabilir: ilk önce basit bir prototip uygulanır, sonra geliştirilir, hatalar düzeltilir, optimizasyonlar yapılır. Uygulama süreci kurumsal açıdan, inovasyonun vücut bulduğu aşamadır; fikirler somut ürüne dönüşür. Bu aşamada yapılan seçimler (hangi programlama dili, hangi ML çerçevesi, bulut vs. yerel çözüm, açık kaynak vs. ticari araçlar) uzun vadeli bakım ve sahip olma maliyetini etkileyeceğinden, kurum stratejisine uygun kararlar alınmalıdır.

Örneğin, bir kamu kurumu açık kaynak çözümleri tercih ederek lisans maliyetlerini düşük tutmak isteyebilir veya tam tersine bir üretici firmasının desteklediği bir platformu seçerek risk azaltmak isteyebilir. Uygulama süreci sonunda, çalışır durumda bir YZ sistemi elde edilir.

10. Entegrasyon Süreci (Integration Process)

YZ sisteminin farklı bileşenlerinin bir araya getirilerek tümleşik bir sistem oluşturulması bu sürecin işlevini tanımlar. Eğer sistem, birden fazla modülden oluşuyorsa, ki çoğu zaman böyledir, bu modüller tek tek geliştirilip test edildikten sonra entegre edilir.

Örneğin, yüz tanıma sistemimizin arayüzü, veri işleme boru hattı ve YZ modeli ayrı ayrı geliştirildikten sonra entegrasyon sürecinde birleştirilir ve birlikte çalışmaları sağlanır. Bu aşamada, bileşenlerin arasındaki arayüz uyumsuzlukları veya beklenmeyen etkileşim sorunları ortaya çıkabilir, bunlar giderilir. Entegrasyon yalnızca yazılım modüllerini değil, gerekiyorsa donanım ile yazılımın birleşimini de kapsar (örneğin YZ sistemi bir IoT cihazında koşacaksa, cihaz entegrasyonu).

Ayrıca, YZ sisteminin mevcut diğer sistemlere entegrasyonu da bu kapsamda düşünülebilir; bir kuruma yeni alınan bir YZ yazılımının var olan veri tabanlarıyla, web servisleriyle, kimlik doğrulama sistemleriyle entegre edilmesi gibi. Entegrasyon süreci, kademeli veya büyük patlama (big-bang) şeklinde yapılabilir; kademeli entegrasyonda sistem parça parça birleştirilip her adım test edilir, büyük patlamada ise tüm parçalar bir defada bir araya getirilir (YZ projelerinde genellikle kademeli entegrasyon tercih edilir, zira hatayı izole etmek daha kolaydır). Kurumsal anlamda entegrasyon süreci, yeni geliştirilen YZ sisteminin organizasyonun geri kalan ekosistemine sorunsuz bir şekilde eklemlenmesini sağlar. Bu süreçte göz ardı edilecek bir hata, örneğin entegrasyon sonrası bazı verilerin güncellenmemesi veya sistemler arası bir güvenlik açığı, ciddi problemlere yol açabilir. Dolayısıyla, entegrasyon süreci hem teknik risklerin (uyumsuzluk, veri kaybı) yönetilmesi hem de proje çıktısının kullanılabilir bir çözüm haline gelmesi için şarttır.

11. Doğrulama Süreci (Verification Process)

Doğrulama, geliştirilmiş YZ sisteminin belirlenen gereksinimlere uygunluğunu teknik olarak kontrol eden süreçtir. "Doğru şeyi inşa ettik mi?" sorusundan ziyade, "İnşa ettiğimiz şeyi doğru yaptık mı?" sorusuna yanıt arar. Geliştirme sürecinde her bileşen veya alt sistem için planlanmış test ve inceleme faaliyetleri doğrulama kapsamındadır.

Örneğin, yazılım testleri (birim testi, entegrasyon testi, sistem testi), model değerlendirmeleri (ayrılmış test veri kümesi üzerinde modelin doğruluk, hassasiyet, F1 skoru gibi metriklerinin ölçülmesi), güvenlik testleri (penetrasyon testleri, veri sızıntı kontrolleri) bu sürecin parçalarıdır. Doğrulama sürecinde her gereksinim maddesi ele alınır ve o gereksinimin karşılandığını gösteren kanıtlar sunulur.

Örneğin, bir gereksinim "sistem 1 saniyenin altında bir yanıt süresine sahip olmalı" ise, performans testi yapılarak bu şartın sağlandığı gösterilir. Keza "model kararları belirli bir doğruluk eşiğini aşmalı" gereksinimi varsa, test sonuçları raporlanır. Doğrulamada tespit

edilen hatalar veya eksikler, proje ekibine geri bildirim olarak döner ve düzeltmeler yapılır; ardından tekrar doğrulama testlerine tabi tutulur.

Kurumlar için doğrulama süreci, kalite güvencesinin teknik belkemiğidir. Bu sayede, sistem daha işin başındayken (kullanıcıya veya canlı ortama çıkmadan önce) iç kontrol edilmiş olur. Bu da riskleri ciddi ölçüde azaltır; çünkü hatalar erken tespit edilmezse, ileride kullanıcı aşamasında sorunlara, memnuniyetsizliğe veya daha kötüsü güvenlik açıklarına yol açabilir. Ayrıca, bazı sektörlerde (özellikle sağlık, otomotiv, savunma gibi) doğrulama adımları düzenleyici uyum için zorunludur. Standardın tanımladığı doğrulama süreci, bu gerekliliklerin sistematik biçimde yerine getirilmesini sağlamaktadır.

12. Geçiş (Transfer) Süreci (Transition Process)

Bu süreç, doğrulanmış YZ sisteminin operasyonel ortama alınması, yani gerçek kullanım ortamına geçirilmesi ile ilgilidir. Geliştirme ve test faaliyetleri sonucunda hazır hale gelen sistem, son kullanıcıların veya hedef operasyon ortamının hizmetine sunulur. Geçiş sürecinde yapılacak işler arasında, sistemin kurulumu/yüklenmesi, kullanıcıların eğitilmesi, gerekli verilerin (örneğin eski sistemden yeni sisteme) taşınması, başlangıç konfigürasyonunun yapılması ve eğer gerekiyorsa pilot uygulama/deneme kullanımı yer alır.

Örneğin, yüz tanıma sistemimizi ele alırsak, bu aşamada yazılımın hastanedeki sunuculara kurulumu, ilgili personelin yazılımı nasıl kullanacağı konusunda eğitimi, sistemin ilk çalışması için gerekli parametrelerin ayarlanması (belki belli kullanıcı erişim haklarının tanımlanması, kamera entegrasyonlarının yapılması) gerekebilir.

Geçiş süreci sırasında beklenmedik sorunlar çıkabilir (farklı donanım ortamında performans düşüşü, kullanıcıların arayüzde zorlandığı noktalar vb.), bu nedenle yakın izleme yapmak ve hızlı düzeltmeler uygulamak önemlidir. Bu süreçte, proje ekibi ile operasyon ekibi arasında sıkı bir işbirliği gerekir; zira proje bitip ürün teslim edilmek üzereyken, sorumluluk operasyona/devreye alma ekibine geçmektedir. Kurumlar için geçiş süreci, yatırımın hayata geçmesi demektir. Tüm geliştirme çabasının gerçek dünyada değer üretmeye başlaması bu noktada olur. Eğer geçiş iyi yönetilmezse, mükemmel bir teknik ürün bile kullanıcı hataları veya altyapı uyumsuzlukları yüzünden başarısız görünebilir. Bu da potansiyel olarak kabul görmeme riski yaratır. Dolayısıyla, geçiş süreci risk yönetimi açısından da planlanmalıdır:

Örneğin, kritik bir sisteme aşamalı geçiş (ilk önce az sayıda kullanıcıyla pilot, sonra tam ölçekli dağıtım) yapılarak risk azaltılabilir.

13. Geçerlilik (Validasyon) Süreci (Validation Process)

Geçerlilik, sistemin amaca uygunluğunu ve paydaşların beklentilerini karşılayıp karşılamadığını doğrulayan süreçtir. Doğrulamadan farkı, teknik gereksinimlerin ötesine geçip, "Doğru şeyi inşa ettik mi?" sorusunu yanıtlamasıdır. Bu aşamada, özellikle sistem gerçek kullanım koşullarında test edilir ve paydaşlarla birlikte değerlendirilir.

Örneğin, geliştirdiğimiz yüz tanıma sistemi hastane ortamında belirli bir süre çalıştırılır ve kullanıcılardan (doktorlar, hemşireler, IT personeli) geri bildirim toplanır. Sistemin gerçekten teşhis sürecine beklenen katkıyı verip vermediği, kullanıcıların memnuniyeti, sistemi kullanırken ortaya çıkan beklenmedik durumlar validasyonun konusu olabilir. Eğer paydaş ihtiyaçları tam karşılanmıyorsa (örneğin, kullanıcılar arayüzü karmaşık bulduysa veya sistem bazı yüzleri tanımakta zorluk çekiyorsa), bu geri bildirimler sistemde iyileştirmeler yapmak üzere değerlendirilir. Validasyon süreci genellikle bir kabul testi veya kullanıcı kabulü şeklinde de uygulanabilir. Müşteri veya temsilcileri, sistemin kontratta belirtilen kriterleri karşılayıp karşılamadığını sınarlar.

Kurumsal bakımdan validasyon, projenin başarı ölçütlerinin dış gözle (kullanıcı gözüyle) onaylandığı aşamadır. Bu süreçte aynı zamanda YZ sisteminin etik ve yasal uygunluğu da gözden geçirilir; örneğin bir etki değerlendirmesi yapılıp sistemin önyargısız çalışıp çalışmadığı, GDPR gibi yasalara uyduğu tekrar kontrol edilebilir. Validasyon tamamlanıp başarılı olduğunda, kurum artık güvenle sistemi canlı kullanıma devam edebilir. Aksi halde, yani validasyon sorunlu geçerse, sistem tam olarak devreye alınmadan düzeltici adımlar atılabilir (ve gerekiyorsa bir daha validasyon yapılır). Bu süreç, kalite yönetiminin ve risk yönetiminin son kontrol noktası olarak görülebilir. Yanlış giden bir şey varsa en son burada yakalama şansı vardır.

14. Sürekli Geçerlilik Süreci (Continuous Validation Process)

ISO/IEC 5338, klasik yazılım yaşam döngüsü süreçlerine yenilikçi bir ekleme yaparak sürekli geçerlilik kavramını getirmiştir. Bu süreç, YZ sistemi operasyonel kullanıma girdikten sonra da performansının, doğruluğunun ve uygunluğunun düzenli aralıklarla gözden geçirilmesini içerir. Makine öğrenmesi tabanlı sistemler, dağıtıma girdikten sonra zamanla "bayatlayabilir" (model stale) hale gelebilir. Yani gerçek dünyadaki değişimlere bağlı olarak başlangıçtaki doğruluklarını yitirebilirler. Sürekli geçerlilik sürecinde, sistem çalışmaya devam ederken, belirlenen metrikler izlenir (örneğin modelin tahmin doğruluğu, gelen verinin

istatistiksel dağılımı, kullanıcı memnuniyeti gibi). Eğer sistem performansında bir düşüş veya beklenmeyen bir sapma gözlemlenirse, bu bir uyarı olarak ele alınır.

Örneğin, yüz tanıma sistemimiz bir süre sonra yeni gelen hasta profillerinde daha fazla hata yapmaya başladıysa, bu durum veri dağılımının değiştiğini veya modele yeni bir güncelleme gerektiğini gösterebilir. Sürekli validasyon süreci kapsamında, düzenli aralıklarla yeniden testler yapılabilir, kullanıcı geri bildirimleri toplanabilir veya modelin tahminleri üzerinde kalite kontrolleri (ör. rastgele örneklerin insan tarafından doğrulanması) yürütülebilir. Bu sürecin çıktısına göre, modele yeniden eğitim (başka bir teknik süreç olan bakım süreciyle bağlantılı olarak) veya sisteme ek iyileştirmeler gündeme gelir.

Kurumlar için sürekli geçerlilik, YZ sisteminin ömrü boyunca güvenilir ve etkili kalmasının anahtarıdır. Bu sayede, sistem devreye alındıktan sonra "kur ve unut" yapılmaz; aksine, sistem sanki canlı bir organizmaymış gibi düzenli kontrol ve bakım altında tutulur. Bu da sürdürülebilirlik ve sürekli iyileştirme ilkelerine uygun bir yönetişim sağlar. Özellikle yüksek riskli YZ uygulamalarında (örneğin, sağlık teşhis, otonom sürüş), sürekli validasyon yapılmaması ciddi sonuçlar doğurabilir; standardın bu süreci öngörmesi, kurumsal risk yönetimine önemli bir katkıdır.

15. İşletim Süreci (Operation Process):

YZ sisteminin günlük operasyonel kullanımını kapsayan süreçtir. Sistem devreye alındıktan sonra, kullanıcılar tarafından gerçek ortamında kullanılırken, operasyon süreci kapsamında sistemin izlenmesi, kullanıcı desteği sağlanması, olay ve sorun yönetimi yapılması sağlanır.

Örneğin, IT operasyon ekipleri sistemin sunucularının düzgün çalıştığını, yanıt sürelerinin beklenen düzeyde olduğunu takip eder; gerektiğinde sistemi yeniden başlatma, log'ları izleme, kapasite artırımı gibi işleri yapar. Kullanıcılardan gelen destek talepleri (örneğin "şu vakayı tanımlamakta zorlandı, ne yapmalıyım?" gibi sorular) yanıtlanır. İşletim süreci, sistemin hizmet düzeyi hedeflerini (SLA) tutturmasına odaklıdır; eğer YZ sistemi bir hizmet sunuyorsa, belirli bir kesintisizlik ve performans seviyesi vaat edilmiş olabilir, bu süreç bunları izler.

Ayrıca, yeni çıkan güvenlik yamalarının uygulanması gibi rutin işler de işletim sürecine dahildir (bu durum biraz bakım süreciyle de ilişkilidir). Kurum açısından işletim süreci, YZ sisteminin vaat ettiği faydayı sürekli ürettiği aşamadır. Bu sürecin iyi yönetilmesi, sistemin güvenilir ve kullanıcı dostu kalmasını sağlar.

Örneğin, bir bankanın kredi başvurularını değerlendiren YZ sistemi düşünelim; operasyon sürecinde eğer sistem bir süre cevap vermemeye başlarsa hemen müdahale edilmesi, kullanıcıların (kredi birimi çalışanlarının) işinin aksamasını önler. İşletim süreci aynı zamanda sistemle ilgili olay kayıtlarının (incident) tutulduğu, bunların raporlandığı bir süreçtir. Bu veriler daha sonra iyileştirme için kullanılabilir. Bu süreçteki bulgular, sürekli validasyon ve bakım süreçlerine de girdi sağlar; örneğin sık sık aynı tür hatalar raporlanıyorsa, bu belki modelin güncellenmesi gerektiğine işaret eder.

16. Bakım Süreci (Maintenance Process)

YZ sisteminin operasyonel kullanım sırasında ortaya çıkan hata düzeltme, iyileştirme, adaptasyon veya güncelleme faaliyetlerini kapsar. Hiçbir sistem durağan değildir; yazılımlar hata içerebilir, kullanıcı ihtiyaçları değişebilir veya dış koşullar (iş kuralları, veri yapısı, yasal mevzuat) evrilebilir. Bakım süreci, YZ sisteminin bu değişimlere ayak uydurmasını sağlar. Bu kapsamda, tespit edilen hataların giderilmesi (örneğin bir yüz tanıma algoritmasının belirli ışık koşullarında çalışmama hatası varsa, bunun düzeltilmesi), performans iyileştirmeleri (daha hızlı çalışması için kod optimizasyonu), yeni özellik eklenmesi (kullanıcılar talep etmişse) veya modelin yeniden eğitilmesi (örneğin modele yeni gelen verilerle periyodik olarak öğrenme yeteneği kazandırmak) gibi işler yer alır.

YZ sistemlerinde bakımın bir özel yönü, modelin yeniden eğitimi veya yeni veriyle adapte edilmesidir ki bu klasik yazılım bakımından farklılık gösterir. Yazılım kodunu hiç değiştirmeden, sadece modeli yeni veriyle güncellemek de bir bakım faaliyeti olabilir. Ayrıca, YZ sisteminin bağlı olduğu dış servislerdeki değişiklikler (API versiyon değişikliği gibi) de bakım kapsamında ele alınır. Kurumlar için bakım süreci, YZ sisteminin ömür boyu değer üretmeye devam etmesi için kritik bir süreçtir. İyi bir bakım yönetimi, sistemin ömrünü uzatır ve toplam sahip olma maliyetini kontrol altında tutar.

Risk yönetimi açısından ise, bakım yapılmaması veya ihmal edilmesi büyük riskler doğurur: Örneğin güvenlik açıkları yamalanmazsa veri ihlali riski artar, model güncellenmezse doğruluk düşer ve yanlış kararlar riski artar. Bu nedenle, standardın tanımladığı bakım süreci, YZ projelerinin uzun soluklu başarısı için bir nevi sigorta poliçesidir. Sistem yaşlandıkça bile güvenli, verimli ve ilgili kalmasını temin eder.

17. İmha (Elden Çıkarma) Süreci (Disposal Process)

Son teknik süreç olarak, YZ sisteminin ömrünün sonunda güvenli ve düzenli bir şekilde devreden çıkarılması bu sürecin konusudur. Her teknolojinin bir yaşam döngüsü olduğu gibi, YZ sistemi de bir noktada ya güncelliğini yitirebilir, yerine yeni bir sistem gelebilir veya proje iptal edilebilir. İmha süreci, bu durumda sistemin devre dışı bırakılmasını planlar ve uygular. Bu kapsamda, sistemin kullandığı verilerin arşivlenmesi veya silinmesi (özellikle kişisel veya hassas veriler varsa, bunların güvenli imhası kritik), donanım bileşenleri varsa bunların sökümü, yazılım lisanslarının iptali, kullanıcıların bilgilendirilmesi gibi adımlar yer alır.

Örneğin, hastanemizdeki yüz tanıma sistemi daha gelişmiş bir sürümle değiştirilecekse, eski sistem imha sürecinde tamamen kapatılırken o sistemdeki hasta verileri uygun şekilde yeni sisteme aktarılıp sonra eski ortamdan silinebilir, eski sunucular kapatılıp geri dönüştürülebilir, kullanıcı erişimleri kaldırılır. İmha süreci, aynı zamanda kazanılan derslerin son bir değerlendirmesini yapmayı da içerebilir. Proje boyunca öğrenilenler, hatalar ve başarılar bir raporla belgeleyip kurumsal bilgi havuzuna konulur (bu, 6.2'deki bilgi yönetimiyle de bağlantılıdır). Kurum için imha süreci, riskli bir geçiş aşaması olabilir; zira bir sistemi kapatırken hizmet kesintisi yaşanmaması, veri kaybı olmaması önemlidir. Bu yüzden imha süreci de en az diğer süreçler kadar planlı yapılmalıdır.

Ayrıca, yasal açıdan, verilerin uygun şekilde saklanması/imhası, sözleşmelerin sonlandırılması gibi konular da burada ele alınır. Sürdürülebilirlik perspektifinden, eski sistemin bertaraf edilmesi çevresel ve etik şekilde yapılmalıdır (örneğin elektronik atıkların doğru yönetilmesi). Standardın bu süreci dahil etmesi, YZ sistem yaşam döngüsüne tümleşik bir bakış açısı getirdiğinin göstergesidir. Sistem sadece kurulup bırakılmaz, kullanım ömrü bitene kadar ve bittikten sonra bile sorumlu bir yönetim yaklaşımıyla ele alınır.

Görüldüğü gibi teknik süreçler, bir YZ sisteminin yaşam döngüsündeki her adımı kapsayarak, projenin fikir aşamasından değer üretimine ve sonlandırılmasına kadar sistematik bir yol haritası sunmaktadır. Kurumlar için bu süreçlerin her biri, farklı risklerin ele alındığı ve yönetildiği birer evredir.

İş analizi ve gereksinim tanımlama ile:

- Amaç ve kapsam riski kontrol altına alınır;
- Mimari ve tasarım ile teknik çözüm riski azaltılır;

- Veri mühendisliği ile veri kaynaklı riskler (kalite, önyargı, gizlilik) yönetilir;
- Uygulama ve entegrasyon ile uygulama hataları riski düşürülür;
- Doğrulama ve validasyon ile kalite ve uygunluk riskleri giderilir;
- Sürekli validasyon ve bakım ile performans düşüşü veya eskime riskleri önlenir;
- İşletim ile operasyonel riskler (kesinti, hata) kontrol edilir;
- İmha ile de sonlandırma riskleri (veri sızıntısı, kalıcı hasar) engellenir.

Bu bütüncül yaklaşım, ISO/IEC 5338'in kuruluşlara sağladığı en önemli katkılardan biridir: YZ sistemlerinin geliştirilmesi ve yönetiminde, süreç odaklı, disiplinli ve sorumlu bir çerçeve sunmak. Sonuç olarak, 6.4 bölümündeki teknik süreçler, YZ projelerinin sadece teknik başarısını değil, aynı zamanda etik, güvenlik, şeffaflık ve sürdürülebilirlik boyutlarını da gözeten bir uygulama kılavuzu niteliğindedir. Bu sayede kuruluşlar, YZ teknolojilerini yenilikçi bir biçimde kullanırken, ortaya çıkabilecek sorunlara karşı hazırlıklı olabilir ve YZ sistemlerinin güvenilirliğini toplum nezdinde artırabilirler.

1.2. Kaynakça

Institute of Electrical and Electronics Engineers. (2021). *IEEE Standard Model Process for Addressing Ethical Concerns During System Design* (IEEE Standard 7000-2021). https://doi.org/10.1109/IEEESTD.2021.9536679

International Organization for Standardization & International Electrotechnical Commission. (2022a). *ISO/IEC 23053:2022 Framework for Artificial Intelligence (AI) Systems Using Machine Learning (ML).* https://www.iso.org/standard/74438.html

International Organization for Standardization & International Electrotechnical Commission. (2022b). *ISO/IEC 22989:2022 Information technology — Artificial intelligence — Artificial intelligence concepts and terminology.* https://www.iso.org/standard/74296.html

International Organization for Standardization & International Electrotechnical Commission. (2022c). *ISO/IEC TR 24368:2022 Information technology — Artificial intelligence — Overview of ethical and societal concerns.* https://www.iso.org/standard/78507.html

International Organization for Standardization & International Electrotechnical Commission. (2023a). *ISO/IEC 5338:2023 Information technology — Artificial intelligence — YZ system life cycle processes.* https://www.iso.org/standard/81118.html

International Organization for Standardization & International Electrotechnical Commission. (2023b). *ISO/IEC 23894:2023 Information technology — Artificial intelligence — Guidance on risk management.* https://www.iso.org/standard/77304.html

International Organization for Standardization, International Electrotechnical Commission, & Institute of Electrical and Electronics Engineers. (2023). *ISO/IEC/IEEE 15288:2023 Systems and software engineering — System life cycle processes.* https://www.iso.org/standard/81702.html


Etik Yapay Zekâ ve
Sosyal Bilimler

2. ETİK YAPAY ZEKÂ VE SOSYAL BİLİMLER

Dr. Duygu Özalp[a], Ulaş Çakır[b], Nidanur Cabbar[c]

[a] *Başkent Üniversitesi, Sosyoloji Bölümü, h.duygu.ozalp@gmail.com*
[b] *Savunma Sanayi Başkanlığı, ulas_cakir@yahoo.com*
[c] *TBD Genç Denetleme Koordinatörü, kadriyenidanurc@gmail.com*

2.1. Yapay Zekâ Yaşam Döngüsü ve Sorumlu YZ

ISO/IEC 5338:2023 standardı, yapay zekâ (YZ) sistemlerinin tanımlanması, kontrolü, yönetimi, yürütülmesi ve iyileştirilmesini kapsayan 'YZ yaşam döngüsü' için kapsamlı bir süreç seti tanımlamaktadır (International Organization for Standardization-ISO, 2023). Burada ele alınan YZ yaşam döngüsünün yalnızca teknik bir süreç değil, aynı zamanda yönetişim ve sorumluluk boyutlarını da içeren çok katmanlı bir yapı olduğu görülmektedir. Söz konusu sürecin her aşamasında şeffaflık, hesap verebilirlik ve etik ilkelerin gözetilmesinin, YZ'nin sürdürülebilir ve güvenilir bir şekilde toplum yararına çalışması için gerekli olduğu anlaşılmaktadır. Çünkü YZ yaşam döngüsü, YZ teknolojilerinin sistematik ve güvenilir şekilde geliştirilmesini sağlayan bir çerçeve sağlamaktadır. Döngüdeki her aşama birbirini tamamlamakta, sürekli geri besleme mekanizmalarıyla daha güçlü ve daha etik YZ uygulamalarının ortaya çıkmasına katkıda bulunmaktadır. Bu nedenle araştırmacılar, mühendisler, politika yapıcılar ve kullanıcılar açısından YZ yaşam döngüsünü anlamak ve uygulamak büyük önem taşımaktadır.

Aşağıda, ilk olarak sosyal bilimlerin YZ yaşam döngüsündeki rolüne yer verilmektedir. Ardından ISO/IEC 5338:2023'te yer alan YZ yaşam döngüsü ve (*eski adıyla*) ABD Savunma Bakanlığı'nın (DoD) Sorumlu YZ Araç Seti (Responsible AI Toolkit-RAI) ve/veya etik uzmanlığı perspektifinden alıntılanan alt süreçler ele alınmaktadır. Alıntılanan her alt süreç, eğitim ve değerlendirme bağlamlarında üretken YZ araçlarını kullanan bir sosyal bilimler/beşeri bilimler araştırmacısının bakış açısından ele alınmakta etik boyutlarla ilişkilendirilmektedir.

2.1.1. Sosyal Bilimlerin YZ Yaşam Döngüsündeki Rolü

YZ yaşam döngüsü genellikle teknik bir çerçeve olarak kurgulansa da bu sürece sosyal bilim araştırmacılarının katkı sağlaması önem taşımaktadır. Sosyo-teknik sistem yaklaşımı çerçevesinde değerlendirildiğinde YZ sistemleri teknik, toplumsal ve kurumsal unsurların etkileşiminden oluşmaktadır (Bijker, Hughes ve Pinch, 1987) ve sosyal bilim araştırmacıları, bu etkileşim etrafında oluşan problemlerin toplumsal bağlamını ve paydaş beklentilerini, tüm sistemin yer aldığı toplumsal, kültürel, ekonomik, siyasal vd. çevre açısından analiz etmektedir. Başka bir ifadeyle sosyal bilim araştırmacıları etik ilkelerin; toplumsal adalet, kapsayıcılık ve temsil açısından uygulanabilirliğini incelemekte, kullanıcıların teknolojiyle etkileşim biçimlerini çok yönlü bir yaklaşımla değerlendirmektedir. Sistem çıktılarının toplumsal sonuçlarını ve olası etkilerini inceleyerek, toplumsal temsiliyet için katkı sağlamaktadır.

Özellikle sosyoloji, kamu yönetimi, eğitim, sosyal politika, sağlık, psikoloji gibi sosyal ve beşeri bilimlerin katkısı, YZ projelerinin toplumsal kabul edilebilirliğini ve adalet boyutunu güçlendirmektedir.

Sosyal bilimler;

(i) problemlerin ve bağlamın tanımlanması,

(ii) verinin yönetişimi,

(iii) toplumsal etki değerlendirme ve

(iv) izleme ve yeniden değerlendirme gibi süreçlere katılım ve katkı sağlayarak

bir YZ sisteminin geliştirilmesi aşamalarında önemli görevler almaktadır.

Böylelikle sosyal bilim araştırmacıları araştırma sorularının ve/veya uygulamaların yalnızca teknik doğruluk değil, toplumsal fayda ve etik sorumluluk açısından da geçerli olmasını sağlamaktadır. Bu aşamada kullanılan kavramların kültürel ve politik bağlamı göz önünde bulundurulmakta böylelikle toplumsal uyumlaştırma gerçekleştirilmektedir. Ayrıca kişisel verilerin korunması (KVK), bilgilendirilmiş onam süreçleri ve önyargı analizlerinin sosyal bilimcilerin desteği olmaksızın yapılması bu analizlerin eksik kalması anlamına gelecektir. Bunlara ek olarak kavramların çok-anlamlılığının yaratabileceği kavramsal belirsizliklerin çözülmesi kodlama kılavuzları, dipnot (anotasyon) yönergeleri ve bağımsız kodlayıcı uyum testleri kullanılarak sosyal bilimcilerin alan bilgisiyle desteklenmeye açık alanlar olarak belirmektedir.

Bir diğer önemli konu, YZ çıktılarının yalnızca doğruluk değil, adalet, çeşitlilik ve hesap verebilirlik ilkeleri çerçevesinde ölçülmesi gerekliliğinden doğmaktadır. Sosyal bilimler

burada farklı gruplar üzerindeki etkileri analiz edebilecek ayrıca kullanıcı deneyimi verilerinin toplumsal anlamını, temsil düzeyini ve etik etkilerini analiz ederek, sistemin insan ve toplum odaklı geliştirilebilmesine katkıda bulunabilecektir.

ISO/IEC 5338, yaşam döngüsüne dair kapsamlı bir teknik şema sunsa da sosyal bilimler açısından 'yeterlilik' ile 'zorunluluk' arasındaki fark belirleyicidir. Standart; paydaş ihtiyaçlarının sistematik belirlenmesi, bilgi edinimi ve risk yönetimi için metodolojik bir çatı/iskelet sağlamaktadır. Ancak sosyal bilimler açısından, bu çatının/iskeletin etik kurullar, disiplinler arası işbirliği ve katılımcı yöntemlerle zenginleştirilmesi gerekmektedir. Bu nedenle ISO/IEC 5338, sosyal bilim araştırmacıları için temel bir referans niteliği taşımakla birlikte tek başına yeterli görülmemekte; UNESCO (2021) Etik YZ Tavsiyeleri ve OECD (2019) YZ İlkeleri gibi tamamlayıcı çalışmalarla desteklenmesi gerektiği düşünülmektedir.

Sonuç olarak sosyal bilim araştırmacılarının, YZ yaşam döngüsünde yalnızca 'etik danışman' değil, sürecin kurucu ortağı olarak görülmesi önem taşımaktadır. Bu rol, geliştirilmesinden problem tanımlarına hatta imha süreçlerine kadar tüm aşamalarda metodolojik, etik ve toplumsal perspektiflerle katkı sunmayı içermektedir. ISO/IEC 5338 standardı bu katkıyı çerçevelemek için değerli bir araç olsa da sosyal bilimler bağlamında eleştirel, katılımcı ve disiplinler arası yaklaşımlarla tamamlanmadığında sınırlı kalması olasıdır.

2.1.2. Risk Yönetimi Süreci

YZ sistemleri için risk yönetimi süreci (ISO/IEC 5338 – Bölüm 6.3.4) çok disiplinli bir şekilde yürütülmelidir. Risklerin belirlenmesi, değerlendirilmesi ve ele alınması, çeşitli alanlardan uzmanların işbirliğini içermelidir. Bu süreçte işbirliği içinde olması gereken uzmanlık alanları hukuk, bilgi güvenliği alanlarının yanı sıra yanı sıra etik, sosyoloji ve psikoloji gibi sosyal ve beşeri bilimler alanlarını kapsamaktadır. Örneğin hukuk alan uzmanları, kişisel verilerin korunması, gizlilik haklarının korunması; bilgi güvenliği alan uzmanları veri ihlalleri, siber tehditler ve güvenlik açıklarıyla ilgili risklerin belirlenmesi ve azaltılması; etik alan uzmanları ayrımcılığın yasaklanması, sosyoloji ve psikoloji alan uzmanları ise etkilenen toplumsal grupların risk algısı da dâhil olmak üzere toplumsal ve bireysel etkilerin analizi gibi konular üzerine çalışabileceklerdir.

Burada anılan tüm sosyal ve beşeri bilim alanları, risk analizine teknik hususların ötesinde toplumsal sonuçları da dâhil ederek katkıda bulunmalıdır. Örneğin, bir YZ sistemi tarafından yapılan yanlış sınıflandırmanın bir topluluğun damgalanmasına veya mevcut

eşitsizliklerin görünürlüğünün artmasına yol açıp açmayacağına yönelik çok yönlü değerlendirmeler sosyal bilimcilerin bilgi, bakış açısı ve deneyimi ile irdelenmelidir.

Risk değerlendirmesinin de teknik faaliyetlerle sınırlı olmaması önemlidir. Potansiyel sonuçların tamamının ele alınmasını sağlamak için hukuki, etik ve toplumsal perspektifleri taşımalıdır. Kapsamlı ve etkili bir risk yönetimi elde etmek için çoklu disiplinlerin aktif katılımı gerekmektedir.

Dolayısıyla sosyal bilimciler bu süreçte,

- Geliştiricilerin ve karar vericilerin sosyo-teknik riskler konusunda bilgilendirilmesi, (başka bir ifadeyle, sosyolojik içgörüler ile teknik tasarım arasında köprü kurulması),
- Paydaşlar üzerindeki etkinin ve/veya meydana gelen zararın değerlendirilmesi amacıyla etik etki değerlendirmelerin yapılması,
- Algoritmik önyargı ve gizlilik endişeleri dâhil olmak üzere potansiyel zararların ortaya konulması ile
- Azaltma stratejileri önerilmesi ve paydaşların sosyo-teknik riskler hakkında bilgilendirilmesi gibi rolleri üstlenebileceklerdir.

Anlaşılacağı üzere burada sosyal bilim araştırmacılarının, YZ yaşam döngüsünün risk yönetimi aşamasına dâhil olmaları gerektiği düşüncesi öne çıkmaktadır. Bu süreç, YZ projesindeki risklerin belirlenmesi, analiz edilmesi ve azaltılmasını içermektedir. Uygulamada sosyal bilimler, etik ve toplumsal riskleri (örneğin önyargı, gizlilik ihlâli, dezenformasyon, topluluklar üzerinde eşit olmayan etki gibi) belirleyerek ve bunların olasılığını ve ciddiyetini değerlendirmeye yardımcı olmaktadır.

Dolayısıyla sosyal bilim araştırmacılarının mevcut çalışma yöntemlerinde resmi sosyal etki analizlerini ve paydaşların erken aşamadaki görüşlerini entegre ederek uygulamalarını güncellemesi gerekliliği de doğmaktadır. Örneğin, etik riskler için genişletilerek yapılandırılmış risk çerçeveleri (ISO 31000 risk yönetimi kavramlarından yararlanarak) benimsenmeli, risk kriterlerini güncellemek için devam eden YZ denetimlerinden elde edilen bulgular da dâhil edilmelidir. Ayrıca ISO 31000 gibi YZ'ye uyarlanmış yerleşik risk standartları, tutarlılığı sağlamak için bu konuda yardımcı olmaktadır. ISO 5338'in kendisinde etik kılavuzlar belirtilmiştir: "*YZ ile ilgili etik ve toplumsal kaygılara genel bakış*" ISO/IEC TR 24368'de yer almaktadır ve "*sistem tasarımı sırasında etik kaygıların ele alınması*" IEEE 7000-

2021'de ele alınmaktadır. Bunlar, hangi standartların dikkate alınması gerektiğini de belirtmektedir.

Riskleri etkili bir şekilde yönetmek için işbirliği yapılmasının hayati önemine metin boyunca dikkat çekilmektedir. Bu aşamada sosyal bilimcilerin veri bilimcileri, YZ mühendisleri, proje yöneticileri ve gizlilik alan uzmanlarıyla yakın bir şekilde çalışmaları bir gereklilik olarak belirmektedir. Örneğin, model önyargı metriklerini yorumlamak için bir YZ mühendisiyle, uyumluluk konularında hukuk uzmanlarıyla ve toplumsal etkileri anlamak için sosyologlarla ekip oluşturması önemli görülmektedir. *RAI Toolkit*'in etik alanı da bu disiplinler arası yaklaşımı vurgulamakta; örneğin, yüksek riskli YZ bileşenlerinde açıklanabilirlik ve insan denetimi gerektirerek, ekiplerin "*etik ilkeleri somut geliştirme uygulamalarına bağlamalarını*" teşvik etmektedir (Slagg, 2025).

Kısaca, yöntemlerin güncellenmesi, yapılandırılmış sosyal etki analizlerinin ve paydaşların erken aşamadaki katkılarının entegrasyonunu gerektirmektedir. Risk yönetimi için ISO 31000 ve etik tasarım için IEEE 7000-2021 gibi standartlar, tutarlı ve sorumlu uygulamaları sağlamak için kullanılmaktadır (IEEE, 2021; ISO, 2018). İşbirliği, veri bilimcileri, YZ mühendisleri, hukuk uzmanlarını, sosyologları ve ilgili tüm sosyal ve beşeri bilimler alan uzmanlarını kapsamalıdır. Önerilen yöntemler arasında risk kayıtları, kontrol listeleri ve insan faktörünün dâhil olduğu inceleme çerçeveleri yer almaktadır (Floridi ve ark., 2018).

2.1.3. Kalite Güvence Süreci

Kalite güvence sürecinde (ISO/IEC 5338 – Bölüm 6.3.8) işbirliği kalite güvencesi alan uzmanları, veri bilimcileri, kullanıcı deneyimi tasarımcıları (user experience design-UX) ve etik uzmanlarını içermektedir. Bu işbirliği ayrıca istatistik ve metodoloji uzmanları, bilgi sistemleri uzmanları ve bağımsız iç ve dış denetçiler olmak üzere çeşitli uzmanlık alanlarını da kapsadığında daha gerçekçi ve uygulanabilir çalışmalar yapması olanaklı olacaktır.

Sosyal bilim araştırmacılarının bu sürece katkısı kalite güvence planına geçerlik-güvenirlik kriterlerini eklemeleri, veri toplama araçlarının standardizasyonunu gözetmeleri konularında olabilecektir. Böylesi bir işbirliğinin varlığı, araştırma sonuçlarının yeniden üretilebilirliğini kuvvetlendirecektir. Anlaşılacağı üzere, sosyal bilimcilerin kalite güvence süreçlerine katılımı, teknik kalitenin yanı sıra etik ve toplumsal normların da dikkate alınmasını sağlama noktasında önemli görülmektedir. Çünkü görevler arasında etik kalite kriterlerinin tanımlanması, veri kümelerinin incelenmesi, sistem çıktılarının doğrulanması ve kullanıcıya

sunulan açıklamaların test edilmesi yer almaktadır. Yöntemlerin güncellenmesi, adalet ölçütlerinin ve toplumsal testlerin dâhil edilmesini gerektirmektedir. Önerilen yöntemler arasında model kartları, veri kümeleri için veri sayfaları, akran değerlendirmeleri bulunmaktadır (Mitchell et al., 2019).

Dolayısıyla kalite güvencesinde sosyal bilimcilerin rolü, YZ sisteminin sadece teknik kalite standartlarını değil, aynı zamanda adalet, hesap verebilirlik, şeffaflık gibi etik normları da karşıladığından emin olmaktır. Sosyal bilimciler kalite güvencesi denetimlerine ve test planlarına katılarak (i) önyargı eşikleri gibi etik ve toplumsal kriterleri tanımlamakta, (ii) temsil ve etik kaygılar açısından eğitim ve test verileri kümelerini incelemekte, (iii) sistemin çıktılarının toplumsal yapı ve değerler ile uyumlu olduğunu doğrulamakta ve (iv) kullanıcılara sunulan açıklamaların veya belgelerin onlar tarafından anlaşılır olup olmadığını test etmektedir.

Yukarıda da yer verildiği gibi sosyal bilimcilerin sürecin aşamalarına katılımı ve işbirliği, kalite güvencesinde önemli görülmektedir. Örneğin, adalet kısıtlamalarını uygulamak için YZ mühendisleriyle, insan-yapay zekâ etkileşimini iyileştirmek için UX tasarımcılarıyla ve kullanıcı verilerini yorumlamak için etik uzmanları veya sosyologlarla koordinasyon sağlamaları değerli olduğu ayrıca RAI Toolkit'in de bu tür ekip çalışmasını vurguladığı görülmektedir.

2.1.4. İş veya Görev Analizi Süreci

Bu aşamada (ISO/IEC 5338 – Bölüm 6.4.1), görevler YZ sisteminin etik ve toplumsal hedeflerle uyumlu olmasının sağlanması, iş/görev hedeflerinin ortaya çıkarılması ve etik kısıtlamaların belirlenmesi ile ilgilidir. Etik etki analizini iş/görev planlamasına entegre etmek için de mevcut yöntemlerin güncellenmesi gerekmektedir. Standartlar arasında sosyal sorumluluk için ISO 26000 ve değer temelli tasarım için IEEE 7000-2021 bulunmaktadır (ISO, 2010; IEEE, 2021). Bu çalışmalarda disiplinler arası işbirliği, iş analistleri, müfredat tasarımcıları/eğitim bilimciler, politika yapıcılar, kamu yönetimi uzmanları, sosyologlar ve etik uzmanlarını içermektedir. Örneğin, bir sosyolog eğitim alanındaki bir YZ tasarımını toplumsal ihtiyaçlarla uyumlu hale getirmek için müfredat tasarımcıları, politika yapıcılar ve eğitimciler ile işbirliği yaparak çalışmalarını sürdürebilecektir.

İş/görev analizinde sosyal bilimcinin rolü, YZ sisteminin amaçlanan uygulamasının etik ve toplumsal hedeflerle uyumlu olmasını sağlama konusunda ortaya çıkmaktadır. Bu süreç, sistemin neden oluşturulduğunu ve ne amaçla çalıştığını tanımlamaktadır. Görevler (i) iş/görev hedeflerini ortaya çıkarmayı ve bunları etik bir bakış açısıyla yorumlamayı, (ii) etik

kısıtlamaları erken aşamada belirlemeyi, iş/görev için performansın ötesinde bir sosyo-etik başarı kriterleri formüle etmeyi içermektedir. Örneğin, eğitim veya değerlendirmede YZ kullanımının önyargı yaratmadan öğrenmeyi nasıl teşvik ettiğini açıklığa kavuşturmak ve iş/görevin kullanıcı gizliliğine ve erişim eşitliğine saygı gösterdiğinden ve toplumsal eşitsizlikleri artırmadığından emin olmak gibi içerikleri bulunmaktadır.

Burada yöntemleri güncellemek, görev planlamasına sosyo-etik etki analizi eklemek anlamına gelmektedir. Dolayısıyla iş modeli şablonlarının, değerler ve toplumsal faydalar/zararlar ile ilgili bölümler eklenerek revize edilmesi önemli görülmektedir. Standartlar açısından, kurumsal sosyal sorumluluk çerçeveleri (örneğin ISO 26000, ISO14001 gibi) ISO 5338'in görev analizi ile birlikte kullanılarak toplumsal değerleri yakalayabilecektir.

Önerilen yöntemsel yaklaşımlar arasında, etik senaryolarla zenginleştirilmiş iş/görev analizi yöntemlerinin kullanılması yer almaktadır. Örneğin, YZ sisteminin istenmeyen kullanımlarını öngörmek için etik uzmanları ve son kullanıcılarla senaryo planlaması yapılabilecektir. Bu amaçla RAI Toolkit, değerleri ortaya çıkarmak için paydaşların hedeflerinin haritalandırılmasını yansıtan *Ethics Canvas* gibi araçlar veya kolaylaştırılmış atölye çalışmaları önermektedir. Uygulamada, etik kontrol listeleri ve kılavuz ilkelerle desteklenen yapılandırılmış iş analizi çerçeveleri (örneğin değer zinciri analizi gibi) kullanarak, misyon analizi sürecinin etik kriterleri somut tasarım gereksinimleri olarak kodlaması sağlanmaktadır.

2.1.5. Paydaş İhtiyaçları ve Gereksinimleri Tanımlama Süreci

Paydaş ihtiyaçlarının ve gereksinimlerin tanımlanması sürecinde sosyal bilim araştırmacıları, paydaşların belirlenmesi, etik gereksinimlerin tespit edilmesi ve teknik ve toplumsal beklentiler arasında uyumlaştırma yapılması gibi rolleri üstlenmektedirler (ISO/IEC 5338 - Bölüm 6.4.2). Bu açıdan yöntem güncellemeleri, kapsayıcı tasarım ve katılımcı yaklaşımlar gerektirmektedir. İlgili standartlar arasında deneyim kalitesi için ISO/IEC 25059 ve etik gereksinimleri türetmek için IEEE 7000-2021 bulunmaktadır (ISO, 2022; IEEE, 2021). İşbirliği, ağırlıkla son kullanıcılar, sosyologlar, UX tasarımcıları, alan uzmanları, iletişim bilimciler ve hukuk danışmanlarını kapsamaktadır. Örneğin, anlaşılabilirlik gereksinimlerini test etmek için insan-bilgisayar etkileşimi uzmanlarıyla, veri işleme ihtiyaçlarını belirlemek için gizlilik alan uzmanlarıyla çalışabilecektir. RAI Toolkit, bu aşamada "*etik ilkeleri operasyonel hale getirmek*" için bunları gereksinimlere entegre ederek disiplinler arası ekip çalışmasını vurgulamaktadır.

Paydaş gereksinimleri süreci, farklı kullanıcıların ve paydaşların YZ sisteminden neye ihtiyaç duyduğunu tanımlamaktadır. Burada sosyal bilim araştırmacılarının katkısı aşağıda sıralanan alanları kapsamaktadır:

- Tüm ilgili paydaşların (savunmasız veya marjinal gruplar dâhil) belirlenmesi, paylaşılan endişelerinin ve değerlerinin ortaya çıkarılması,
- Gereksinim sürecinin adalet, gizlilik ve şeffaflık ihtiyaçlarını açıkça içermesinin sağlanması, örneğin, sistemin çıktılarının uzman olmayanlara da açıklanabilir olması,
- Teknik talepler ile toplumsal beklentiler arasında arabuluculuk ve/veya düzenleyici rol üstlenilmesi, etik değerlerin ölçülebilir gerekliliklere dönüştürülmesi.

Önerilen yöntemler etik filtrelerle desteklenmiş standart gereksinim mühendisliği araçlarını içermektedir. Bu yöntemler arasında kapsayıcı tasarım yaklaşımlarının gözden geçirilmesi, etik odak grupları ve paydaş etki değerlendirmeleri yer almaktadır (Shilton, 2018). RAI Toolkit, yakalanan gereksinimlerin etik hususları yansıtmasını sağlamak için algoritmik etki değerlendirmeleri veya paydaş etki çerçeveleri kullanılmasını önermektedir.

Kısacası, tamamen sorumlu bir gereksinim spesifikasyonu oluşturmak için etik kontrol listeleri ve adalet kriterleriyle desteklenmiş resmi veri toplama protokolleri (örneğin görüşmeler, anketler) kullanılmaktadır. Sosyal bilim uzmanları, bu süreçlerden elde edilen verilerin toplumsal temsiliyet, etik uygunluk ve paydaşlar arası değer dengesi açısından değerlendirilmesine katkı sağlamakta ve böylelikle, gereksinimlerin yalnızca teknik açıdan değil, toplumsal sorumluluk ilkeleriyle uyumlu biçimde tanımlanması mümkün hale gelmektedir.

2.1.6. Bilgi Edinme Süreci

Bilgi edinme sürecindeki (ISO/IEC 5338 – Bölüm 6.4.7) görevler, etik bir şekilde alan bilgisi edinmeyi, kaynakların çeşitliliğini sağlamayı ve bilişim alan uzmanlarıyla işbirliği yapmayı içermektedir. Yöntemlerin güncellenmesi, veri kaynaklarının etik incelemesini ve önyargı denetimlerini gerektirmektedir. İlgili standartlar arasında veri kalitesi için ISO/IEC 8000 ve YZ kullanım örnekleri için ISO/TR 24030 bulunmaktadır (ISO, 2019; ISO, 2023). İşbirliği, bilişim alan uzmanları, veri bilimi uzmanları, dilbilimciler, içerik analizi uzmanları ve sosyologları içermektedir. Önerilen yöntemler arasında, bilgi kodlama sırasında kaynak izleme ve etik denetim içeren bilgi grafikleri bulunmaktadır (Binns, 2018).

Bilgi edinme, YZ için gerekli alan bilgisinin (örneğin kurallar, uzman görüşleri, arka plan verileri gibi) elde edilmesi ve kodlanmasını içermektedir. Bu süreçte, sosyal bilim araştırmacılarının katkısı, kavramların doğru tanımlanmasını ve tutarlı biçimde kodlanmasını sağlayabilecektir. Örneğin, 'veri' kavramı farklı disiplinlerde farklı anlamlara gelebilmekte; sosyal bilimci, bu kavramın bağlamını sabitleyerek bilgi edinme sürecini güçlendirmektedir. Sosyal bilimcilerin katkısı şu şekilde özetlenebilir:

- YZ'nin nasıl kullanılacağı ile ilgili toplumsal, kültürel normlar veya bağlamsal faktörler gibi dâhil edilecek etik bilgilerin belirlenmesi ve
- Önyargılı veya eksik bilgi kaynaklarına karşı koruma sağlanması.

Bilgi kaynaklarının etik incelemesini standartlaştırarak yöntemlerin güncellenmesi bilgi edinme sürecine ilişkin en hayati konuların başında gelmektedir. Bu açıdan veri edinme sürecine veri kaynağı kontrolleri ve önyargı denetimleri eklenmelidir. Örneğin, metin verileri toplarken, nefret söylemi veya basmakalıp davranışların (*stereotiplerin*) kontrol edilmesi ve bunların etiketlenmesi sosyal bilimciler tarafından sürdürülebilecektir. Ayrıca bir toplumsal süreç hakkındaki bilgileri modellemek için, modelin kültürel farkları dikkate almasını sağlamak amacıyla sosyal bilim araştırmacıları ile birlikte çalışılması önem taşımaktadır.

Uygulamada, şeffaf bilgi toplama protokolleri kullanarak ve YZ denetim yazılımı gibi araçlardan yararlanarak, elde edilen bilginin tanımlanan etik gereklilikleri karşıladığından emin olunması mümkün hale gelmektedir.

2.1.7. İmha (Yok Etme) Süreci

İmha sürecindeki (ISO/IEC 5338 - Bölüm 6.4.17) görevler arasında veri saklama politikalarının tanımlanması, alınan/öğrenilen derslerin (lessons learned) belgelenmesi ve etik donanım ve yazılımın uygun şekilde ve zamanında imha edilmesinin sağlanması yer almaktadır. Yöntem güncellemeleri, etik ömür sonu kontrol listelerini içermektedir. İlgili standartlar arasında bilgi güvenliği için ISO/IEC 27001 ve medya temizliği için NIST SP 800-88 yer almaktadır (ISO, 2022; NIST, 2014). Bu süreçte işbirliği, bilgi güvenliği ve bilişim alan uzmanları, arşiv ve bilgi yönetimi alan uzmanları, hukuk/uyum alan uzmanları ve etik alan uzmanlarını içermektedir. Örneğin, veri silme işleminin onaylanması için siber güvenlik uzmanlarıyla, saklama sürelerinin düzenlemelere uygunluğunun sağlanması için hukuk alan uzmanlarıyla çalışılması önemli görülmektedir. Modeller hassas veriler içeriyorsa, veri gizliliği alan sorumlularıyla koordinasyon sağlanmalıdır. RAI Toolkit'in izlenebilirliğe verdiği önem, bu aşamada yönetişim ve denetim rollerinin dâhil edilmesini gerektirmektedir.

İmha süreci, bir YZ sisteminin (ve verilerinin/modellerinin) ömrü dolduğunda güvenli bir şekilde kullanımdan kaldırılmasını kapsamaktadır. Etik açıdan bu, veri gizliliğinin, çevresel sorumluluğun ve kullanım sonrası hesap verebilirliğin sağlanmasını içermektedir. Bu sürece sosyal bilim araştırmacıları, katılımcı haklarının korunup korunmadığının gözetilmesi, verilerin etik saklama veya imha yöntemlerine uyumunun sağlanması açısından katkı sağlayabilecektir. Böylelikle kullanımın sona ermesinin ardından mevcut verilerin kötüye kullanım riski en aza indirilmektedir. Sosyal bilimciler ayrıca veri saklama politikalarının tanımlanması, öğrenilen derslerin belgelenmesi ve bilgilerin etik bir şekilde korunması ve kötüye kullanımın engellenmesi, donanım ve yazılımın yönetmeliklere uygun olarak imha edilmesinin sağlanması konularına da destek olabilmektedir.

Önerilen yöntemler arasında güvenli silme protokolleri, geri dönüşüm prosedürleri ve imha doğrulama kontrol listeleri oluşturulması bulunmaktadır (Crawford & Calo, 2016) ve bu süreçte resmî imha protokollerinin kullanılması da gerekli görülmektedir. Sertifikalı veri silme yazılımı, donanım geri dönüşüm hizmetleri ve imha işlemlerini kaydetmek için belgeleme şablonları gibi araçlar kullanılmalıdır. RAI Toolkit yaklaşımı, projeyi sonlandırmadan önce tüm etik yükümlülüklerin (örneğin veri silme, paydaşlara bildirim, çevresel kılavuzlar gibi) yerine getirildiğini doğrulayan bir hizmet dışı bırakma kontrol listesini takip etmektir.

2.1.8. Kaynakça

Binns, R. (2018). Fairness in machine learning: Lessons from political philosophy. *Proceedings of Machine Learning Research, 81*, 149–159.

Crawford, K., & Calo, R. (2016). There is a blind spot in YZ research. *Nature, 538*(7625), 311–313.

Floridi, L., Cowls, J., Beltrametti, M., Chatila, R., Chazerand, P., Dignum, V., ... & Vayena, E. (2018). AI4People—An ethical framework for a good YZ society: Opportunities, risks, principles, and recommendations. *Minds and Machines, 28*(4), 689–707.

IEEE. (2021). *IEEE 7000-2021 – Model process for addressing ethical concerns during system design*. Institute of Electrical and Electronics Engineers.

International Organization for Standardization. (2010). *ISO 26000:2010 Guidance on social responsibility*. ISO.

International Organization for Standardization. (2015). *ISO 9001:2015 Quality management systems – Requirements*. ISO.

International Organization for Standardization. (2018). *ISO 31000:2018 Risk management – Guidelines*. ISO.

International Organization for Standardization. (2019). *ISO/IEC 8000:2019 Data quality*. ISO.

International Organization for Standardization. (2022). *ISO/IEC 25059:2022 Quality of experience framework*. ISO.

International Organization for Standardization. (2023). *ISO/IEC 5338:2023 Artificial intelligence lifecycle processes*. ISO.

Jobin, A., Ienca, M., & Vayena, E. (2019). The global landscape of YZ ethics guidelines. *Nature Machine Intelligence, 1*(9), 389–399.

Mitchell, M., Wu, S., Zaldivar, A., Barnes, P., Vasserman, L., Hutchinson, B., ... & Gebru, T. (2019). Model cards for model reporting. *Proceedings of the Conference on Fairness, Accountability, and Transparency*, 220–229.

Morley, J., Floridi, L., Kinsey, L., & Elhalal, A. (2021). From what to how: An overview of YZ ethics tools, methods and research to translate principles into practices. *Science and Engineering Ethics, 27*(4), 1–25.

National Institute of Standards and Technology. (2014). *NIST Special Publication 800-88: Guidelines for media sanitization*. NIST.

Shilton, K. (2018). Values and ethics in human-computer interaction. *Foundations and Trends in Human–Computer Interaction, 12*(2), 107–171.

Slagg, A. (2025). Building Responsible YZ for Agencies. FedTech, *https://fedtechmagazine.com/article/2025/04/dod-responsible-ai-rai-toolkit-perfcon#:~:text=%E2%80%9CEthics%20in%20AI%20is%20often,ensuring%20human%20oversight%20in%20specific* (20/08/2025).

U.S. Department of Defense. (2022). *Responsible Artificial Intelligence (RAI) toolkit*. DoD.

2.2. Yapay Zekâ Etiği ve Uluslararası Standartlar

YZ teknolojileri hayatımızın her alanına hızla entegre olurken, bu ilerlemenin getirdiği etik sorumluluklar da giderek daha fazla önem kazanmaktadır. YZ sistemlerinin sadece teknik olarak başarılı olması yeterli görülmemekte; aynı zamanda adil, güvenilir, şeffaf ve insan odaklı olması da önemli görülmektedir. Bu nedenle, YZ'nin sorumlu gelişimi ve kullanımı için uluslararası standartlar ve belirli etik ilkeler ortaya konulmuştur. Aşağıdaki tablo, YZ sistemlerinin yaşam döngüsünü düzenleyen ISO/IEC 5338 gibi teknik standartların yanı sıra, YZ etiğini doğrudan ele alan ISO/IEC 22989, ISO/IEC 42001, ISO/IEC AWI 25590 ve

ISO/IEC AWI TR 42103 gibi diğer önemli uluslararası standartların ve yaklaşımların, temel etik ilkelerle nasıl bağlantılı olduğunu detaylandırmaktadır.

YZ Etiği ve Uluslararası Standartlar: İlişkilendirme Tablosu

Standart Adı	Etik İlke
ISO/IEC 22989:2022	Şeffaflık (Transparency)
	Açıklanabilirlik (Explainability)
	Sağlamlık (Robustness)
	Risk Yönetimi (Risk Management)
ISO/IEC 42001:2023	Sorumlu YZ Kullanımı (Responsible YZ Use)
	Veri Yönetişimi (Data Governance)
	İnsan Gözetimi (Human Oversight)
	Risk Bazlı Yaklaşım (Risk-Based Approach)
IEEE 7000 Serisi Standartlar	Değer Odaklı Mühendislik (Value-based Engineering - VBE)
	Önyargı Azaltma (Bias Mitigation)
	Hesap Verebilirlik (Accountability)
ISO/IEC AWI 25590	Veri Kalitesi (Output Data Quality)

2.2.1. YZ Etiğinin Temel İlkeleri

YZ etiği, teknik gelişmeleri yönlendiren ve YZ'nin topluma faydalı olmasını sağlayan temel prensipler bütünü olarak tanımlanabilecektir. Anılan tabloda adı geçen her bir etik ilkenin kısa açıklamalarına aşağıda yer verilmektedir:

- **Şeffaflık (Transparency)** Bir YZ sisteminin nasıl çalıştığı, ne tür verilerle eğitildiği ve kararlarını nasıl aldığı konusunda açık ve anlaşılır olunması anlamına gelmektedir. Bu 'kara kutu' sistemlerden kaçınmayı veya en azından iç işleyişinin belirli düzeylerde açıklanabilmesini hedeflemektedir. Şeffaflık, güven inşa etmek ve potansiyel sorunları erken tespit etmek için kritik önem taşımaktadır.
- **Açıklanabilirlik (Explainability)** YZ sisteminin belirli bir karara veya sonuca neden ulaştığının insan tarafından anlaşılabilecek bir şekilde izah edilebilmesidir. Özellikle YZ'nin tıbbi teşhis, adli kararlar, kredi onayı gibi tıbbi, hukuki ve/veya ekonomik anlamda alınacak son derece kritik kararlarda kullanıldığı durumlarda,

sistemin neden o kararı verdiğini bilmek hem hesap verebilirlik hem de bireylerin hakları açısından temel bir gerekliliktir.

- **Sağlamlık (Robustness)** Bir YZ sisteminin, beklenmedik veya alışılmadık girdilere, veri gürültüsüne veya kötü niyetli saldırılara rağmen doğru ve tutarlı performansını sürdürebilme yeteneğidir. Sağlam bir YZ sistemi, güvenilirliği ve güvenliği artırarak, yanlış veya tehlikeli sonuçlar üretme riskini azaltacaktır.
- **Risk Yönetimi (Risk Management)** YZ sistemleriyle ilişkili olası zararları (örneğin önyargı, gizlilik ihlali, güvenlik açıkları gibi) proaktif olarak tanımlama, değerlendirme ve azaltma sürecidir. Bu ilke, YZ'nin potansiyel olumsuz etkilerinin öngörülmesini ve bunlara karşı önlemler alınmasını sağlayarak, etik ihlallerin önüne geçmeyi amaçlamaktadır.
- **Sorumlu YZ Kullanımı (Responsible YZ Use)** YZ teknolojilerinin geliştirilmesi, dağıtımı ve kullanımının; yasalara, etik ilkelere ve toplumsal değerlere uygun, hesap verebilir ve denetlenebilir bir şekilde yapılmasıdır. Bu ilke, YZ'nin insan haklarına saygı duymasını, olumsuz toplumsal etkilerden kaçınmasını ve genel refaha katkıda bulunmasını sağlamak için bir şemsiye görevi görmektedir.
- **Veri Yönetişimi (Data Governance)** YZ sistemlerinde kullanılan verilerin toplanması, depolanması, işlenmesi ve korunmasına yönelik kapsamlı kurallar ve süreçler bütünüdür. Veri kalitesi, gizliliği, güvenliği ve etik kullanımı, YZ'nin adil ve güvenilir sonuçlar üretmesi için temeldir. Çünkü önyargılı veya güvenli olmayan veriler, etik olmayan YZ davranışlarına yol açabilme riskini doğurmaktadır.
- **İnsan Gözetimi (Human Oversight)** YZ sistemlerinin tamamen otonom olmaktan ziyade, insanlar tarafından denetlenebilmesi, yönetilebilmesi ve gerektiğinde müdahale edilebilmesi gerekliliğidir. Bu ilke, insan otonomisini korumayı, YZ hatalarının düzeltilmesini sağlamayı ve nihai kararların ve sorumluluğun insanlarda kalmasını güvence altına almayı amaçlamaktadır.
- **Risk Bazlı Yaklaşım (Risk-Based Approach)** YZ sistemlerinin potansiyel zararlarının seviyesine göre farklı seviyelerde denetim, doğrulama ve önlem stratejisi belirlenmesine yönelik bir gerekliliktir. Daha yüksek riskli YZ uygulamaları (örneğin sağlık, güvenlik gibi) daha sıkı etik ve güvenlik kontrollerine

tabi olmalıdır. Bu yaklaşım, kaynakları en çok ihtiyaç duyulan alanlara yönlendirerek etik riskleri optimize etmeye yardımcı olmayı önceliklendirmektedir.

- **Değer Odaklı Mühendislik (Value-based Engineering - VBE)** YZ sistemlerinin tasarım ve geliştirme süreçlerinin en başından itibaren etik değerleri (örneğin adalet, gizlilik, özerklik gibi) sistem gereksinimlerine dönüştürmeyi ve bunları teknik implementasyona dâhil etmeyi hedeflemektedir. Bu, etik kaygıların döngünün ayrılmaz bir parçası olmasını sağlamaktadır.
- **Önyargı Azaltma (Bias Mitigation)** YZ algoritmalarında veya eğitim verilerinde var olabilecek 'belli gruplara karşı ayrımcı eğilimler' gibi önyargıların tespit edilmesi, ölçülmesi ve azaltılması için uygulanan yöntemlerdir. Bu ilke, YZ sistemlerinin adil olmasını ve herhangi bir kişi veya grubun cinsiyet, ırk/etnik köken, yaş, inanç gibi özelliklerine dayalı ayrımcılığa maruz kalmamasını sağlamayı amaçlamaktadır.
- **Hesap Verebilirlik (Accountability)** YZ sistemlerinin kararları ve eylemlerinden kimin sorumlu olduğunun açıkça belirlenmesi ve gerektiğinde bu sorumluluğun yerine getirilmesidir. Bir YZ sisteminin neden olduğu bir zarar veya hata durumunda geliştirici, dağıtıcı, kullanıcı olmak üzere ilgili tarafların hesap verebilmesi ve gerekli düzeltmeleri yapabilmesi gerekmektedir. Bu ilke, şeffaflık ve açıklanabilirlik ilkesiyle yakından ilişkilidir.
- **Veri Kalitesi (Output Data Quality)** Üretken YZ (Generative AI) uygulamalarının ürettiği metin, görsel, ses gibi içeriklerin doğruluğu, tutarlılığı, güvenilirliğinin sağlanması ile potansiyel zararlı veya yanıltıcı içerik üretilmesinin önlenmesi gibi konuların yönetilmesidir.
- **Veri Güvenliği ve Gizliliği (Data Security & Privacy)** YZ sistemleri bağlamında sentetik verilerin kullanımı da dâhil olmak üzere, verilerin yetkisiz erişim, kullanım, ifşa, değişiklik veya yok edilmeye karşı korunmasıdır. Bu ilke, bireylerin kişisel bilgilerinin korunmasını ve YZ sistemlerinin veri işleme süreçlerinin yasal ve etik standartlara uygun olmasını sağlamayı amaçlamaktadır.

2.2.2. Etikle İlgili Sonuç

YZ sistemleri hayatımızın her alanına derinlemesine nüfuz etmeye devam ederken, bu teknolojilerin sorumlu, adil ve insan odaklı bir şekilde geliştirilmesi ve kullanılması hayati bir

öneme sahiptir. ISO/IEC 5338 gibi teknik standartlar YZ sistemlerinin yaşam döngüsü süreçlerini yapılandırarak güvenli ve yönetilebilir bir temel sağlarken, YZ etiğini doğrudan ele alan ISO/IEC 42001, ISO/IEC 22989 ve IEEE 7000 Serisi gibi diğer uluslararası standartlar, bu teknik çerçevenin etik prensiplerle nasıl birleştirileceğini göstermektedir.

Bu standartlar, şeffaflık, açıklanabilirlik, sağlamlık, adalet, veri yönetişimi, insan gözetimi ve hesap verebilirlik gibi temel etik ilkelerin YZ sistemlerinin tasarımından dağıtımına kadar her aşamasına entegre edilmesini sağlamayı amaçlar. Bu arada amaç, YZ'nin önyargı, ayrımcılık, gizlilik ihlalleri gibi potansiyel risklerini en aza indirirken, insan haklarına saygı duyan ve toplumsal refahı artıran bir teknoloji olmasını sağlamaktır.

YZ'nin hızla gelişen doğası göz önüne alındığında, bu etik ilkelerin ve standartların sürekli olarak gözden geçirilmesi ve güncellenmesi gerekmektedir. YZ'nin getirdiği karmaşık etik sorunlara karşı proaktif bir yaklaşım benimsemek, yalnızca yasal uyumluluğu sağlamakla kalmayacak, aynı zamanda YZ'ye olan toplumsal güveni artıracak ve bu güçlü teknolojinin faydalarını herkes için en üst düzeye çıkaracaktır. Nihayetinde, YZ'nin geleceği, onu yaratıcı ve yenilikçi bir şekilde kullanma yeteneğimiz kadar, etik değerlere ne kadar sıkı sarıldığımıza da bağlı olacaktır.

2.3. YZ Yönetişiminde Etik İlkeler, Normatif Çerçeveler, Sonuç ve Politika Önerileri

YZ, karar alma süreçlerini, hizmet sunumunu ve toplumsal etkileşimleri etkileyerek hem kamu hem de özel sektör faaliyetlerine giderek daha fazla entegre olmaktadır. YZ, verimlilik, inovasyon ve erişilebilirliğin artırılması için fırsatlar sunarken, aynı zamanda etik, toplumsal ve yönetişimle ilgili endişeleri de gündeme getirmektedir. Bu çalışmada, (i) kamu ve özel sektörde YZ için etik ilkeler ve hesap verebilirlik mekanizmalarının oluşturulması ve (ii) özellikle kamuda YZ'nin benimsenmesi için şeffaflık ve güvenilirlik standartlarının tanımlanması olmak üzere iki temel alanda öneriler sunulmaktadır.

2.3.1. Etik İlkelerin ve Hesap Verebilirlik Mekanizmalarının Oluşturulması

1. YZ için Etik İlkeler

Etik ilkelerin ve hesap verebilirlik mekanizmalarının oluşturulması hem kamu hem de özel sektörde YZ'nin sorumlu bir şekilde geliştirilmesi ve kullanılması için temel öneme sahiptir. Burada başvurulacak ilkelerin, OECD YZ İlkeleri (2019), AB YZ Etik Kılavuzları Üst

Düzey Uzman Grubu (2019) ve UNESCO YZ Etiği Tavsiye Kararı (2021) gibi uluslararası kabul görmüş çerçevelerle uyumlu olması önemli görülmektedir.

Bahsi geçen çerçevelerin temelinde, YZ sistemlerinin insan haklarına, insan onuruna ve toplumsal refaha saygı duymasını sağlayan "*insan merkezli tasarımı*" yatmaktadır (Jobin ve ark., 2019). Etik yönetişim ayrıca adalet ve ayrımcılık yapmamayı gerektirmekte, bu da mevcut yapısal eşitsizliklerin yeniden üretilmesini veya artmasını önlemek anlamına gelmektedir (Barocas ve ark., 2023). Gizlilik ve veri koruma da aynı derecede önemlidir ve YZ uygulamalarının kişisel bilgileri koruması ve veri yönetimi düzenlemelerine tam olarak uyması gerekmektedir (Floridi & Cowls, 2022). Bir diğer önemli unsur ise hesap verebilirlik ve açıklanabilirliktir. Sistemlerin denetlenebilir olması ayrıca tasarım, uygulama ve denetim sorumlulukları açıkça tanımlanabilir olması önem taşımaktadır (Cath, 2018). Bunlara ek olarak, YZ'nin, sürdürülebilirlik ve sosyal sorumluluk ilkeleri tarafından yönlendirilmesine özen gösterilmesi, çevresel zararı en aza indirirken toplumsal kalkınmaya da olumlu katkıda bulunmasının gözetilmesi önemli görülmektedir (UNESCO, 2021).

Kısaca hem kamu kurum ve kuruluşlarının hem de özel kuruluşların OECD YZ İlkeleri (2019), AB YZ Etik Kılavuzları Üst Düzey Uzman Grubu (2019) ve UNESCO YZ Etiği Tavsiye Kararı (2021) gibi uluslararası çerçevelere uygun etik ilkeleri benimseme konusunda teşvik edilmesi önemli görülmektedir.

2. Hesap Verebilirlik Mekanizmaları

İlkeleri uygulamaya geçirmek, iyi tanımlanmış hesap verebilirlik mekanizmaları gerektirmektedir. Etik komiteleri veya sektörler arası temsilcilerin bulunduğu düzenleyici kurumlar gibi bağımsız denetim organları, etik standartlara uyumu sağlamaya yardımcı olabilecektir (Winfield & Jirotka, 2018). Ayrıca, uygulama öncesinde zorunlu YZ Etki Değerlendirmelerinin (Artificial Intelligence Impact Assessment-AIIA) gerçekleştirilmesi ve adalet, önyargı, güvenlik ve insan hakları gibi alanlarda potansiyel risklerin değerlendirilmesi önemlidir (OECD, 2019). Sistemler faaliyete geçtikten sonra, ortaya çıkan riskleri belirlemek ve etik uyumu sürdürmek için, tercihen bağımsız üçüncü taraf denetçiler tarafından yürütülen denetim ve izleme yapıları kurulması da ayrıca gerekli görülmektedir (Floridi, 2019).

Hesap verebilirlik, özellikle haklarını etkileyen bağlamlarda, bireylere YZ destekli kararlara itiraz etmek için açık ve erişilebilir yollar sağlayan sağlam düzeltme mekanizmaları da gerektirmektedir (Wachter & Mittelstadt, 2019). Bu kurumsal önlemlerin yanı sıra, YZ yaşam döngüsünün tamamında geliştiriciler, uygulayıcılar ve karar vericiler için mesleki

sorumluluk açıkça tanımlanmalı ve etik taahhütlerin hem teknik hem de örgütsel uygulamalara dâhil edilmesi sağlanmalıdır (Avrupa Komisyonu, 2021).

Bu şekilde, etik ilkeler sadece değerler düzeyinde ifade edilmez, aynı zamanda yönetişim yapıları ve mesleki sorumluluklar aracılığıyla da işlevselleştirilmekte ve güvenilir YZ için kapsamlı bir çerçeve oluşturulmaktadır.

2.3.2. Kamuda YZ için Şeffaflık ve Güvenilirlik Standartlarının Belirlenmesi

1. Şeffaflık Standartları

Kamuda, özellikle vatandaşları doğrudan etkileyen karar alma süreçlerinde YZ'nin benimsenmesi, şeffaflık ve güvenilirliğe daha fazla dikkat edilmesini gerektirmektedir. Şeffaflık, hesap verebilirlik için olduğu kadar kamu kurumları ile toplum arasında güvenin tesis edilmesi için de gerekli görülmektedir. Bunu başarmak için devletler, karar alma süreçlerinde YZ'nin kullanımının açıkça belirtilmesini sağlayarak vatandaşların algoritmik sistemlerin ne zaman ve nasıl kullanıldığını bilmelerine olanak tanımalıdır (OECD, 2019). Ancak, açıklık tek başına yeterli değildir. Vatandaşlar, YZ odaklı sonuçların ardındaki mantığı da anlayabilmelidir. Bu, hesap verebilirliği belirsizleştiren aşırı teknik ifadelerden kaçınarak, sonuçları erişilebilir ve anlaşılır bir dille aktaran açıklanabilirlik uygulamaları gerektirmektedir (Doshi-Velez & Kim, 2017).

Ayrıca, kamu kurum ve kuruluşları, kullandıkları veri setleri, algoritmalar ve karar alma çerçeveleri ile ilgili teknik ve etik belgeleri yayınlamayı taahhüt etmelidir. Güvenlik ve gizlilik ilkeleriyle uyumlu olarak uygulandığında bu tür bir açıklık, YZ uygulamalarının varsayımlarını ve sınırlamalarını incelemeyi mümkün kılacaktır (Brundage et al., 2020). Açıklama ve belgelemenin ötesinde, şeffaflık aynı zamanda katılımı da içermektedir. Bu doğrultuda sivil toplum kuruluşları ve akademik paydaşların YZ sistemlerinin yönetişimini gözden geçirmelerine, eleştirmelerine ve katkıda bulunmalarına olanak tanıyan mekanizmalar oluşturulmalı ve böylece daha demokratik ve kapsayıcı bir denetim süreci yaratılmalıdır (Zerilli et al., 2019).

2. Güvenilirlik Standartları

Güvenilirlik, kamuda sorumlu YZ kullanımı için ikinci temel dayanak noktasıdır. Vatandaşlar, kırılgan veya hataya açık sistemlere güvenme eğiliminde değildir ve güvenilmez araçlar, kamu kurum ve kuruluşlarına olan güveni hızla zedeleyebilecektir. Bu nedenle, YZ

uygulamaları, hatalara, kötü niyetli saldırılara veya veri manipülasyonuna dayanabileceğinden emin olmak için sıkı sağlamlık testlerinden geçmelidir (Goodfellow et al., 2015). Güvenilirlik ayrıca, YZ yaşam döngüsü süreçleri için ISO/IEC 5338 ve güvenilirlik için ISO/IEC 24028'de özetlenenler gibi standartlaştırılmış değerlendirme çerçeveleri ve kalite güvence protokollerine de bağlıdır (ISO/IEC, 2023).

Güvenilirliğin bir diğer önemli unsuru, veri kümelerinde ve sistem çıktılarında önyargının sürekli olarak izlenmesi ve azaltılmasıdır. Önyargının azaltılması tek seferlik bir müdahale olarak ele alınmamalı, bunun aksine YZ sistemlerinin sürekli denetimine entegre edilmelidir (Mehrabi ve ark., 2021). Son olarak, sağlık, adalet veya sosyal hizmetler gibi kamu alanlarında kullanılan YZ'ler, mevcut kurumsal altyapılarla birlikte çalışabilirliği ve güvenlik risklerine karşı güçlü korumaları sağlamalıdır. Bu ve benzeri konularda önlemlerin alınmaması halinde, kritik kamu çalışmalarında YZ'nin etkinliği ve meşruiyeti ciddi şekilde zedelenebilecektir (Avrupa Komisyonu, 2021).

Özetle kamu kurum ve kuruluşları YZ yönetişimine şeffaflık ve güvenilirlik standartlarını dâhil ederek, hem hesap verebilirliği hem de güveni güçlendirebilecek ayrıca teknolojik yeniliklerin adil, etkili ve demokratik hizmet sunumuna katkıda bulunmasını sağlayabilecektir.

2.3.3. Sonuç ve Politika Önerileri

Etik hesap verebilirlik ve şeffaf güvenilirlik standartlarına olan ikili ihtiyaç, YZ için kapsamlı bir yönetişim çerçevesinin gerekliliğini vurgulamaktadır. Özel sektör inovasyonu yönlendirirken, kamu kurum ve kuruluşları insan haklarını korumak ve güveni sağlamak konusunda tartışmasız ve benzersiz sorumluluklar taşımaktadır.

İlgili literatürden yola çıkılarak oluşturulan ve öne çıkan öneriler şu şekilde sunulabilir:

- Küresel normlarla uyumlu, kamu ve özel kurum ve kuruluşlar genelinde bütünleşik/entegre etik ilkelerin benimsenmesi,
- Denetimler, gözetim organları ve telafi sistemleri aracılığıyla hesap verebilirlik mekanizmalarının kurumsallaştırılması,
- Kamuda YZ kullanımında şeffaflığın, açıklanabilirlik ve katılımcı yönetişim yoluyla zorunlu hale getirilmesi,

- Sıkı testler, ISO tabanlı kalite çerçeveleri ve önyargı azaltma protokolleri yoluyla güvenilirlik standartlarının uygulanması,
- YZ yönetişiminde birlikte çalışabilirliği ve küresel güveni sağlamak için standartların uluslararası uyumlaştırılmasının teşvik edilmesi.

Etik, hesap verebilirlik, şeffaflık ve güvenilirliği birleştiren dengeli bir yaklaşım, YZ'nin kamu yararına hizmet ederken özel sektörde inovasyonu teşvik etmesine yardımcı olacaktır.

2.3.4. Kaynakça

Barocas, S., Hardt, M., & Narayanan, A. (2023). *Fairness and machine learning*. MIT Press.

Brundage, M., Avin, S., Clark, J., Toner, H., Eckersley, P., Garfinkel, B., … & Amodei, D. (2020). Toward trustworthy YZ development: Mechanisms for supporting verifiable claims. *arXiv preprint arXiv:2004.07213*.

Cath, C. (2018). Governing artificial intelligence: Ethical, legal and technical opportunities and challenges. *Philosophical Transactions of the Royal Society A, 376*(2133), 20180080. https://doi.org/10.1098/rsta.2018.0080

Doshi-Velez, F., & Kim, B. (2017). Towards a rigorous science of interpretable machine learning. *arXiv preprint arXiv:1702.08608*.

European Commission. (2021). *Proposal for a regulation laying down harmonised rules on artificial intelligence (Artificial Intelligence Act)*. COM/2021/206 final. https://eur-lex.europa.eu/legal-content/EN/TXT/?uri=CELEX:52021PC0206

Floridi, L. (2019). Establishing the rules for building trustworthy AI. *Nature Machine Intelligence, 1*(6), 261–262. https://doi.org/10.1038/s42256-019-0055-y

Floridi, L., & Cowls, J. (2022). A unified framework of five principles for YZ in society. *Harvard Data Science Review, 4*(1). https://doi.org/10.1162/99608f92.8cd550d1

Goodfellow, I., Shlens, J., & Szegedy, C. (2015). Explaining and harnessing adversarial examples. *arXiv preprint arXiv:1412.6572*.

ISO/IEC. (2023). *ISO/IEC 5338: Information technology — Artificial intelligence — YZ system life cycle processes*. International Organization for Standardization.

Jobin, A., Ienca, M., & Vayena, E. (2019). The global landscape of YZ ethics guidelines. *Nature Machine Intelligence, 1*(9), 389–399. https://doi.org/10.1038/s42256-019-0088-2

Mehrabi, N., Morstatter, F., Saxena, N., Lerman, K., & Galstyan, A. (2021). A survey on bias and fairness in machine learning. *ACM Computing Surveys, 54*(6), 115:1–115:35. https://doi.org/10.1145/3457607

OECD. (2019). *OECD principles on artificial intelligence*. OECD Publishing.

UNESCO. (2021). *Recommendation on the ethics of artificial intelligence*. UNESCO Publishing.

Wachter, S., & Mittelstadt, B. (2019). A right to reasonable inferences: Re-thinking data protection law in the age of big data and AI. *Columbia Business Law Review, 2019*(2), 494–620.

Winfield, A. F. T., & Jirotka, M. (2018). Ethical governance is essential to building trust in robotics and artificial intelligence systems. *Philosophical Transactions of the Royal Society A, 376*(2133), 20180085. https://doi.org/10.1098/rsta.2018.0085

Zerilli, J., Knott, A., Maclaurin, J., & Gavaghan, C. (2019). Transparency in algorithmic and human decision-making: Is there a double standard? *Philosophy & Technology, 32*(4), 661–683. https://doi.org/10.1007/s13347-018-0330-6.

2.4. Etikle İlişkili Bazı Standartlar, Teknik Raporlar ve Teknik Şartnameler

2.4.1. ISO/IEC TR 24368:2022 Teknik Raporu ve ISO/IEC AWI TS 25571 Teknik Şartnamesi

1. Giriş

Bu bölüm, YZ etiği ve yönetişimi alanında iki kritik ISO/IEC standardı olan ISO/IEC TR 24368:2022 ve ISO/IEC AWI TS 25571'e ilişkin açık kaynaklı bilgileri kapsamlı bir şekilde incelemektedir. ISO/IEC TR 24368:2022, YZ'nin etik ve toplumsal endişelerine dair yüksek düzeyli bir genel bakış sunarak, bu alandaki ilk uluslararası tanınmış standart olarak öne çıkmaktadır. Temel ilkeleri ve süreçleri ana hatlarıyla belirleyerek, YZ'nin sorumlu gelişimi

için kavramsal bir çerçeve sağlamaktadır. Buna karşılık, geliştirme aşamasındaki ISO/IEC AWI TS 25571, YZ sistem yaşam döngüsünün her aşamasında potansiyel etik sorunları belgelemek için pratik bir şablon sunmayı amaçlamaktadır. Bu, soyut etik ilkelerden eyleme geçirilebilir uygulama araçlarına doğru stratejik bir ilerlemeyi temsil etmektedir.

Bu bölüm, bu iki belgenin YZ'nin güvenilir ve sorumlu bir şekilde geliştirilmesi ve dağıtılması için nasıl tamamlayıcı roller oynadığını detaylandırmaktadır. ISO/IEC TR 24368:2022, etik pusula görevi görürken, ISO/IEC AWI TS 25571, kuruluşların etik hususları operasyonel süreçlerine entegre etmeleri için somut bir mekanizma sunmaktadır. Uluslararası standardizasyonun bu çabaları, YZ'nin hızla gelişen ortamında güven oluşturmak, riskleri azaltmak ve küresel uyumu teşvik etmek için hayati önem taşımaktadır. Sunulan tüm bilgiler, yalnızca kamuya açık kaynaklardan elde edilmiş olup, belgenin tam metinlerine erişim kısıtlamaları gibi sınırlamalar açıkça belirtilmiştir.

2. Yapay Zekâ Standardizasyonuna ve Etik Zorunluluklara Giriş

YZ teknolojilerinin hızlı ilerlemesi, küresel toplum ve ekonomi üzerinde dönüştürücü bir etki yaratmıştır. İş dünyası ve sağlık hizmetlerinden ulaştırma ve siber güvenliğe kadar çok çeşitli sektörlerde yenilikleri teşvik eden YZ yeteneklerinde önemli bir artış gözlemlenmektedir. YZ sistemleri, genellikle görevleri bilgilendirerek, tavsiyelerde bulunarak veya basitleştirerek faydalı etkiler yaratmak için kullanılmaktadır. Bununla birlikte, bu teknolojilerin yaygınlaşması, kamu yararını korumaya yönelik artan bir odaklanmayı da beraberinde getirmiştir. Büyük ekonomilerin YZ girişimlerine yaptığı ağır yatırımlar, bu gelişmelerin önemini pekiştirmektedir.

Yapay Zekâ Risklerini Azaltmada Standardizasyonun Kritik Rolü

YZ sistemlerinin güvenilirliğini ve kalitesini sağlamak, YZ hatalarından kaynaklanan potansiyel riskleri azaltmak için hayati öneme sahiptir. 2019 tarihli bir IDC anketi, çoğu kuruluşun bazı projelerinde YZ hataları bildirdiğini ve dörtte birinin yüzde 50'ye varan bir başarısızlık oranına sahip olduğunu ortaya koymuştur. Bu tür başarısızlıklar, yeterli yazılım testinin ve dolayısıyla YZ standartlarının geliştirilmesi için güçlü bir motivasyon kaynağı olmuştur. YZ ürünlerinin üretimine yönelik gereksinimler, bu sistemlerin önemi ve yüksek maliyetleri nedeniyle giderek artan bir şekilde standartlar şeklinde sunulmaktadır. Standartlar, teknoloji geliştirme oranlarını ölçmek için önemli göstergeler haline gelmiş ve ulusal kültürler ile ürünlerin piyasaya sürülmesine yönelik resmi kurallardan etkilenmektedir. YZ ürünleri için

izleme çerçeveleri ve değerlendirme standartları oluşturma çabaları, bu zorluklara doğrudan bir yanıt niteliğindedir.

YZ standardizasyonunun arkasındaki itici güçler incelendiğinde, hem proaktif hem de reaktif bir dinamik olduğu ortaya çıkmaktadır. Bir yandan, YZ projelerindeki belgelenmiş yüksek başarısızlık oranları ve "potansiyel riskleri azaltma" ihtiyacı, mevcut sorunlara ve gözlemlenen zararlara karşı reaktif bir yanıtı düşündürmektedir. Standartlar, belirlenen eksiklikleri gidermek ve tekrarlanmasını önlemek için ortaya çıkmaktadır. Öte yandan, "kamu yararını koruma" ve etik sorunların proaktif olarak belirlenmesi (ISO/IEC AWI TS 25571'in amacında görüldüğü gibi) vurgusu, ileriye dönük, önleyici bir yaklaşımı işaret etmektedir. Bu, YZ gelişimini en başından itibaren sorumlu bir şekilde yönlendirmeye yönelik stratejik bir çabayı göstermektedir. Bu sürekli uyum, hızla gelişen bir teknolojiyi düzenlemenin karmaşıklığının altını çizmekte ve çevik ve duyarlı standardizasyon süreçlerinin gerekliliğini vurgulamaktadır.

ISO/IEC Ortak Teknik Komitesi 1 (JTC 1), YZ Alt Komitesi 42 (SC 42)'nin Rolü

ISO/IEC JTC 1/SC 42, ISO/IEC TR 24368:2022 de dahil olmak üzere YZ standartlarının geliştirilmesinden sorumlu teknik komitedir. Bu komite, YZ yönetişiminin küresel manzarasını şekillendirmede merkezi bir rol oynamaktadır. Standardizasyonun, "bu teknolojilerin dağıtımında güven ve itimat oluşturma" ile açıkça bağlantılı olduğu ve "müşteriler ve paydaşlarla güven oluşturmanın" etik YZ çerçevelerini benimsemenin temel bir faktörü olarak tanımlandığı görülmektedir. Bu, sadece teknik uyumluluğun ötesine geçerek, uluslararası standartların varlığının ve benimsenmesinin, geliştiriciler ve dağıtıcılar için güvenilirliğin ve etik taahhüdün bir sinyali olarak daha geniş bir kitleye (kullanıcılar, düzenleyiciler ve halk dahil) hizmet ettiğini göstermektedir. YZ'nin yaygın toplumsal kabul görmesi ve tam ekonomik potansiyelini ortaya çıkarması için güvenilir olarak algılanması gerekmektedir. Standartlar, özellikle etik kaygıları ele alanlar, YZ geliştiricileri/dağıtıcıları ile toplum arasındaki güven boşluğunu kapatmak için kritik mekanizmalar olarak hizmet etmektedir. Sorumlu uygulamaları göstermek için ortak, doğrulanabilir bir çerçeve sağlayarak pazar güvenini artırmakta ve sorumlu yeniliği hızlandırmaktadır.

3. ISO/IEC TR 24368:2022 – Yapay Zekâ'da Etik ve Toplumsal Endişelere Genel Bakış

ISO/IEC TR 24368:2022, YZ alanında etik ve toplumsal endişeleri ele alan temel bir belgedir. Bu Teknik Rapor, YZ'nin sorumlu bir şekilde geliştirilmesi ve dağıtılması için uluslararası bir referans noktası sağlamaktadır.

A. Resmi Tanım ve Yayın Detayları

ISO/IEC TR 24368:2022'nin tam başlığı "Bilgi Teknolojisi – Yapay Zekâ - Etik ve toplumsal endişelere genel bakış" şeklindedir. Belgenin yayın türü Teknik Rapor (TR) olarak belirtilmiştir. Bu tanımlama, belgenin uygunluk için zorunlu gereksinimler yerine bilgilendirici materyal, rehberlik veya en iyi uygulamaları sağladığını göstermektedir. Belge, ilk baskısı (1.0) olarak 19 Ağustos 2022 tarihinde yayınlanmıştır. ISO/IEC JTC 1/SC 42 YZ Teknik Komitesi tarafından geliştirilmiş olup uluslararası YZ standardizasyonundaki yetkili kökenini teyit etmektedir. Belge 48 sayfadan oluşmakta ve 2.44 MB dosya boyutuna sahiptir.

ICS 35.020 (Genel olarak bilgi teknolojisi) altında sınıflandırılmıştır. Belge, ISO/IEC 2022 tarafından telif hakkı ile korunmaktadır ve tüm hakları saklıdır. Önceden yazılı izin olmaksızın çoğaltılması veya kullanılması yasaktır. Kamu erişimi, önizleme ve özetlerle sınırlıdır. Tam yayının satın alınması gerekmektedir (örneğin, çok kullanıcılı bir kopya için 177 CHF).

Bazı PDF dosyaları, birden fazla bilgisayarda açılmayı veya yazdırma/kopyalama/ yapıştırma işlevlerini sınırlayabilen *Dijital Haklar Yönetimi* (DRM) ile korunmaktadır. Bir belgenin "Uluslararası Standart" (IS) yerine "Teknik Rapor" (TR) olarak sınıflandırılması bilinçli bir stratejik tercihtir. TR'ler genellikle bilgilendirici olup rehberlik sağlarken, Uluslararası Standart (IS) belgeleri uygunluk için gereksinimler belirlemektedir. Bu belge, açıkça "yüksek düzeyli bir genel bakış" sağladığını ve "bu alandaki ilke, süreç ve yöntemlerle ilgili bilgi" sunduğunu belirtmektedir. Ayrıca, "belirli bir değerler kümesini (değer sistemleri) savunmayı amaçlamadığı" da ifade edilmektedir. YZ gibi karmaşık ve etik açıdan hassas bir alan için bu reçeteci olmayan, bilgilendirici yaklaşım stratejik olarak yaşamsal öneme sahiptir. TR'nin evrensel bir referans çerçevesi olarak hizmet etmesine, farklı kültürel, yasal ve operasyonel bağlamlarda ortak anlayışı ve diyaloğu teşvik etmesine olanak tanımaktadır. Bu, daha geniş uluslararası uzlaşmayı ve benimsemeyi kolaylaştırarak, farklı yargı bölgelerinin kendi özel düzenlemelerini YZ etiğine dair ortak bir temel anlayış üzerine inşa etmelerini sağlamaktadır.

Kullanıcının "açık kaynak bilgi" talebine rağmen, ISO/IEC TR 24368:2022 belgesinin tam metni telif hakkı ile korunmakta ve satın alınması gerekmektedir, hatta bazı versiyonları *Dijital Haklar Yönetimi* korumalıdır. Bu durum, küresel standardizasyon ekosistemindeki temel bir gerilimi ortaya koymaktadır. Standart Geliştirme Kuruluşları (SDO'lar), bu uluslararası referans noktalarını geliştirmek için gereken kapsamlı, işbirlikçi ve uzman odaklı süreçleri

finanse etmek için genellikle belge satışlarından elde edilen gelirlere güvenmektedir. Bu durum, bu tür standartların *özetleri* ve *kapsamları* kamuya açık olsa da, tam uygulamayı sağlayan ayrıntılı içeriğin genellikle bir maliyetle geldiğini vurgulamaktadır. Bu ticari model, özellikle küçük kuruluşlar, akademik araştırmacılar veya sivil toplum grupları için yaygın erişime bir engel oluşturabilir ve standardın erişimini ve etkisini potansiyel olarak sınırlayabilmektedir. Bu, önemli kamu yararı olan bir alanda temel etik rehberliğin erişilebilirliği konusunda etik bir değerlendirme ortaya koymaktadır.

B. Kapsam ve Amaç

ISO/IEC TR 24368:2022'nin temel işlevi, YZ'nin etik ve toplumsal endişelerine dair yüksek düzeyli bir genel bakış sunmaktır. Belge, YZ etiği alanındaki ilke, süreç ve yöntemlerle ilgili bilgiler sunmaktadır. Teknoloji uzmanları, düzenleyiciler, çıkar grupları ve genel olarak toplum dahil olmak üzere geniş bir paydaş yelpazesini hedeflemektedir. Belge, açıkça belirli bir değerler kümesini (değer sistemlerini) savunmayı amaçlamadığını belirtmektedir. Bu tarafsızlık, farklı kültürel ve operasyonel bağlamlara saygı göstermek için korunmaktadır.

Ayrıca, YZ'nin etik ve toplumsal endişelerinden kaynaklanan sorunları ele alan Uluslararası Standartlara genel bir bakış içermektedir.

Belge, kuruluşlara YZ yaşam döngüsü boyunca etik endişeleri belirlemek ve ele almak için yapılandırılmış bir yaklaşım sunmaktadır. YZ'yi çeşitli toplumsal değerlerle uyumlu hale getirmeye, güveni ve yeniliği teşvik etmeye yardımcı olmayı amaçlamaktadır. *Nemko Digital* (https://digital.nemko.com/), bu belgeyi YZ yönetişiminde önemli bir ilerleme olarak, "YZ etik ve toplumsal endişelerine özel olarak ayrılmış ilk uluslararası tanınmış standart" olarak tanımlamaktadır. Esnekliği ve değer-tarafsız duruşu, sağlık hizmetleri, finans, otonom araçlar ve akıllı şehirler dahil olmak üzere çeşitli sektörlerde uygulanabilir olmasını sağlamaktadır.

TR 24368'in "YZ etik ve toplumsal endişelerine özel olarak ayrılmış ilk uluslararası tanınmış standart" olarak tanımlanması, onu dönüm noktası niteliğinde bir belge olarak konumlandırmaktadır. "YZ'nin etik ve toplumsal endişelerinden kaynaklanan sorunları ele alan Uluslararası Standartlara genel bir bakış" içermesi, onun bir meta-standart veya yol gösterici belge olarak hareket ettiğini, daha geniş YZ standartlarına bakışı haritalandırdığını göstermektedir. Bu Teknik Rapor, YZ etiğinin karmaşık alanında yol alan kuruluşlar ve politika yapıcılar için kritik bir giriş noktası görevi görmektedir. YZ'nin etik sorunlarının *ne* olduğunu anlamak için ortak bir dil ve çerçeve sağlamakta, *nasıl* yönetileceğine (diğer, daha spesifik standartlar tarafından ele alınan) geçmeden önce bu temeli oluşturmaktadır. Bu temel rol, tutarlı

ve birbirine bağlı bir küresel YZ yönetişim çerçevesi oluşturmak için elzemdir ve daha sonraki, daha kuralcı standartların ortak bir etik anlayış üzerine inşa edilmesini sağlamaktadır.

Belgenin "değer-tarafsız duruşu" ve "belirli bir değerler kümesini savunmayı amaçlamadığı" yönündeki tekrarlanan vurgu, oldukça stratejik bir tasarım tercihidir. YZ etiği, farklı bölgeler ve topluluklar arasında önemli ölçüde değişen çeşitli kültürel, toplumsal ve felsefi değerlerle iç içe geçmiştir. Tek, kuralcı bir değer sistemini dayatmaya çalışmak muhtemelen direnişe yol açacak ve uluslararası benimsemeyi engelleyecektir. Bu tarafsızlık, TR'nin belirli ahlaki sonuçları dikte etmek yerine, etik kaygıları belirlemek ve tartışmak için evrensel bir çerçeve olarak işlev görmesine olanak tanımaktadır. Çeşitli paydaşların YZ etiğini kendi özel değer sistemleri ve düzenleyici ortamları içinde bağlamsallaştırmaları için ortak bir zemin sağlamaktadır. Bu pragmatik yaklaşım, geniş uluslararası işbirliğini sağlamak ve YZ teknolojilerinin küresel olarak sorumlu bir şekilde dağıtımını kolaylaştırmak için kritik öneme sahiptir.

C. Temel Temalar ve İlkeler

Belgenin girişi, YZ'nin gizlilik ihlalleri ve ayrımcı sonuçlar gibi potansiyel faydalarını ve risklerini tartışmaktadır. Taraflı veriler, opak karar verme, izlenebilirlik eksikliği ve toplumsal etkilerin yetersiz anlaşılması gibi etik kaygıların temel kaynaklarını tanımlamaktadır. Ayrıca, bu sorunları ele almak için kapsayıcı, disiplinlerarası bir yaklaşıma duyulan ihtiyacı vurgulamaktadır. Rapor, insan hakları hakkında genel bilgiler içermekte ve *RPC Legal* (https://www.rpclegal.com/), uygun veri kullanımını sağlamak, gizliliğe saygı göstermek ve YZ'nin ayrımcılık yapmasını önlemek de dahil olmak üzere insan haklarını korumanın önemini özellikle belirtmektedir.

ISO/IEC TR 24368:2022'de ana hatlarıyla belirtilen temel temalar ve ilkeler şunlardır :

- **Hesap Verebilirlik:** Bu ilke, etik denetimde netliği vurgulamaktadır. Etkili hesap verebilirlik sistemleri, YZ etik denetimi için net rol tanımları, etik YZ stratejisi için yönetici sorumluluğu, YZ zararlarını ele almak için olay müdahale protokolleri, düzenli denetim ve performans izleme ve karar izlenebilirliği için dokümantasyon standartlarını içermektedir. Hesap verebilirlik, teknik ekiplerin ötesine geçerek organizasyonel liderlik ve yönetişim yapılarını da kapsamaktadır.
- **Adillik ve Ayrımcılık Yapmama (Önyargı Azaltma):** Ayrımcı sonuçlara yol açabilecek YZ ile ilgili risklerin proaktif olarak belirlenmesini ve azaltılmasını gerektirmektedir. Adillik, bireysel, grup ve karşıolgusal adilliği kapsayan çok boyutlu bir kavram olarak

görülmektedir. Pratik unsurlar arasında potansiyel önyargıyı belirlemek için demografik etki testi ve itiraz edilen kararlar için itiraz süreçleri bulunmaktadır.

- **Şeffaflık ve Açıklanabilirlik:** Şeffaflık gereksinimlerinin YZ sisteminin risk seviyesiyle orantılı olması gerektiğini vurgulamaktadır. Bu, hem teknik şeffaflığı (örneğin, model mimarilerini ve eğitim verilerini belgeleme) hem de operasyonel şeffaflığı (örneğin, sistem sınırlamalarını ve kullanım durumlarını iletme) içermektedir. Özel uygulamalar arasında yüksek riskli kararlar için algoritmik açıklanabilirlik, veri kaynağı ve kalite dokümantasyonu, farklı demografik gruplardaki model performans metrikleri ve sistem sınırları ile etkileşim protokollerinin tanımlanması bulunmaktadır.
- **Mesleki Sorumluluk:** Temel bir tema olarak kabul edilmektedir.
- **İnsan Değerlerinin Teşviki:** Etik çerçevenin merkezindedir.
- **Gizlilik ve Güvenlik Önlemleri:** Hem doğrudan veri kullanımını hem de çıkarım risklerini ele alan tasarımdan gizlilik ilkelerini teşvik etmektedir. YZ modellerini düşmanca saldırılar ve veri zehirlenmesi gibi tehditlerden korumak için sağlam siber güvenlik önlemlerinin önemini vurgulamaktadır. Gelişmiş anonimleştirme teknikleri kullanarak gizlilik korumasını içermektedir.
- **Teknolojinin İnsan Kontrolü:** YZ'nin insan refahına hizmet etmesini sağlamak için temel bir ilkedir.
- **Toplum Katılımı ve Gelişimi (Katılım Stratejileri):** Etkili katılım stratejileri, farklı bakış açılarını bir araya getiren çok paydaşlı çalıştaylar, sürekli rehberlik ve denetim için topluluk danışma kurulları, temsili popülasyonlarla kullanıcı testleri, alan uzmanları ve etikçilerle uzman danışmanlığı ve şeffaflık ve girdi için kamu yorumu dönemlerini içermektedir.
- **İnsan Odaklı Tasarım:** YZ sistemlerini insan refahı ve ihtiyaçları ön planda tutularak tasarlamayı vurgulamaktadır.
- **Hukukun Üstünlüğüne ve Uluslararası Davranış Normlarına Saygı:** Yasal çerçevelere ve yerleşik uluslararası davranışlara bağlılık. *RPC Legal*, imza atan ülkelerin YZ'ye özgü riskleri düzenleme, vatandaşları koruma ve YZ'nin güvenli bir şekilde kullanılmasını sağlama yükümlülüğünü vurgulamaktadır.
- **Çevresel Sürdürülebilirlik:** YZ'nin çevresel etkisinin dikkate alınması.
- **İşgücü Uygulamaları:** YZ'nin istihdam ve işgücü üzerindeki etkisinin ele alınması.

Temalar ve ilkelerin kapsamlı listesi, ilgili ayrıntılı uygulamalarla birlikte, etik YZ'nin tek bir sorun değil, çeşitli hususların karmaşık bir etkileşimi olduğunu göstermektedir. Örneğin,

"Adillik" genellikle "Şeffaflık" (önyargıyı belirlemek için) ve "Hesap Verebilirlik" (onu ele almak için) ilkelerini gerektirmektedir. Belgenin "kapsayıcı, disiplinlerarası bir yaklaşım" vurgusu, bu bütünsel bakış açısını daha da pekiştirmektedir. Bu durum, kuruluşların YZ etiğini yalıtılmış olarak veya parçalı teknik düzeltmelerle ele alamayacağını göstermektedir. Gerçekten etik bir YZ çerçevesi, ilkelerin birbirini güçlendirdiği ve çözümlerin teknik, organizasyonel ve toplumsal boyutları kapsadığı entegre bir strateji gerektirmektedir. Bu, çapraz fonksiyonlu işbirliğini ve kapsamlı bir yönetişim yapısını zorunlu kılmaktadır.

Her yüksek düzeyli ilke (örneğin, Şeffaflık, Hesap Verebilirlik, Adillik) için, belgeler somut, eyleme geçirilebilir uygulama örnekleri sunmaktadır (örneğin, "Algoritmik açıklanabilirlik", "Net rol tanımları", "Demografik etki testi"). Bu, tartışmayı felsefi ideallerin ötesine, geliştiricilerin, dağıtıcıların ve düzenleyicilerin üstlenebileceği pratik adımlara taşımaktadır. ISO/IEC TR 24368:2022, YZ etiğinin soyut alanı ile YZ sistem geliştirme ve dağıtımının somut gerçekleri arasında hayati bir köprü görevi görmektedir. Pratik örnekler sunarak ve belirli önlemleri ana hatlarıyla belirterek, kuruluşların etik taahhütlerini doğrulanabilir süreçlere, metriklere ve yönetişim yapılarına dönüştürmelerini sağlamaktadır. Bu, etik YZ'yi daha az iddialı bir kavram ve daha çok ulaşılabilir bir operasyonel hedef haline getirmekte, mevcut organizasyonel iş akışlarına entegrasyonunu kolaylaştırmaktadır.

D. Etik ve Sosyal Olarak Kabul Edilebilir YZ Oluşturma ve Kullanma Uygulamalarına İlişkin Örnekler

Belge, etik YZ uygulamalarına yönelik genel organizasyonel uygulamalar sunmaktadır :

- Dahili süreçleri YZ ilkeleriyle uyumlu hale getirme.
- Bir kuruluşun takip edemeyeceği uygulamaları tanımlama (etik kırmızı çizgiler).
- Yeni YZ projeleri için bir inceleme süreci oluşturma.
- İlgili personele YZ etiği eğitimi sağlama.

Belge ayrıca belirli kullanım durumlarında uygulama unsurlarına da değinmektedir :

- **Sağlık Hizmetleri:** Temel özellikler arasında klinisyenlerin anlayabileceği ve doğrulayabileceği açıklanabilir teşhisler, farklı hasta popülasyonlarında önyargı izleme, gelişmiş anonimleştirme teknikleri kullanılarak gizlilik koruması ve klinik sonuçlara karşı sürekli doğrulama bulunmaktadır.

- **Finans:** Uygulama unsurları arasında kredi karar süreçlerinde algoritmik şeffaflık, potansiyel önyargıyı belirlemek için demografik etki testi, itiraz edilen kararlar için itiraz süreçleri ve tüketici savunuculuk gruplarıyla paydaş katılımı bulunmaktadır.

YZ yönetişim yolculuğu için aşamalı bir yaklaşım da ana hatlarıyla belirtilmiştir.

- **Aşama 1:** Değerlendirme ve Planlama: Mevcut YZ sistemlerini ve geliştirme süreçlerini değerlendirmeyi, etik riskleri ve uyumluluk gereksinimlerini belirlemeyi ve yönetişim yapılarını oluşturmayı ve sorumlulukları atamayı içermektedir.
- **Aşama 2:** Uygulama ve Entegrasyon: Etik çerçeveleri YZ geliştirme projelerine uygulamaya, izleme ve değerlendirme sistemlerini uygulamaya ve paydaş katılım süreçlerini oluşturmaya odaklanmaktadır.

Sağlık hizmetleri ve finans gibi farklı sektörler için belirli uygulama örneklerinin dahil edilmesi, temel etik ilkeler evrensel olsa da, pratik uygulamalarının önemli ölçüde bağlamsallaştırma gerektirdiğini göstermektedir. "Açıklanabilir teşhisler," şeffaflığın sağlık hizmetlerine özgü bir tezahürüdür; tıpkı "kredi kararlarında algoritmik şeffaflık"ın finans için olduğu gibi. Bu durum, YZ etiği için "tek beden herkese uyar" yaklaşımının yetersiz olduğunu göstermektedir. Sağlam bir genel çerçeve sağlarken, kuruluşları kendi özel sektörleri, düzenleyici ortamları ve YZ uygulamalarının benzersiz toplumsal etkileri ışığında ilkeleri uyarlamaya ve yorumlamaya teşvik etmektedir. Bu uyarlanabilirlik, standardın çeşitli alanlarda geniş uygulanabilirliği ve etkinliği için kritik öneme sahiptir.

YZ yönetişimine yönelik aşamalı bir yaklaşımın ana hatlarıyla belirtilmesi, "Değerlendirme ve Planlama" ile "Uygulama ve Entegrasyon" dahil olmak üzere, etik YZ'yi tek seferlik bir uyumluluk kontrolü yerine devam eden bir süreç olarak açıkça çerçevelemektedir. Bu, standardın etik kaygıları "YZ yaşam döngüsü boyunca" ele alma amacına uygun düşmektedir. YZ sistem yaşam döngüsünün 10 aşamasını tanımlayan ISO/IEC 8183'ten bahsedilmesi de bunu daha da pekiştirmektedir. **Etik YZ statik bir durum değil, dinamik, yinelemeli bir yolculuktur.** Kuruluşların, ortaya çıkan etik riskleri ele almak ve toplumsal değerlerle sürekli uyumu sağlamak için YZ sistemlerini ve süreçlerini sürekli olarak izlemesi, değerlendirmesi ve uyarlaması beklenmektedir. Bu, tüm YZ yaşam döngüsü boyunca etik bütünlüğü sürdürmek için sağlam yönetişim yapıları, sürekli eğitim ve entegre izleme sistemleri gerektirmektedir.

E. Küresel YZ Düzenlemesi ve Standardizasyon Ortamındaki Bağlam

ISO/IEC TR 24368:2022, ISO ve IEC tarafından yayınlanan çeşitli YZ ile ilgili standartlardan biridir. Diğer temel standartlarla entegre olmak ve onları tamamlamak üzere tasarlanmıştır:

- **ISO/IEC TR 24028:2020:** YZ sağlayan veya kullanan sistemlerin güvenilirliğini etkileyebilecek faktörlerin analizine odaklanmaktadır.
- **ISO/IEC 23894:2023:** YZ'nin geliştirilmesi ve kullanımıyla ilgili riskleri yönetmek için tüm sektörlerdeki kuruluşlara stratejik rehberlik sunmakta, risk yönetiminin YZ odaklı faaliyetlere nasıl entegre edileceği de dahil olmak üzere.
- **ISO/IEC 42001:2023:** YZ yönetim sistemleri için bir çerçeve sağlamaktadır.
- **ISO/IEC 8183:** Fikir aşamasından hizmet dışı bırakmaya kadar YZ sistem yaşam döngüsünün 10 aşamasını tanımlamakta, verimli veri yönetimini ve uyumluluğu sağlamaktadır.
- **ISO/IEC 25059:** Doğruluk, dayanıklılık, güvenilirlik ve fonksiyonel güvenlik gibi yönleri ele alan YZ Sistemleri için bir Kalite Modelidir.

Birleşik Krallık hükümeti, ISO ve IEC gibi uluslararası standart geliştirme kuruluşlarıyla işbirliğinin önemini vurgulamıştır. ISO/IEC TR 24368 gibi etik YZ çerçevelerinin benimsenmesinin temel faktörleri şunlardır:

- Dünya çapında ortaya çıkan YZ yasalarına düzenleyici uyumluluk.
- Önyargı, ayrımcılık ve gizlilik ihlallerine karşı risk azaltma.
- Müşteriler ve paydaşlarla güven oluşturma.
- Bilinçli pazarlarda rekabetçi farklılaşma.
- Yeteneği çekme, çünkü en iyi profesyoneller etik odaklı kuruluşları tercih etmektedir.

Stanford İnsan Merkezli YZ Enstitüsü'nden yapılan bir araştırma, bu etik çerçeveleri kullanan şirketlerin, yeni bölgesel YZ düzenlemelerine uyum sağlarken %31 daha düşük uyumluluk maliyetleri elde ettiğini göstermektedir. TR 24368, etik "genel bakışı" sağlamaktadır. Ancak, güvenilirlik (TR 24028), risk yönetimi (23894), yönetim sistemleri (42001), yaşam döngüsü yönetimi (8183) ve kalite (25059) gibi daha geniş bir ISO/IEC standartları paketi içinde konumlandırılmıştır. Bu, YZ standardizasyonuna yönelik bilinçli, modüler ve entegre bir yaklaşımı işaret etmektedir. Kuruluşların gerçekten kapsamlı ve sağlam bir YZ yönetişimi arayışında bu standartları ayrı ayrı görmemesi gerektiği anlaşılmaktadır. TR

24368, etik temeli oluşturmakta ve daha teknik ve yönetimsel standartların uygulanmasına rehberlik etmektedir. Bu entegre çerçeve, tüm YZ yaşam döngüsünü kapsayan bütünsel bir yaklaşıma olanak tanıyarak, yalnızca teknik performansı değil, aynı zamanda etik uyumu, sorumlu dağıtımı ve genel güvenilirliği de sağlamaktadır. Etik hususların teknik tasarımı bilgilendirdiği ve bunun tersi de geçerli olduğu bir sinerji yaratmaktadır.

Düzenleyici uyumluluk açık bir itici güç olsa da, belgeler aynı zamanda cazip stratejik ve ekonomik faydaları da vurgulamaktadır: "artan paydaş güveni, gelişmiş sistem performansı, rekabetçi farklılaşma ve yeteneği çekme". "%31 daha düşük uyumluluk maliyetleri" gibi ölçülebilir fayda, iş gerekçesini daha da sağlamlaştırmaktadır. Bu, sadece cezalardan kaçınmanın ötesine geçerek aktif olarak değer yaratmayı ifade etmektedir. Etik YZ, hızla "sahip olunması gereken" veya tamamen uyumluluk odaklı bir faaliyetten stratejik bir zorunluluğa dönüşmektedir. TR 24368 gibi çerçevelerin benimsenmesi, tüketicilerin, çalışanların ve düzenleyicilerin YZ aktörlerinden daha fazla sorumluluk talep ettiği bir pazarda giderek artan bir rekabet avantajı olarak görülmektedir. Bu, YZ teknolojilerinin gelecekteki başarısının ve yaygın benimsenmesinin, uluslararası standartların benimsenmesiyle yönlendirilen, kanıtlanabilir etik sağlamlıkları ve güvenilirlikleriyle ayrılmaz bir şekilde bağlantılı olacağını göstermektedir.

Aşağıdaki tabloda, ISO/IEC TR 24368:2022'de ana hatlarıyla belirtilen temel etik ilkelerin ve temaların, kamuya açık belgelerde açıklanan pratik tezahürleriyle birlikte açık, yapılandırılmış ve kolayca anlaşılır bir genel bakışı sunulmaktadır (Tablo-1).

Tablo 1: ISO/IEC TR 24368:2022'deki Temel Etik İlkeler ve İlgili Uygulamalar

İlke/Tema	**Açıklama/Temel Yön**	**İlgili Uygulamalar/Uygulama Unsurları**
Hesap Verebilirlik	Etik denetimde netlik ve sorumluluk.	YZ etik denetimi için net rol tanımları; etik YZ stratejisi için yönetici sorumluluğu; YZ zararlarını ele almak için olay müdahale protokolleri; düzenli denetim ve performans izleme; karar izlenebilirliği için dokümantasyon standartları.
Adillik ve Ayrımcılık Yapmama	Ayrımcı sonuçlara yol açabilecek YZ ile ilgili risklerin proaktif olarak belirlenmesi ve azaltılması.	Potansiyel önyargıyı belirlemek için demografik etki testi; itiraz edilen kararlar için itiraz süreçleri; farklı hasta popülasyonlarında önyargı izleme.

İlke/Tema	Açıklama/Temel Yön	İlgili Uygulamalar/Uygulama Unsurları
Şeffaflık ve Açıklanabilirlik	YZ sistemleri hakkında paydaşlara anlamlı, doğru, kapsamlı, erişilebilir ve anlaşılır bilgi sağlanması.	Yüksek riskli kararlar için algoritmik açıklanabilirlik; veri kaynağı ve kalite dokümantasyonu; farklı demografik gruplardaki model performans metrikleri; sistem sınırları ve etkileşim protokolleri.
Gizlilik ve Güvenlik Önlemleri	Hem doğrudan veri kullanımını hem de çıkarım risklerini ele alan tasarımdan gizlilik ilkelerinin teşviki.	Gelişmiş anonimleştirme teknikleri kullanarak gizlilik koruması; YZ modellerini tehditlerden korumak için sağlam siber güvenlik önlemleri.
İnsan Hakları Uygulamaları	İnsan haklarının korunması, uygun veri kullanımı, gizliliğe saygı ve YZ ayrımcılığının önlenmesi.	Genel insan hakları bilgileri; YZ'ye özgü riskleri düzenleme, vatandaşları koruma ve güvenli YZ kullanımını sağlama yükümlülüğü.
Toplum Katılımı ve Gelişimi	YZ sistemlerinin tasarımı, geliştirilmesi ve dağıtımında çeşitli paydaşların aktif katılımı.	Çok paydaşlı çalıştaylar; sürekli rehberlik ve denetim için topluluk danışma kurulları; temsili popülasyonlarla kullanıcı testleri; alan uzmanları ve etikçilerle uzman danışmanlığı; şeffaflık ve girdi için kamu yorumu dönemleri.
Mesleki Sorumluluk	YZ geliştiricileri ve dağıtıcıları için etik davranış ve uygulamalar.	Dahili süreçleri YZ ilkeleriyle uyumlu hale getirme; etik kırmızı çizgileri tanımlama; yeni YZ projeleri için inceleme süreci; YZ etiği eğitimi.
İnsan Değerlerinin Teşviki	YZ'nin insan refahını ve toplumsal ilerlemeyi desteklemesini sağlama.	
Teknolojinin İnsan Kontrolü	YZ sistemlerinin insan gözetimi ve kontrolü altında kalmasını sağlama.	
İnsan Odaklı Tasarım	YZ sistemlerini insan ihtiyaçları ve refahı ön planda tutularak tasarlama.	
Hukukun Üstünlüğüne ve Uluslararası Davranış Normlarına Saygı	Yasal çerçevelere ve yerleşik uluslararası davranışlara bağlılık.	
Çevresel Sürdürülebilirlik	YZ'nin çevresel etkisinin dikkate alınması.	

İlke/Tema	Açıklama/Temel Yön	İlgili Uygulamalar/Uygulama Unsurları
İşgücü Uygulamaları	YZ'nin istihdam ve işgücü üzerindeki etkisinin ele alınması.	

4. ISO/IEC AWI TS 25571 – YZ Sistem Yaşam Döngüsündeki Etik Sorunları Belgeleme Şablonu

ISO/IEC AWI TS 25571, YZ alanında geliştirilmekte olan önemli bir Teknik Şartnamedir. Bu belge, YZ sistemlerinin etik boyutlarını somut bir şekilde ele almayı amaçlamaktadır.

A. Resmi Tanım ve Durum

Belgenin resmi tanımı ISO/IEC AWI TS 25571'dir. Bazı belgelerde ISO/IEC NP TS 25571 olarak da geçmektedir; bu da bir "Yeni Çalışma Maddesi Önerisi" (NP) ve bir "Teknik Şartname" (TS) olduğunu göstermektedir. "AWI" (Onaylanmış Çalışma Maddesi), önerinin onaylandığını ve çalışmanın başladığını ifade etmektedir. Belgenin türü Teknik Şartname (TS) olarak belirtilmiştir. Teknik Şartnameler, genellikle tam Uluslararası Standartlardan daha az resmiyet taşır ancak daha ayrıntılı teknik özellikler veya rehberlik sağlar ve genellikle tam standardizasyonun erken olabileceği yeni teknolojiler için kullanılır. Belge şu anda "Öneri" aşamasındadır.

Konusu (YZ etiği) göz önüne alındığında, ISO/IEC TR 24368:2022'den sorumlu aynı komite olan ISO/IEC JTC 1/SC 42 YZ altında olması muhtemeldir. Ancak, 25571 için doğrudan komite belgede açıkça belirtilmemiştir. Geliştirme aşamasında bir standart olduğu için, yalnızca kapsamı, amacı ve tahmini zaman çizelgesi kamuya açıktır. Şablonun ayrıntılı içeriği bu aşamada kamuya açık değildir.

ISO/IEC TR 24368:2022, YZ'nin etik kaygılarına dair yüksek düzeyli bir "genel bakış" sunarak ilkeleri ve temaları belirlemektedir. Buna karşılık, ISO/IEC AWI TS 25571, "paydaşların potansiyel etik sorunları belgelemeleri için örnek bir şablon" sağlamayı amaçlamaktadır. Bu, standardizasyonda doğal ve gerekli bir evrimi temsil etmekte, etik YZ'nin *ne* olduğunu tanımlamaktan, onu operasyonelleştirmek için pratik *nasıl yapılır* araçları sağlamaya doğru ilerlemektedir. Teknik Şartname (TS) formatının seçimi bunun için uygundur, çünkü bir TR'den daha spesifik rehberliğe izin verirken, gelişen bir alan için esnekliği

korumaktadır. Bu ilerleme, ISO/IEC içinde YZ yönetişimine yönelik olgunlaşan bir yaklaşımı ifade etmektedir. Soyut etik ilkeler ile bunların gerçek dünya YZ sistemlerinde somut uygulaması arasındaki boşluğu kapatmaya yönelik stratejik bir niyeti göstermektedir. TS, TR 24368'de ana hatlarıyla belirtilen ilkelerin doğrudan bir uygulama yardımı olarak hizmet etmesi muhtemeldir ve kuruluşların etik hususları YZ yaşam döngüsü boyunca operasyonelleştirmeleri için yapılandırılmış bir metodoloji sunmaktadır.

TS 25571 için ana hatlarıyla belirtilen geliştirme zaman çizelgesi, "Öneri"den "Taslak", "Kamu Yorumları", "Yorum Çözümü", "Onay" ve "Yayın"a doğru ilerlemektedir. Bu, uluslararası standart geliştirmenin çok aşamalı, yinelemeli ve uzlaşmaya dayalı bir süreç olduğunu açıkça göstermektedir. "Kamu Yorumları" aşaması, dünya çapındaki paydaşlardan çeşitli girdiler ve geri bildirimler toplamak için kritik bir mekanizmadır. Bu titiz süreç, potansiyel olarak zaman alıcı olsa da, nihai standardın sağlamlığını, geniş uygulanabilirliğini ve uluslararası kabulünü sağlamak için tasarlanmıştır. YZ etiği kadar hassas ve hızla gelişen bir alan için, bu işbirlikçi yaklaşım, çeşitli bakış açılarını dahil etmek, potansiyel istenmeyen sonuçları ele almak ve ortaya çıkan rehberlik için meşruiyet oluşturmak için elzemdir. Paydaşlar bu devam eden süreci tanımalı ve bu önemli araçların geliştirilmesini etkilemek için proaktif olarak katılmalıdır.

B. Kapsam ve Amaç

ISO/IEC AWI TS 25571'in temel işlevi, YZ sistem yaşam döngüsünün her aşamasında ortaya çıkabilecek potansiyel etik sorunları belgelemek için paydaşlara örnek bir şablon sağlamaktır. Şablonun etkili kullanımı için ilgili talimatları da içermektedir. Birincil amacı, potansiyel etik sorunları proaktif olarak belirleyerek ve ele alarak YZ sistemlerinin güvenilirliğini ve etik bütünlüğünü artırmaktır. Hizmet sağlayıcılar, kullanıcılar ve geliştiriciler için etik YZ uygulamalarına hazırlanmaları ve bunlara uymaları için bir şablon sunmaktadır. Önerilen standart, kayıt yönetimine uygulanmakta ve YZ sistemleri içindeki etik ve toplumsal hususları ele almak için sistematik bir yaklaşım sunmaktadır. Şablon, etik ve toplumsal kaygıların belirlenmesine yardımcı olmak üzere tasarlanmıştır, bu da risk yönetimi süreçlerini bilgilendirebilir. Tamamlanmış bir şablondan toplanan bilgiler, risk yönetiminin ayrılmaz bir parçası olan YZ sistem etki değerlendirmelerini de bilgilendirebilir.

TS 25571'in açık amacı, "potansiyel etik sorunları proaktif olarak belirlemek ve ele almak" ve "risk yönetimi süreçlerini bilgilendirmek"tir. Bu, etik hususları YZ geliştirme yaşam döngüsünün ayrılmaz, önceden belirlenmiş bir parçası olarak yerleştirmeye yönelik bir değişimi

işaret etmektedir. Bu standart, YZ geliştiren ve dağıtan kuruluşlar içinde etik durum tespiti süreçlerini kurumsallaştırmayı amaçlamaktadır. Yapılandırılmış bir şablon sağlayarak, paydaşları bir YZ projesinin en erken aşamalarından itibaren etik riskleri sistematik olarak haritalandırmaya, tahmin etmeye ve azaltmaya yetkilendirmektedir. Bu proaktif yaklaşım, YZ hatalarının olasılığını azaltmak, sistem güvenilirliğini artırmak ve daha sorumlu ve sürdürülebilir yeniliği teşvik etmek için kritik öneme sahiptir.

Standardın "kayıt yönetimi"ne uygulanması ve "YZ sistemleri içindeki etik ve toplumsal hususları ele almak için sistematik bir yaklaşım" sağlama rolü oldukça önemlidir. Ne kadar iyi niyetli olursa olsun, etik hususlar uygun dokümantasyon ve izlenebilirlik olmadan soyut kalabilir. Şablonun çıktısının "YZ sistem etki değerlendirmelerini" bilgilendirmesi, doğrulanabilir kayıtlara duyulan ihtiyacı daha da vurgulamaktadır. Bu Teknik Şartname, YZ için etkili etik yönetişimin sağlam dokümantasyon uygulamalarıyla ayrılmaz bir şekilde bağlantılı olduğunu kabul etmektedir. Etik hususları kayıt yönetimine entegre ederek, hesap verebilirliği, denetlenebilirliği ve sürekli öğrenmeyi kolaylaştırmaktadır. Bu, kuruluşların sadece etik ilkelere bağlılık iddia etmekle kalmayıp, aynı zamanda bunu doğrulanabilir kayıtlar aracılığıyla gösterebilecekleri anlamına gelmektedir; bu da düzenleyici uyumluluk, dahili denetim ve dış güven oluşturmak için giderek daha hayati hale gelmektedir. Etik YZ'yi bir politika beyanından somut bir uygulamaya dönüştürmektedir.

C. İlgili Tartışmalar ve Kavramsal Temeller

"Şeffaf ve Açıklanabilir YZ'ye Doğru: ISO/IEC standartlarının geliştirilmesini bilgilendirmek için çalıştay" başlıklı bir çalıştayın tartışmaları oldukça önemlidir. Bu çalıştay, açıkça ISO/IEC AWI 12792 ve ISO/IEC AWI TS 6254'ü beslemeyi amaçlasa da, kavramsal tartışmalar etik dokümantasyonu ele alan herhangi bir standart için temel niteliktedir.

Çalıştaydan önerilen tanımlar şunlardır :

- **Şeffaflık:** "Paydaşlara, bir YZ sisteminin ilgili yönleri hakkında anlamlı, doğru, kapsamlı, erişilebilir ve anlaşılır bilgi mevcudiyeti" olarak tanımlanmıştır. Tartışmalar, "tam" açıklamanın her zaman anlamlı veya erişilebilir olup olmadığını ve "doğru" ile "gerçekçi" temsil arasındaki farkı içermiş, "gerçekçi" kelimesinin "doğru" veya "dürüst" bir şekilde temsil etmekten daha büyük bir yükümlülük taşıdığı önerilmiştir.
- **Yorumlanabilirlik (algoritmalar):** "Bir paydaşın, bir YZ sisteminin amacını, sistemin davranış nedenlerini ve amacına uygun olarak ve paydaş beklentileri doğrultusunda çalışıp çalışmadığını ve farklı girdilerin farklı sonuçlara nasıl yol açabileceğini

zamanında anlama kolaylığı" olarak tanımlanmıştır. Teknik yaklaşımlar, açıklanabilirlik yöntemlerinin yanı sıra diğer analiz veya görselleştirme yöntemlerini de içermektedir. Anlama için eğitim gerekebilir ve paydaşın dikkate alınması önemlidir.

- **Açıklanabilirlik (politika):** "Bir YZ sisteminin paydaşlarına, YZ sisteminin sonuçlarının ötesinde özlü, erişilebilir, yeterli ve faydalı açıklayıcı bilgi sağlama yeteneği" olarak tanımlanmıştır. Bu bilgi, model davranış gerekçelerini, geliştirme sürecine ilişkin dokümantasyonu, sistem sınırlamalarının beyanını veya uzman olmayan kullanıcılara temel algoritmik eğitim sunmayı içerebilir.
- **Açıklanabilirlik (algoritmalar):** "Bir YZ sisteminin kendi davranış nedenlerini zamanında doğru bir şekilde üretme yeteneği, amacına uygun olarak ve paydaş beklentileri doğrultusunda çalışıp çalışmadığının ve farklı girdilerin farklı sonuçlara nasıl yol açabileceğinin incelenmesine olanak tanır" olarak tanımlanmıştır.
- **Yorumlanabilirlik ve Açıklanabilirlik Arasındaki Ayrım:** Tartışmalar, özellikle model davranışını anlama konusunda net bir ayrım ihtiyacını vurgulamıştır. Modeller, iç çalışma mekanizmaları ve dolayısıyla sonuçları alan uzmanları tarafından anlaşılabilirse yorumlanabilirdir.

Genel YZ sisteminin doğruluğu, dayanıklılığı ve güvenilirliği ve fonksiyonel güvenliği, çalışma grubuna gayri resmi olarak geri bildirilecek konular olarak tartışılmıştır. Bu yönler, yakın zamanda yayınlanması beklenen ISO/IEC 25059 – YZ Sistemleri için Kalite Modeli'nde açıkça ele alınmaktadır.

"Şeffaflık," "yorumlama" ve "açıklanabilirlik" gibi terimler için yapılan kapsamlı tartışmalar ve önerilen tanımlar, ortak ve kesin bir kelime dağarcığının sadece akademik bir egzersiz değil, etkili etik standardizasyon için kritik bir ön koşul olduğunu vurgulamaktadır. "Doğru" ve "gerçekçi" arasındaki tartışma gibi nüanslar, pratik dokümantasyonun temelini oluşturacak kavramları tanımlamak için gereken derinlemesine düşünmeyi göstermektedir. TS 25571 gibi bir şablonun etik sorunları belgelemede etkili olabilmesi için, temel kavramların açık ve net bir şekilde tanımlanması gerekmektedir. Bu tanımlama çabaları, farklı disiplinlerden ve bağlamlardan paydaşların standardın rehberliğini tutarlı bir şekilde yorumlayabilmesini ve uygulayabilmesini sağlamaktadır. Bu temel çalışma, ortak bir anlayış oluşturmak ve etik yönetişim çabalarını zayıflatabilecek yanlış yorumlamaları önlemek için kritik öneme sahiptir.

Şeffaflık/açıklanabilirlik tartışmaları ile doğruluk, dayanıklılık, güvenilirlik ve fonksiyonel güvenliği ele alan ISO/IEC 25059 (YZ Sistemleri için Kalite Modeli) arasındaki

açık bağlantı, etik hususların YZ sistemlerinin teknik performansından ve güvenilirliğinden ayrı olmadığını ortaya koymaktadır. Etik sorunları belgelemek, genellikle sistem kalitesinin etik sonuçları nasıl etkilediğini anlamayı gerektirmektedir (örneğin, hatalı bir model haksız sonuçlara yol açabilir). Bu, YZ yönetişiminin bütünsel görünümünü pekiştirmektedir. Etik sorunları belgelemek için kullanılan şablon (TS 25571), YZ sistem kalitesi ve güvenilirliği ile ilgili değerlendirmelerden girdi almayı ve bunlara katkıda bulunmayı gerektirecektir. Örneğin, bir önyargı sorununu belgelemek, 25059 tarafından kapsanan modelin performans metriklerinin değerlendirilmesini gerektirebilir. Bu entegre yaklaşım, etik YZ'nin sadece uyumlulukla ilgili değil, aynı zamanda şeffaf ve açıklanabilir, sağlam, güvenilir ve toplumsal açıdan faydalı YZ sistemleri oluşturmakla ilgili olmasını sağlamaktadır.

5. Karşılaştırmalı Analiz ve Birlikte Çalışabilirlik

ISO/IEC TR 24368:2022 ve ISO/IEC AWI TS 25571, YZ etiği ve yönetişiminin geniş çerçevesi içinde birbirini tamamlayan roller oynamaktadır. Bu iki belge, YZ'nin sorumlu bir şekilde geliştirilmesi ve dağıtılması için hem kavramsal bir temel hem de pratik bir uygulama aracı sağlamaktadır.

<u>YZ Etik Yönetişiminde Tamamlayıcı Roller</u>

- **ISO/IEC TR 24368:2022 ("Ne" ve "Neden"):** YZ'nin etik ve toplumsal endişelerine dair "yüksek düzeyli bir genel bakış" sunmakta ve "ilke, süreç ve yöntemlerle ilgili bilgi" sağlamaktadır. Etik hususların geniş manzarasını ana hatlarıyla belirleyen temel bir etik pusula görevi görmektedir.
- **ISO/IEC AWI TS 25571 ("Nasıl"):** YZ sistem yaşam döngüsü boyunca "potansiyel etik sorunları belgelemek için paydaşlara örnek bir şablon" sağlamayı amaçlamakta ve kullanımı için talimatları içermektedir. Etik hususları uygulamak için eyleme geçirilebilir bir araç olarak tasarlanmıştır.

TR 24368, YZ etiğinin "ne" ve "neden"ini tanımlayan entelektüel ve etik temeli oluşturmaktadır. Geliştirilmekte olan TS 25571 ise, "nasıl"ı kuruluşların bu ilkeleri YZ yaşam döngüsü boyunca uygulamaları için somut, yapılandırılmış bir mekanizma sağlamaktadır. Bu, sağlam yönetişim çerçevelerinin geliştirilmesinde mantıksal ve gerekli bir ilerlemeyi temsil etmektedir. Pratik araçları olmayan yüksek düzeyli bir etik çerçeve, yalnızca iddialı kalma riski taşırken, yol gösterici ilkeleri olmayan bir araç, gerçek etik düşünce olmaksızın sadece uyumluluğa yol açabilir. Bu sinerji, YZ etiğini teoriden pratiğe taşımak için kritik öneme sahiptir. Bu standartlar, birlikte, kuruluşların etik hususları YZ uygulamalarına entegre etmeleri

için daha eksiksiz ve eyleme geçirilebilir bir yol sunarak, hem etik zorlukların derinlemesine anlaşılmasını hem de bunları sistematik olarak ele alma operasyonel yeteneğini teşvik etmektedir. Bu bütünsel yaklaşım, anlamlı ve kanıtlanabilir sorumlu YZ'ye ulaşmak için hayati öneme sahiptir.

İlkelerden Uygulamaya Köprü Kurma

TR 24368, şeffaflık, hesap verebilirlik ve adillik gibi temel etik ilkeleri ana hatlarıyla belirtmektedir. TS 25571, etik sorunları belgeleme odağıyla, bu ilkelerin işlevselleştirilmesini doğrudan desteklemektedir. Şeffaflık ve açıklanabilirlik için önerilen tanımlar etrafındaki tartışmalar, bu ilkelerin TS şablonu kullanılarak nasıl belgeleneceği ve değerlendirileceği ile doğrudan ilgilidir.

Risk Yönetimi ve Etki Değerlendirmelerini Bilgilendirme

TR 24368, YZ risklerini anlamak için bir bağlam sağlamaktadır. TS 25571, şablonunun etik ve toplumsal kaygıların belirlenmesine yardımcı olabileceğini ve "risk yönetimi süreçlerini" ve "YZ sistem etki değerlendirmelerini" bilgilendirebileceğini açıkça belirtmektedir. Bu, pratik risk azaltma stratejileriyle doğrudan bir bağlantıyı göstermektedir.

Güven ve Sorumlu YZ'yi Teşvik Etme

Her iki standart da, etik hususları ele almak için yapılandırılmış yaklaşımlar sağlayarak YZ dağıtımında güven ve itimat oluşturma genel hedefine katkıda bulunmaktadır.

Daha Geniş YZ Standartları Ekosistemi ile Entegrasyon

Nemko Digital, TR 24368'in ISO/IEC 23894:2023 (YZ risk yönetimi) ve ISO/IEC 8183 (YZ sistem yaşam döngüsü) gibi tamamlayıcı standartlarla entegrasyon yollarına sahip olduğunu vurgulamaktadır. TS 25571, yaşam döngüsü dokümantasyonuna odaklanarak ve risk değerlendirmelerini bilgilendirerek, bu entegrasyon noktalarını daha da güçlendirmektedir.

TR 24368, etik kaygıları "YZ yaşam döngüsü boyunca" ele almak için yapılandırılmış bir yaklaşımı vurgulamaktadır. TS 25571, sorunları "YZ sistem yaşam döngüsünün her aşamasında" belgelemek için açıkça bir şablon sağlamaktadır. Her iki belgede de tüm yaşam döngüsüne (konseptten hizmet dışı bırakmaya kadar) yapılan bu tutarlı vurgu, ISO/IEC YZ standardizasyonundaki temel bir felsefenin altını çizmektedir. Etik hususlar, tek seferlik bir kontrol veya YZ geliştirmenin tek bir aşamasıyla (örneğin, dağıtım) sınırlı değildir. Bunun yerine, YZ sisteminin tüm varlığı boyunca sürekli olarak yerleştirilmeli, izlenmeli ve

yönetilmelidir. Bu, sürekli uyanıklığı, yinelemeli etik incelemeleri ve sorumlu YZ'ye proaktif ve entegre bir yaklaşıma geçişi kolaylaştıran araçların (TS 25571 gibi) kullanılmasını zorunlu kılmaktadır.

6. Sonuç: Standardizasyon Yoluyla Etik YZ'nin İlerlemesi

ISO/IEC TR 24368:2022 ve ISO/IEC AWI TS 25571'in YZ etiği ve yönetişiminin görünen kritik rolleri kapsamlı bir şekilde incelenmiştir. Her iki standart da, farklı ancak tamamlayıcı işlevleriyle, YZ teknolojilerinin sorumlu bir şekilde geliştirilmesi ve dağıtılması için uluslararası bir çerçeve oluşturma çabalarına önemli katkılarda bulunmaktadır.

ISO/IEC TR 24368:2022'nin Önemi

ISO/IEC TR 24368:2022, YZ'nin etik ve toplumsal endişelerine özel olarak ayrılmış ilk uluslararası tanınmış standart olarak öne çıkmaktadır. Yüksek düzeyli bir genel bakış sunarak, YZ etiği için temel ilkeleri, süreçleri ve yöntemleri ana hatlarıyla belirleyerek kavramsal bir temel sağlamaktadır. Stratejik değer-tarafsız duruşu, geniş uluslararası uygulanabilirliği ve uzlaşmayı mümkün kılmaktadır. Bu Teknik Rapor, YZ'nin karmaşık etik manzarasında yol alan kuruluşlar ve politika yapıcılar için bir kılavuz görevi görmektedir.

ISO/IEC AWI TS 25571'in Önemi

Geliştirme aşamasındaki ISO/IEC AWI TS 25571, YZ sistem yaşam döngüsü boyunca potansiyel etik sorunları belgelemek için pratik bir araç olarak tasarlanmıştır. Proaktif tanımlama ve yapılandırılmış dokümantasyon yoluyla YZ sistemlerinin güvenilirliğini ve etik bütünlüğünü artırmayı amaçlamaktadır. Ayrıca, risk yönetimi süreçlerini ve YZ sistem etki değerlendirmelerini bilgilendirmede önemli bir rol oynamaktadır. Bu, etik hususları YZ geliştirme ve dağıtımının operasyonel süreçlerine entegre etmek için somut bir mekanizma sunmaktadır.

Yapay Zekâ Yönetişimi Üzerindeki Tamamlayıcı Doğa ve Etki

TR 24368, temel etik çerçeveyi sağlarken, TS 25571, bunun uygulanması için pratik araçları sunmaktadır; bu, YZ standardizasyonunda kritik bir ilerlemeyi temsil etmektedir. Bu iki standart, soyut ilkelerden eyleme geçirilebilir uygulamalara geçiş yaparak güvenilir ve sorumlu YZ'yi teşvik etmeye toplu olarak katkıda bulunmaktadır. Kuruluşlara YZ'nin karmaşık etik manzarasında rehberlik etmede, yeniliğin toplumsal değerlerle uyumlu olmasını sağlamada önemli bir rol oynamaktadırlar.

YZ etiği ve yönetişimi, yayınlanmış bir Teknik Raporun ve aktif geliştirme aşamasındaki bir Teknik Şartnamenin bir arada varlığı, devam eden tanımlama tartışmaları ve standart geliştirmenin yinelemeli doğası ile oldukça dinamik ve gelişen bir alandır. Bu durum, teknolojik ilerlemenin hızlı temposu ve çeşitli bölgesel YZ düzenlemelerinin ortaya çıkmasıyla daha da karmaşık hale gelmektedir. Kuruluşlar, politika yapıcılar ve araştırmacılar için YZ etiği, statik bir uyumluluk egzersizi olarak ele alınamaz. Yeni ve gelişen standartların sürekli izlenmesi, kamu yorumu dönemlerine proaktif katılım ve dahili uygulamaların çevik bir şekilde uyarlanması, uyumlu kalmak, riskleri etkili bir şekilde azaltmak ve YZ sistemlerine olan kamu güvenini sürdürmek için elzem olacaktır. Bu, YZ yönetişiminde sürekli öğrenme, uyum ve işbirlikçi evrimden oluşan bir ekosistemi teşvik etmektedir.

Açık Kaynak Bilgilerin Önemi ve Etik Veri İşleme

Kamuya açık bilgilerin, bu standartların kapsamını ve amacını anlamada değeri büyüktür. Ancak, tam belge erişiminin ticari yönleri ve Standart Geliştirme Kuruluşlarının finansman modelleri ile erişilebilirliği dengelemenin devam eden zorluğu kabul edilmelidir. Bilgilerin etik, doğru ve fikri mülkiyet haklarına ve erişim sınırlamalarına tam saygı gösterilerek sunulmasına olan bağlılık sürdürülmelidir.

Bu standartların uluslararası kuruluşlar (ISO/IEC) tarafından geliştirilmesi, TR 24368 tarafından benimsenen değer-tarafsız yaklaşım ve ulusal hükümetlerin (örneğin, Birleşik Krallık) SDO'larla işbirliğine verdiği önem, sorumlu YZ için ortak bir zemin oluşturmaya yönelik uyumlu bir küresel çabayı işaret etmektedir. Bölgesel düzenlemeler belirli nüanslar getirebilse de, TR 24368 ve yakında çıkacak TS 25571 gibi uluslararası standartlar, YZ etiği için kritik bir temel ve ortak bir dil sağlamaktadır. Bu uyumlaştırma, küresel ölçekte yeniliği ve sorumlu dağıtımı teşvik etmek, parçalanmayı ve etik standartlarda potansiyel "en dibe doğru yarışı" önlemek için hayati öneme sahiptir. Bu uluslararası çabalar, küresel olarak güvenilir ve faydalı bir YZ geleceği inşa etmek için temeldir.

7. Kaynakça

- https://journal.standard.ac.ir/article_195913b9928026c481968c5f0681d46a029595.pdf
- https://standardsdevelopment.bsigroup.com/projects/9024-11178
- https://webstore.iec.ch/en/publication/78464
- https://webstore.iec.ch/en/iec_catalog/product/preview/?id=L3B1Yi9wZGYvcHJldmlldy9pbmZvX2lzb2llY3RyMjQzNjh7ZWQxLjB9ZW4ucGRm
- https://digital.nemko.com/standards/iso-iec-24368

- https://webstore.ansi.org/standards/iso/isoiectr243682022
- https://www.rpclegal.com/thinking/artificial-intelligence/ai-guide/part-5-ai-regulation-globally/
- https://aistandardshub.org/forums/topic/output-from-workshop-on-iso-iec-standards-for-ai-transparency-and-explainability/
- https://www.open-std.org/jtc1/sc22/wg14/www/projects
- https://share.ansi.org/ISOT/Updated%20ISO-IEC-ITU%20coordination/2024%20ISO-IEC-ITU%20New%20work%20items/ISO%20IEC%20ITU%20coordination%20-%20New%20work%20items%20(2024-11)..pdf
- https://arxiv.org/pdf/2506.13839

2.4.2. ISO/IEC TR 24027:2021, ISO/IEC TR 24028:2020 Teknik Raporları ile ISO/IEC CD TS 22443 Teknik Şartnamesi

Bu bölümde ISO/IEC TR 24027:2021 ve ISO/IEC TR 24028:2020 teknik raporları ile ISO/IEC CD TS 22443 teknik şartnamesi incelenmektedir. Sıralanan metinlerin incelenmesinde sırasıyla teknik raporun ve/veya şartnamenin;

(i) kapsamı, uygulama alanı, hedef kitlesi, içeriği,
(ii) dünyada ve Türkiye'de uygulanma durumu, getirdiği zorluklar ve sağladığı faydalar,
(iii) Türkiye açısından mevcut ve olası etkileri,
(iv) gelecekteki durumuna yönelik değerlendirmeler "Genel Değerlendirme" alt başlığında irdelenmektedir.

Bunlara ek olarak teknik rapora ve/veya şartnameye ilişkin Türkiye'nin politikalarına yönelik öneriler "Türkiye İçin Öneriler" alt başlığında sunulmaktadır. Burada önemle dikkat çekilmesi gereken husus, ISO/IEC CD TS 22443 teknik şartnamesinin 11 Eylül 2025 tarihinde iptal edilmiş olmasıdır. Bu durum ilgili şartnamenin değerlendirildiği alt başlıkta irdelenmiş olmakla birlikte söz konusu bilgiye burada da yer verilmesi önem taşımaktadır.

1. ISO/IEC TR 24027:2021 Bilgi Teknolojileri – Yapay Zekâ – YZ Sistemlerindeki Önyargılar ve YZ Destekli Karar Alma

ISO/IEC TR 24027:2021, YZ'de önyargıların kaynaklarını haritalayan, ölçüm tekniklerini ve azaltma yöntemlerini açıklayan, üst düzey etik ilkeler ile daha kuralcı

yönetim/güvence standartları arasında kavramsal ve operasyonel bir köprü görevi gören, pratik, yaşam döngüsü odaklı bir teknik rapordur.

Genel Değerlendirme

ISO/IEC TR 24027, ISO/IEC JTC 1/SC 42 tarafından yayınlanan bir teknik rapordur (Kasım 2021). Bu teknik raporun amacı, tüm YZ yaşam döngüsü (veri toplama, eğitim, sürekli öğrenme, tasarım, test, değerlendirme ve kullanım) boyunca "*YZ sistemlerinde ve YZ destekli karar verme süreçlerinde önyargıyı tespit etmek ve ele almak için mevcut en iyi uygulamaları açıkça belirtmektir*". Temel hedef kitlesi, belirli uygulamalarda önyargı risklerini ortaya çıkarmak, ölçmek ve azaltmak için uygulanabilir yöntemlere ihtiyaç duyan geliştiriciler, dağıtıcılar, satın alma görevlileri, denetçiler ve düzenleyicilerdir. Rapor, önyargıya yol açan sınıflandırma ve tanımlar, kaynaklar ve nedensel zincirler, ölçüm yaklaşımları aileleri (grup adalet ölçütleri, kalibrasyon, hata oranı analizleri vb.) ve ön işleme, işleme ve son işleme seçenekleri olarak gruplandırılmış yenileme stratejileri ile birlikte, ödünleşimleri belgeleme ve izleme konusunda kılavuzluk içermektedir.

Küresel açıdan incelendiğinde TR 24027'nin, uluslararası YZ standardizasyonunda ve ulusal standart kuruluşları tarafından temel bir belge olarak yaygın bir şekilde referans alındığı görülmektedir; CEN/CLC, AB politika araçlarıyla (*AB YZ Yasası kapsamında uyumlaştırılmış standartlar için hazırlık çalışmaları dâhil*) ilgisini artıran bir Avrupa PD (*PD CEN/CLC ISO/IEC TR 24027:2023*) yayınlamıştır. Önemli standart yayıncıları ve gözlemcileri (OECD.AI, BSI Group, ulusal standart kuruluşları) bu raporu adaletin işlevselleştirilmesi için pratik bir başlangıç noktası olarak indekslemekte ve tanıtmaktadır.

Türkiye açısından incelendiğinde ise Türkiye'nin, TSE'nin ayna komitesi (MTC-195) aracılığıyla ISO/IEC SC 42 çalışmalarına katıldığı ve bu sayede TR 24027'yi benimsemek veya uyum sağlamak için resmî kanallara sahip olduğu görülmektedir. TR 24027'ye açıkça atıfta bulunan özel bir Türk ulusal baskısına veya kamuya açık bir TSE kılavuz notuna çevrimiçi olarak yaygın bir şekilde ulaşılamamış olsa da, bu teknik raporu benimseme ve yerelleştirme kurumsal kapasitesinin, ayna komitesi ve TSE kanalları aracılığıyla mevcut olduğunu söylemek mümkündür. Bu durum da söz konusu teknik raporun Türkiye'ye uyarlanmasını sağlayacak olan ulusal mevzuat için bir fırsat olarak da değerlendirilmektedir.

Raporun taşıdığı zorluklar ve fırsatlar birlikte değerlendirildiğinde raporun kavramsal bir yapıda olması önemli bir zorluk olarak görülmektedir. Bu kavramsal yapısı raporun özelliklerini ölçülebilir performans tanımlayıcıları ve denetim kanıtlarına dönüştürmek için

daha fazla çeviri gerekliliğine işaret etmektedir (örneğin eşlik eden kontrol katalogları veya süreç standartları). Bir diğer zorluk, önyargının ölçülmesine dayanmaktadır. Başka bir deyişle önyargının ölçümünün bağlama göre oluşması metrik seçimi ve yorumlama konularını bir belirsizlik alanına dönüştürebilmektedir. Ayrıca yerel veri sınırlamaları genelleştirilmiş metriklerin doğrudan uygulanmasını zorlaştırmaktadır.

Bunlara karşın rapor, ISO/IEC risk ve yönetim standartları (örneğin, ISO/IEC 23894 ve ISO/IEC 42001) ile birleştirilebilen, satın alma ve ihracat için denetlenebilir kanıtlar oluşturmak üzere önyargı çalışmaları için genel kabul görmüş bir terminoloji ve metodoloji sunmaktadır ve bu durum, AB pazarlarını hedefleyen veya sınır ötesi kamu ihalelerine katılan kuruluşlar için özellikle değerli görülmektedir.

Teknik raporun Türkiye'de kamu özelinde yaratabileceği etkilerin TSE, bakanlıklar ve ihale makamlarının ihale belgelerinde TR 24027'ye atıfta bulunması ile ilişkili olduğu değerlendirilmektedir. Böylesi bir durumda ihale tarafları birbiriyle uyumlu ve/veya tutarlı karar alma sistemleri/planlarına sahip olacak, bu da karşılaştırılabilirliği artıracak ve vatandaşlara yönelik sistemler (sosyal hizmetler, yardımların dağıtımı, istihdam taraması) için daha bilinçli ihale kararlarının alınmasını destekleyecektir. Ayrıca teknik raporun rehberliği, AB'nin adalet ve ayrımcılık yapmama beklentileriyle uyum sağlamayı da mümkün hale getirebilecektir.

Özel sektör açısından ise TR 24027 uygulamalarını benimseyen Türk YZ satıcıları ve entegratörleri, uluslararası müşterilere daha iyi bir şekilde gerekli özeni ve risk yönetimini gösterebilecek, böylelikle AB ve adil belgelerin beklendiği diğer yargı bölgelerine pazar erişimini iyileştirebilecektir. Ayrıca bu teknik rapor önyargılı sonuçlardan kaynaklanan itibar, yasal ve operasyonel riskleri azaltmak için bir YZ yönetim sistemi içinde operasyonel kontrollere de dönüştürülebilecektir.

TR 24027'nin, daha kuralcı ve sertifikalandırılabilir ürünler (uyumlaştırılmış standartlar, sektör profilleri, yönetim sistemi standartları) gelişirken, kalıcı bir kavramsal referans olmaya devam edeceği düşünülmektedir. Uygulamada, terminolojisi ve yapısı ulusal kılavuzlarda, satın alma şablonlarında, denetim kontrol listelerinde ve eğitim müfredatlarında yeniden kullanılacak ve yerelleştirilecektir. AB YZ Yasası kapsamında uyumlaştırma süreci ilerledikçe, TR 24027'nin içeriği muhtemelen Avrupa ve komşu pazarların adalet beklentileri için bir mihenk taşı olacaktır.

Türkiye İçin Öneriler

ISO/IEC TR 24027:2021, Türkiye'ye YZ sistemlerinin tasarımına ve yönetimine adalet ve önyargı azaltma ilkelerini yerleştirmek için zamanında ve pratik bir temel sunmaktadır. Bununla birlikte ulusal ve uluslararası çalışmalarda politika düzenleme/eğitim çalışmaları için referans olarak kullanılmasına rağmen, bu belgenin, gereklilikleri belirten tam bir standart değil, bir teknik rapor olduğuna dikkat çekilmesi önemli görülmektedir. Çünkü bu durum, bahse konu belgenin zorunlu veya sertifikalandırılabilir olmaktan çok tavsiye niteliğinde olduğu anlamına gelmektedir.

Yine de bu teknik raporun tam değerini ortaya çıkarmak için, ulusal çabalar yerelleştirme, satın alma ve politikaya entegrasyon ve veri ve uzmanlığa sürekli yatırımlara odaklanılması önemli görülmektedir. Her şeyden önce, Türkiye bu kılavuzu Türk Standartları Enstitüsü (TSE) aracılığıyla benimsemeli ve yerelleştirmelidir. TSE, YZ ayna komitesi (MTC-195) aracılığıyla, TR 24027'nin teknik içeriğini veri koruma (KVK) ve sektöre özgü yönetişim dâhil olmak üzere Türk yasal ve düzenleyici bağlamlara uyarlayan resmî bir kılavuz notu yayınlamak için uygun konumdadır. Yerel olarak ilgili örnekler ve terminoloji içeren Türkçe bir versiyon, hem kamu hem de özel sektör paydaşları tarafından erişilebilirliği ve uygulamayı önemli ölçüde artıracaktır.

İhale, adaleti işlevselleştirmek için etkili bir araçtır. Kamu kurum ve kuruluşları, YZ tedarikçilerinden ihale tekliflerinin bir parçası olarak, TR 24027'nin önerileri temelinde açıkça yapılandırılmış, eksiksiz bir "Önyargı Yönetimi Planı" sunmalarını talep etmelidir. Bu plan, veri seti sayfaları, model kartları, kullanılan adalet ölçütlerinin gerekçeleri ve izleme stratejileri gibi kanıtları içermelidir. Adalet kriterlerini kamu sözleşmelerine dâhil ederek Türkiye, hesap verebilirliği sağlayabilir ve sorumlu YZ için beklentiler belirlemede örnek teşkil edebilir.

Önyargı azaltmanın sadece bir kontrol listesinden daha fazlası olmasını sağlamak için kuruluşlar TR 24027'yi sertifikalandırılabilir yönetişim standartlarıyla entegre etmeye teşvik edilmelidir. Özellikle, ISO/IEC 23894 (AI risk yönetimi) ve ISO/IEC 42001 (AI yönetim sistemleri) adil olma niyetini belgelenmiş, denetlenebilir uygulamalara dönüştürmek için gerekli çerçeveleri sağlamaktadır. TR 24027'yi bu sistemlerle birleştirmek, kuruluşların önyargı riskini belirlemek ve yönetmek için şeffaf, tekrarlanabilir prosedürler oluşturmasına olanak tanıyacaktır.

Yerel veri kalitesi, anlamlı önyargı değerlendirmesinin temel taşıdır. Türkiye, özellikle Türk doğal dil işleme (NLP), bilgisayarlı görme ve konuşma tanıma alanlarında, çeşitli ve temsili veri kümeleri ve karşılaştırma setleri oluşturmaya yatırım yapmalıdır. Buna paralel olarak, kuruluşların kullandıkları verilerin özelliklerini, sınırlamalarını ve temsil gücünü

belgelemeleri zorunlu hale getirilmelidir. Bu, adalet testlerini iyileştirmekle kalmayacak, aynı zamanda Türkiye'nin nüfusunu ve bağlamını yansıtan verilerle eğitilmiş YZ sistemlerine halkın güvenini de güçlendirecektir.

Son olarak, herhangi bir ulusal adalet stratejisinin başarısı, insanlara ve kurumlara bağlıdır. YZ geliştiricileri, denetçiler ve kamu alımları için TSE ve üniversitelerle işbirliği içinde sunulan eğitim programları aracılığıyla kapasite geliştirmeye öncelik verilmelidir. Ayrıca Türkiye, TR 24027 tabanlı kriterlere göre YZ sistemlerini değerlendirebilecek akredite bağımsız test laboratuvarlarının geliştirilmesini desteklemelidir. Bu laboratuvarlar, kamuoyunun güvenini artıran ve pazara erişimi destekleyen tarafsız bir adalet güvencesi sinyali sunan tasdik raporları hazırlayabilecektir.

Özetle, ISO/IEC TR 24027:2021'nin sadece teknik bir rapor değil, stratejik bir belge olarak değerlendirilmesi önemlidir. Bu niteliğiyle Türkiye'nin satın alma, yönetişim ve ihracat hazırlığı stratejilerine entegre edilmeye hazırdır. Başarılı bir şekilde benimsenmesi, bilinçli yerelleştirme, sertifikalandırılabilir standartlarla uyum ve veri, test ve insan kapasitesi için güçlü bir ulusal altyapının oluşturulmasına bağlıdır.

2. ISO/IEC TR 24028:2020 Bilgi Teknolojileri – YZ – YZ'de güvenilirliğe genel bakış

TR 24028, ISO/IEC JTC 1/SC 42 tarafından yayınlanan bir teknik rapor olup, "*güvenilir YZ (trustworthy AI)*"nın temelini oluşturan *kavramları, riskleri ve yöntemleri* incelemek amacıyla hazırlanmıştır. Teknik rapor, şeffaflık, açıklanabilirlik ve kontrol edilebilirlik için yaklaşımlar; yaygın tuzaklar ve risk kaynakları; YZ sistemlerinde kullanılabilirlik, esneklik, güvenilirlik, doğruluk, emniyet, güvenlik ve gizlilik elde etmenin yolları hakkında bilgiler sunmaktadır. Bu rapor, sertifikasyon için bir kontrol listesi değil, güvenilirlik konularının bir haritası olarak görülebilecektir.

Genel Değerlendirme

ISO/IEC TR 24028:2020 gereklilikleri belirlemek yerine ilkeleri ve uygulamaları sentezlediği için, politika yapıcılar, satın alımcılar, geliştiriciler, denetleyiciler ve risk yöneticileri tarafından kavramsal bir temel ve daha kuralcı standartlar (örneğin, ISO/IEC 42001 YZ yönetim sistemleri veya alana özgü güvenlik/emniyet normları) için çapraz referans olarak kullanılmaktadır. Teknik rapor, güvenilirliği etkileyen faktörleri (teknik ve sosyo-teknik), YZ süreçlerinde tipik tehditleri/zayıflıkları ve yaşam döngüsü boyunca hafifletici önlemleri analiz

etmektedir. Ayrıca bu teknik rapor, güvenilirliğin *çok boyutlu* olduğunu ve bazı ödünlerin (örneğin, doğruluk, sağlamlık ve gizlilik) verilmesini gerektirdiğini vurgulamaktadır.

Küresel durum incelendiğinde TR 24028'in, ISO, OECD.AI ve Avrupa standardizasyon yol haritaları tarafından temel bir referans olarak yaygın bir şekilde işlenmiş olduğu görülmektedir. Avrupa'da, CEN-CENELEC JTC 21, AB'nin çalışmalarını, AB YZ Yasası kapsamında nihai kullanım için ISO/IEC çıktılarıyla uyumlu hale getirmektedir; TR 24028, güvenilirlik ortamının bir parçası olarak AB'nin "*devamlı plan (rolling plan)*" ve gözlem materyallerinde yer almaktadır.

Türkiye, TSE bünyesinde ISO/IEC SC 42 ve CEN-CENELEC JTC 21 çalışmalarını takip eden bir ayna komitesine (AI Mirror Committee-MTC-195) sahiptir. Bu komite, 24028 standardını ulusal düzeyde benimsemek veya referans almak ve paydaşların katkılarını koordine etmek için resmi bir yapıdır. TSE'nin açıklamaları, bu komitenin uluslararası/AB YZ standartları ve ulusal YZ stratejisi ile uyum sağlama konusundaki rolünü vurgulamaktadır.

Teknik raporun boşlukları ve/veya zayıf yönleri arasında ilk olarak raporun sertifikalandırılamayan bir yapıda olması gelmektedir. Başka bir ifadeyle bir teknik rapor olarak TR 24028, tek başına uygunluk değerlendirmesini mümkün kılmamakta ve bu nedenle de kuruluşların, sertifikalandırılabilir çerçeveler (örneğin ISO/IEC 42001 AIMS) ve risk standartları aracılığıyla kavramlarını *işlevselleştirmeleri* gerekir. Ayrıca ölçülebilir kriterler ve temel performans göstergeleri (key performance indicators-KPIs) dönüştürmek, özellikle farklı sektörlerde, oldukça zorlayıcı bir iştir. Bu teknik rapor özellikle *nasıl kanıtlanacağı* yerine *nelere dikkat edilmesi gerektiğini* araştırmaktadır. Son olarak güvenilirlik, güvenlik, mahremiyet, sağlamlık ve etik gibi konuları kapsaması dolayısıyla uzmanlar, birden fazla standardı ve AB kılavuzunu karşılaştırmalı olarak incelemek zorundadır ve bu da alan uzmanları için ek bir yük anlamına gelmektedir.

Teknik raporun güçlü yönleri ve/veya taşıdığı fırsatlar açısından akla gelen ilk konu AB YZ Yasası'na hazırlık olarak değerlendirilebilecek olmasıdır. TR 24028, güvenilir YZ ile ilgili Avrupa politika hedefleriyle uyumlu ortak bir terminoloji ve kapsam sunmaktadır. Bu durum, AB ile ticaret yapan veya AB'de faaliyet gösteren Türk kuruluşlar için yararlı bir temel oluşturmaktadır. Ayrıca ISO/IEC 42001 ile birleştirildiğinde, TR 24028'in kavramları yönetim sistemlerine, denetimlere, tedarikçi gözetimine ve sürekli iyileştirmeye entegre edilebilecektir.

Teknik rapor, Türkiye'deki bakanlıklar, belediyeler, kamu iktisadi teşebbüsleri gibi kamu aktörlerinin TR 24028'i politika ve tedarik alanındaki güven gereklilikleri için temel

sınıflandırma olarak kullanabilmesini de mümkün kılmaktadır. Örneğin bu sayılan kamu aktörleri tedarikçilerden sistemlerinin *güvenilirlik, sağlamlık, güvenlik, emniyet, gizlilik, şeffaflık, açıklanabilirlik* ve *kontrol edilebilirlik* konularını nasıl ele aldığını haritalandırmalarını isteyebilecek, ardından bunları risk kayıtları ve izleme planlarıyla ilişkilendirebilecektir. Bu da kurumlar arasında tutarlı beklentiler yaratabilecek ve onları AB odaklı pazar ihtiyaçları ile uyumlu hale getirebilecektir. Buna ek olarak kamusal denetim otoriteleri, denetim kontrol listelerini ve kılavuz notlarını TR 24028'in özelliklerine (*başka bir ifadeyle neyin "iyi" olduğuna*) uyarlayabilirken, güvence kanıtı ve süreç olgunluğu için TR 42001 veya sektörel standartlara başvurabileceklerdir.

Özel sektör için ise pazar erişimi ve iş ortağı güvencesi ile iç yönetişim konuları önemli görülmektedir. Buna göre ürün belgelerinde ve risk açıklamalarında TR 24028'in terminolojisini kullanmak, AB müşterileri nezdinde güvenilirliği artırabilecek ve kuruluşları AB YZ Yasası kapsamında gelecekteki uyumlaştırılmış standartlara hazırlayabilecektir.

Ayrıca ISO/IEC 42001 ile birlikte kullanıldığında, bu raporun konuları politika kontrolleri, standart çalışma prosedürleri (SOPs) ve temel performans göstergeleri (KPIs) (örneğin sağlamlık test planları, model arızalarına yönelik olay müdahalesi, şeffaflık beyanları) haline gelerek denetlenebilirliği ve dayanıklılığı artırmaktadır.

TR 24028'in, daha özel ve sertifikalandırılabilir standartlara işaret eden temel bir genel bakış olmaya devam edeceği öngörülmektedir. AB YZ Yasası'nın uyumlaştırılmış standartları olgunlaştıkça, buradaki kavramlarının Avrupa ve sektörel profiller tarafından referans alınması ve kuruluşların bu standardın özelliklerini YZ yönetim sistemleri (ISO/IEC 42001) içindeki kontrol kataloglarına dönüştürmesi beklenmektedir. Diğer bir deyişle, bu teknik raporun rolünün, tek başına bir güvence aracı olmaktan ziyade, istikrarlı bir *kavramsal iskelet* kurmak olduğu değerlendirilmektedir.

Türkiye İçin Öneriler

YZ sistemlerinin güvenilir olmasını sağlamak, bunların yenilikçi olmasını sağlamak kadar önemlidir. ISO/IEC TR 24028:2020, YZ'nin güvenilirliğini tanımlamak ve değerlendirmek için uluslararası düzeyde tanınan bir çerçeve sunarak, YZ sistemlerini güvenilir, güvenli, açıklanabilir ve kamu kullanımı için uygun hale getiren temel özellikleri özetlemektedir. Türkiye için bu standart, hem düzenleyiciler hem de uygulayıcılar için ortak bir referans noktası haline gelebilir; tedarik süreçlerini şekillendirebilir, özel sektör yönetişimine

rehberlik edebilir, AB ve küresel pazarlarla uyumu sağlayabilir. TR 24028'i Türk kurumları aracılığıyla kurumsallaştırarak ve özelliklerini kurumsal uygulamalara entegre ederek, Türkiye güvenilir YZ için sağlam ve geleceğe dönük bir yönetişim modeli oluşturabilir.

Bunun için, TR 24028'in güvenilirlik özelliklerini mevcut Türk yasal çerçevelerine, özellikle veri korumaya ilişkin KVKK, sektörel kurallara (örneğin finans, sağlık, havacılık) ve AB YZ Yasası'nın temalarına uyarlayan TSE ayna komitesi (MTC-195) aracılığıyla bir ulusal kılavuz notu oluşturulması önerilmektedir. Böyle bir kılavuz notu iki pratik araç sağlayacaktır:

- Satın alanlar için, tüm ihalelerde sistematik olarak güvenilirlik kanıtı talep edilmesini sağlayan hazır bir kontrol listesi,
- Tedarikçiler için, uyumluluğu kolaylaştıran ve YZ sistemlerinin hem ulusal hem de Avrupa gerekliliklerini karşıladığını gösteren ortak bir şablon oluşturulmuş olacaktır.

Kamu kurum ve kuruluşları, YZ güvenilirliğinin temel tanımı olarak ihalelerinde TR 24028'e atıfta bulunmaya teşvik edilmelidir. Aynı zamanda, tedarikçilerden ISO/IEC 42001 (AI Yönetim Sistemi) süreçleri (belgelenmiş politikalar, risk kayıtları, izleme raporları gibi) veya eşdeğer yönetişim mekanizmaları aracılığıyla uyumluluk kanıtı göstermeleri de istenmelidir. Bu ikili yaklaşım, güvenilirliği ölçülebilir bir satın alma gerekliliği haline getirirken, farklı olgunluk düzeylerindeki kuruluşlara esneklik sağlayacaktır.

Güvenilirliği işlevsel hale getirmek için Türkiye, ulusal bir "Güvenilirlik Kontrol Kataloğu" oluşturması önerilmektedir. Katalog, TR 24028'in her bir özelliği (güvenilirlik, sağlamlık, güvenlik, gizlilik, şeffaflık, açıklanabilirlik ve kontrol edilebilirlik) için aşağıdaki hususları tanımlayacaktır:

- Kontrol hedefleri (neye ulaşılması gerektiği),
- Sorumlu sahipler (kimlerin hesap verebilir olduğu),
- Metrikler ve temel performans göstergeleri (KPIs) (uyumluluğun nasıl ölçüleceği),
- Test prosedürleri (uyumluluğun nasıl doğrulanacağı),
- Kanıt belgeleri (hangi belgelerin gerekli olduğu) ve
- İnceleme sıklığı (kontroller ne sıklıkla yapılacağı).

Bu katalog, organizasyonel YZ Yönetim Sistemlerine (AIMS, ISO/IEC 42001) doğrudan entegre edilmeli ve güvenilirlik kontrollerinin günlük yönetişim ve denetim süreçlerinin bir parçası olmasını sağlamalıdır. Aynı derecede önemli olan bir diğer husus, model yaşam döngüsü aşamalarını güven kontrollerine uyumlu hale getirmektir. Örneğin:

- **Eğitimden önce:** veri kümeleri sağlamlık ve temsil edilebilirlik incelemelerinden geçmelidir.
- **Doğrulamadan önce:** modeller tanımlanmış sağlamlık eşiklerini karşılamalıdır.
- **Dağıtımdan önce:** şeffaflık beyanları yayınlanmalı ve olay müdahale planları hazır olmalıdır.
- **Çalışma sırasında:** izleme, sapma, önyargı ve güvenilirlik sorunlarını tespit etmeli ve düzeltici eylemler teşvik edilmelidir.

Yaşam döngüsü kararlarını güvenilirlik testlerine bağlayarak, kuruluşlar veri yayınlamadan üretim aşamasına kadar her aşamada hesap verebilirliği yerleştirirler. Güvenilirlik, yetenekli profesyoneller ve test altyapısı olmadan kök salamaz. Türkiye bu nedenle, TR 24028'i ders programının temelini oluşturacak şekilde, TSE Akademi ve üniversiteler aracılığıyla ulusal eğitim programları başlatması önerilmektedir. Bu programlar, sağlamlık testi, tehdit modelleme, tasarımdan itibaren gizlilik, şeffaflık belgeleri ve güvenilirlik denetimi gibi konuları kapsayan teori ile uygulamalı pratikleri birleştirmelidir. Bu tür bir eğitim, hem kamu görevlilerini (ihale ve denetimleri değerlendirecek olanlar) hem de özel sektör çalışanlarını (uyumluluğu uygulamakla yükümlü olanlar) hazırlayacaktır. TR 24028'i bir öğrenme yolu haline getirerek Türkiye, güvenilir YZ mühendisliği ve doğrulaması yapabilen bir işgücü yetiştirebilir ve standartların sadece kâğıt üzerinde değil, pratikte de uygulanmasını sağlayabilir.

Türkiye'nin uluslararası gelişmeleri aktif olarak izlemesi ve bunlara uyum sağlaması önerilmektedir. Özellikle, CEN-CENELEC JTC 21 çıktılarını ve AB YZ Yasası uyumlaştırma sürecini takip etmek çok önemli olacaktır. Türk profilleri, tedarik şablonları ve kontrol katalogları, uyumsuzluğu önlemek için Avrupa standartlarıyla paralel olarak gelişmelidir. Önemli güncellemeleri Türkçeye çevirmek ve bunları kamu kurumlarına, sektör derneklerine ve KOBİ'lere erken bir aşamada yaymak, yerel paydaşlara rekabet avantajı sağlayacaktır.

Ortak Araştırma Merkezi (JRC) yayınları ve AB kılavuz notları, güncellemeler için düzenli referans noktaları olarak hizmet edebilir. Yakın uyumu sürdürerek Türkiye, YZ yönetişiminin AB pazarları için güvenilir olmasını sağlarken, aynı zamanda yerli inovasyonu da teşvik etmektedir.

ISO/IEC TR 24028:2020, YZ güvenilirliğinin *ortak dilini* sağlamaktadır. Neyin ele alınması gerektiğini (özellikler, riskler, risk azaltma önlemleri) tanımlarken, ISO/IEC 42001 ve sektöre özgü standartlar, kuruluşların pratikte uyumluluğu nasıl kanıtlayabileceklerini göstermektedir. TR 24028'i kılavuz notları aracılığıyla kurumsallaştırarak, tedarik ve kurumsal

yönetişime entegre ederek, eğitim ve yerel uzmanlığa yatırım yaparak ve AB'deki gelişmelerle uyumlu kalarak Türkiye, güvenilir YZ alanında bölgesel lider konumuna gelebilecektir.
Bu kombinasyon, kamu alımlarını güçlendirme, özel sektör yönetişimini iyileştirme, küresel rekabet gücünü artırma ve Türk YZ sistemlerinin sadece yurt içinde değil, uluslararası pazarlarda da güvenilir olmasını sağlama potansiyelini taşımaktadır.

3. *ISO/IEC CD TS 22443 Bilgi Teknolojileri – YZ – Toplumsal Kaygıları ve Etik Hususları Ele Alma Konusunda Rehberlik*

ISO/IEC CD TS 22443 bir teknik şartname (*Technical Specification-TS*) taslağıdır ve ISO özetinde bu belgenin "*kuruluşların YZ sistemlerinin yaşam döngüsü boyunca bireylere ve topluma yönelik olarak doğurabileceği olası toplumsal endişeleri ve etik hususları nasıl belirleyip ele alabileceği konusunda rehberlik sağladığı*" belirtilmektedir. Teknik şartname, Eylül 2025 tarihinde ISO/IEC JTC 1/SC 42 kapsamında "*Komite Taslağı (Committee Draft-CD)*" aşamasında iken (Edition, 1, 2025) ISO proje yaşam döngüsü, Komite Taslağı aşamasının 2025 yılının başlarında tamamlandığını ve projenin 11 Eylül 2025 tarihinde iptal edildiğini göstermektedir. Bu durum, bahse konu teknik şartname üzerindeki resmî çalışmaların durdurulduğu ve sonraki resmî aşamalara (Uluslararası Standart Taslağı, Nihai Uluslararası Standart Taslağı, yayın) ilerlemeyeceği anlamına gelmektedir.

YZ konusunda toplumsal ve etik kaygılarla ilgili özel bir şemsiye kılavuz olması amaçlanan bu teknik şartnamenin iptal edilmesinin muhtemel nedeninin, kapsamı itibariyle halihazırda yayınlanmış olan standartlarla (risk yönetimi, etki değerlendirmesi, yönetişim, şeffaflık ve yönetim sistemleri) büyük ölçüde örtüşmesi olduğu düşünülmektedir.

Bununla birlikte bu başlık altında teknik şartnamenin iptal edilmesinin duyurulmasından önce hazırlanan değerlendirme sunulmaktadır.

Genel Değerlendirme

ISO/IEC CD TS 22443, *"mevcut YZ sistemi yönetişimi, yönetim sistemi ve etki değerlendirme standartlarını genişletmeyi"* amaçlamaktadır ve YZ sistemleri geliştiren veya kullanan tüm kuruluş türleri için geçerlidir. Diğer bir deyişle, standardın bağlamı, etik ve toplumla ilgili YZ yönetişimidir; hedef kitlesi, YZ uygulayan tüm kuruluşlardır (kamu veya özel, büyük veya küçük); ve içeriği, YZ yaşam döngüsü boyunca toplumsal/etik riskleri (önyargı, gizlilik veya güvenlik sorunları gibi) belirleme ve yönetme yöntemlerini kapsamaktadır. Özellikle, SC 42'nin liderliği, YZ'de "*şu anda toplumsal endişeleri ve etik*

hususları ele almaya yönelik rehberlik sunan yeni bir teknik şartname üzerinde çalıştıklarını" doğrulamaktadır. Bu, TS 22443'ün yeni bir politika getirmek yerine, uygulanabilir kılavuzlar (mevcut etik çerçeveleri tamamlayıcı nitelikte) sunarak bir boşluğu doldurmayı amaçladığını vurgulamaktadır.

Standardın mevcut durumu ile taşıdığı zorluklar ve fırsatlara ilişkin olarak değerlendirmelerin öncesinde ISO/IEC TS 22443'ün hâlâ geliştirilme aşamasında olduğuna dikkat çekilmesi önemli görülmektedir. CD danışma süreci Nisan 2025 tarihinde tamamlanmıştır ve belge, onaylanırsa DIS ve yayın aşamalarına geçecektir. SC 42'nin güvenilir YZ (trustworthy AI) konusunda geniş yetkisi (*şeffaflık, önyargı, yönetişim gibi konuları kapsar*), CD TS 22443'ün YZ etik ekosistemindeki birçok birbiriyle bağlantılı standarttan biri olduğu anlamına gelmektedir. Küresel bağlamda incelendiğinde, 60'tan fazla ülkenin SC 42'ye katılarak çok çeşitli bakış açılarını yansıttığı görülmektedir. Burada söz konusu olan çeşitlilik hem bir fırsat hem de bir zorluk oluşturmaktadır: bir yandan, uluslararası düzeyde en iyi uygulamaları uyumlu hale getirebilecekse de diğer yandan, toplumsal değerler (*farklı hukuk sistemleri, kültürler, gelişmişlik düzeyleri*) konusunda uzlaşma sağlanması karmaşık bir konudur. Başka bir ifadeyle uygulamada bu, uluslararası uyumu teşvik edebilecekse de (*örneğin, işletmelerin EU YZ Tüzüğü gibi düzenlemelere uymasına yardımcı olabilir*) aynı zamanda çeşitli ulusal yasalar ve öncelikler arasında denge kurulmasını da gerektirecektir.

Türkiye'de, YZ'yi özel olarak düzenleyen kesinleşmiş bir yasa henüz bulunmamaktadır. Ancak, politika yapıcıların bu alanda çalışmalarını sürdürdüğü bilinmektedir. Türkiye'nin bu konudaki son çalışmaları arasında, Kişisel Verileri Koruma Kurumu tarafından YZ kullanımı ve gizlilikle ilgili sektörel kılavuzların yayınlanması ve güvenlik, şeffaflık, eşitlik, hesap verebilirlik ve gizlilik ilkelerine dayanan bir YZ tasarısının hazırlanması ve Haziran 2024 tarihinde Meclis'e sunulması öne çıkmaktadır. Bu ortamda Türkiye için birtakım fırsatlar bulunmaktadır: TS 22443 kılavuzunun benimsenmesi, Türk kuruluşlarının ortaya çıkan küresel normlara uyum sağlamasına ve gelecekteki düzenlemelere hazırlanmasına yardımcı olabilecektir. Örneğin, yukarıda da yer verildiği üzere TSE halihazırda ISO/IEC JTC 1/SC 42'ye YZ konusunda MTC 195 Teknik Komitesi'ni kurarak katılmaktadır. Bu doğrultuda Türkiye'nin TS 22443'ün geliştirilmesine katkıda bulunması mümkün hale gelmektedir ve bu katılım metnin içeriğinin Türkiye'nin kendi ulusal bakış açısını yansıtmasını sağlayabilecektir. Zorluklar arasında ise uluslararası bir kılavuzun Türk hukuku ve uygulamalarına uyarlanması ile özellikle küçük ve orta ölçekli işletmelerin oldukça ayrıntılı olabilecek bu kılavuzu

uygulayabilmelerinin sağlanması gelmektedir. Bununla birlikte, Türkiye'nin şu anda *"YZ'nin sorumlu kullanımını teşvik etme"* konusuna odaklanması, TS 22443'ün etik ve hesap verebilirlik konusundaki ulusal önceliklerle uyumlu olduğunu göstermektedir.

Henüz taslak halinde olan ISO/IEC TS 22443 teknik şartnamesinin yayımlanması halinde, Türkiye'de önemli etkiler yaratabileceği öngörülmektedir. Özel sektör için, YZ geliştiricilerine ve satıcılarına toplumsal/etik riskleri öngörme ve azaltma konusunda bir yol haritası sunarak, rehberlik edebilecektir (örneğin, YZ etki değerlendirmeleri yaparak). Uluslararası kabul görmüş kılavuzları takip ederek, Türk teknoloji şirketleri YZ ürünlerinin güvenilirliğini artırabilecek ve küresel pazarlara daha kolay erişebilir hale gelebilecektir. Türkiye'nin ulusal stratejik planlarında *"YZ merkezleri"* ve inovasyon desteğinin vurgulandığı göz önünde bulundurulduğunda bunun özellikle değerli olduğu sonucuna ulaşmak mümkündür. Çünkü ISO belgelerine uyum, yeni YZ kuluçka merkezleri veya hızlandırıcılar için bir kalite işareti haline gelebilecektir. Kamu için TS 22443, kamu kurum ve kuruluşları ile kamu hizmetlerine YZ'yi toplumsal açıdan sorumlu bir şekilde entegre etme konusunda bilgilendirici niteliktedir. Teknik şartnamenin yaşam döngüsü yaklaşımı, birtakım bakanlıkların (örneğin sağlık, eğitim, ulaştırma ve altyapı gibi) bu kılavuzları satın alma kriterlerine, strateji belgelerine veya denetim çerçevelerine entegre edebileceği anlamına gelmektedir. Esasen, TS 22443, kamu kurumlarının etik YZ'yi (ethical AI) işlevselleştirmesine yardımcı olarak YZ mevzuat taslağının genel ilkelerini somut süreçlerle tamamlayabilecektir.

Hem kamu hem de özel sektör için ortak sonuç YZ'nin nasıl kullanıldığı konusunda daha fazla tutarlılık ve şeffaflık olacaktır. Bunun gibi uluslararası YZ standartlarının geliştirilmesine yönelik çalışmalar, *"YZ'nin benimsenmesinin önündeki engelleri aşarken etik, toplumsal ve teknik endişeleri ele almak için gereklidir"*. Türkiye için bu, daha sorunsuz bir dijital dönüşüm anlamına gelebilecektir. Özellikle, kuruluşların adalet, gizlilik ve risk yönetimi konusunda ISO'nun kılavuzunu izlemesi ve bunun vatandaşlar tarafından görülmesi, onların YZ destekli hizmetlere daha fazla güvenmeleri teşvik edebilecektir. Ayrıca, TS 22443'ün bir rehber olarak kullanılması rekabet avantajı sağlayabilir. Örneğin, yurtdışında YZ çözümleri satan Türk firmaları, ISO/IEC TS 22443 ile uyumluluğu etik tasarım ve kurumsal sorumluluğun kanıtı olarak kullanabilecektir. Kamuda, bu kılavuzlara uyum, her ayrıntıyı yasallaştırmaya gerek kalmadan politikaların küresel normlarla (ve dolayısıyla AB gibi ortaklarla) uyumunu sağlayabilecektir. Bu tür standartların benimsenmesinin paydaşların güvenini ve birlikte

çalışabilirliği artırma eğiliminde olduğunu ve bunun Türkiye'nin YZ ekosistemi için önemli faydalar sağlama potansiyeli taşıdığını söylemek mümkündür.

ISO/IEC CD TS 22443'ün Uluslararası Standart Taslağı aşamasına geçmesi ve sonunda yayınlanması beklenmektedir. Yakın gelecekte, tutarlılığı sağlamak için ilgili ISO/IEC belgeleriyle (örneğin, YZ yönetim sistemi standardı 42001 veya YZ etki değerlendirmesi 42005) birlikte gözden geçirilmesi muhtemeldir. YZ'nin hızlı gelişimi göz önüne alındığında, kılavuz ilk yayınlanmasını takiben güncellenebilir veya genişletilebilir. Örneğin, yeni toplumsal endişeler (karbon ayak izi veya sosyal medyadaki yanlış bilgiler gibi) hızlı revizyonlara neden olabilir. Teknik şartnamenin "*toplumsal endişeler*"e odaklanması, eşitsizlik veya iklim etkisi gibi konulara uyum sağlamaya devam etmesi gerektiği anlamına gelmektedir.

Küresel olarak, TS 22443, yasal gereklilikler yerine teknik en iyi uygulamaları sağlayarak, AB YZ Yasası, OECD YZ ilkeleri ve UNESCO YZ Etiği Tavsiye Kararı gibi diğer çerçeveleri tamamlayacaktır. Farkındalık arttıkça, daha fazla ülkenin ISO kılavuzunu düzenlemelere veya kurumsal standartlara dâhil etmesi beklenebilir. UNESCO "*sürdürülebilirlik, şeffaflık ve güvenilirlik gibi toplumsal ihtiyaçları çalışmalarımızın merkezine yerleştirmek, YZ'nin tam potansiyelini ortaya çıkarmaya yardımcı olabilir*" diyerek YZ'nin belirleyici yönüne vurgu yapmaktadır. SC 42'nin taslağı başarılı olursa, TS 22443 dünya çapında "*etik YZ yönetişimi*" için temel referanslardan biri haline gelebilir. Türkiye için, YZ konusunda uluslararası işbirliği yoğunlaştıkça, bu tür küresel standartlarla uyum sağlanması giderek daha önemli hale gelecektir.

Türkiye İçin Öneriler

ISO/IEC TS 22443'ten en fazla faydayı sağlamak için Türkiye'nin proaktif adımlar atması önerilmektedir. İlk olarak, Türk paydaşların (TSE ve MTC 195 aracılığıyla) CD'yi inceleyerek ve yorumlar sunarak standardın geliştirilmesine katılınabilir. Aktif katılım, ülkemiz perspektiflerinin (kültürel değerler, yasal bağlam, sektörel ihtiyaçlar) yansıtılmasını sağlayacaktır. İkinci olarak, Türkiye'nin politika yapıcılarının, ulusal YZ girişimlerine karşı taslak kılavuzu haritalandırması öngörülmektedir. YZ Yönetmelik Tasarısı ve veri koruma kılavuzları zaten şeffaflık ve hesap verebilirlik ilkelerini vurgulamaktadır; teknik şartnamenin pratik kılavuzunu bu politikalara entegre etmek, uygulamaya açıklık getirecektir. Örneğin, düzenleyiciler sektörel kılavuzlarda veya hatta YZ ürünleri için akreditasyon programlarında TS 22443'e atıfta bulunabilirler.

Ayrıca hem kamu hem de özel kuruluşlar TS 22443'ün içeriğini bilmeli ve taşıdığı önerileri pilot olarak uygulamayı düşünmelidir. Bu, standartla ilgili eğitim programları veya TSE tarafından yürütülen atölye çalışmaları içerebilecektir. YZ kullanan şirketler, risk yönetimi süreçlerinin bir parçası olarak bu teknik şartnameyi kullanabilir. Kamu kurum ve kuruluşları (örneğin e-devlet projeleri), ISO/IEC YZ etik kılavuzuna uyumu gerektiren sözleşme maddelerini benimseyebilir. Son olarak, eğitim ve araştırma kurumları, teknik şartnamenin kavramlarını müfredatlara ve düşünce kuruluşlarına dâhil ederek gelecekteki işgücünün uluslararası YZ etik normlarına hâkim olmasını sağlayabilecektir.

Genel olarak, Türkiye, ISO/IEC SC 42'deki mevcut rolünü kullanarak küresel en iyi uygulamaları hem etkileyebilir hem de bunlardan faydalanabilir. Böylelikle, YZ teknolojilerinin toplumsal değerleri ve kamu güvenini koruyacak şekilde kullanılmasına katkıda bulunacaktır.

4. Kaynakça

Australian Government Digital Transformation Agency, Australian Government YZ technical standard, Version 1, https://www.digital.gov.au/sites/default/files/documents/2025-08/Australian%20Government%20AI%20technical%20standard.pdf#:~:text=o%20establishing%20a%20bias%20management,in%20societal%20and%20organisational%20culture

CEN-CENELEC JTC 21. (n.d.). *Artificial intelligence — European standardization activities*. European Committee for Standardization and European Committee for Electrotechnical Standardization.

ENISA. (n.d.). YZ *cybersecurity guidelines and threat landscape reports*. European Union Agency for Cybersecurity. https://www.enisa.europa.eu

European Commission. (2021). *Proposal for a regulation laying down harmonised rules on artificial intelligence (Artificial Intelligence Act)* (COM/2021/206 final). Publications Office of the European Union.

European Commission Joint Research Centre. (n.d.). YZ *Watch reports and technical studies*. Publications Office of the European Union.

European Parliament, & Council of the European Union. (2016). *Regulation (EU) 2016/679 of the European Parliament and of the Council of 27 April 2016 on the protection of natural persons with regard to the processing of personal data and on*

the free movement of such data (General Data Protection Regulation). *Official Journal of the European Union*, L 119, 1–88.

European Parliament, & Council of the European Union. (2022). *Regulation (EU) 2022/2065 of the European Parliament and of the Council of 19 October 2022 on a single market for digital services and amending Directive 2000/31/EC (Digital Services Act)*. *Official Journal of the European Union*, L 277, 1–102.

ISO/IEC. (2020). *Artificial intelligence (AI) — Overview of trustworthiness in AI (ISO/IEC TR 24028:2020)*. International Organization for Standardization.

ISO/IEC. (2021). *Artificial intelligence (AI) — Bias in YZ systems and AI-aided decision making (ISO/IEC TR 24027:2021)*. International Organization for Standardization.

ISO/IEC. (2023a). *Information technology — Artificial intelligence — Guidance on risk management (ISO/IEC 23894:2023)*. International Organization for Standardization.

ISO/IEC. (2023b). *Information technology — Artificial intelligence — Management system (ISO/IEC 42001:2023)*. International Organization for Standardization.

ISO/IEC. *Information technology — Artificial intelligence — Trustworthiness of content delivered through IT systems (ISO/IEC CD TS 22443)*. International Organization for Standardization.

KVKK. (2016). *Law on the protection of personal data (Kişisel Verilerin Korunması Kanunu, No. 6698)*. Republic of Türkiye.

Nemko Digital, YZ Regulation in Turkey: Turkish YZ Laws and KVKK Compliance, https://digital.nemko.com/regulations/ai-regulation-in-turkey#:~:text=Tackling%20Discrimination%20and%20Bias

OECD.AI, Catalogue of Tools & Metrics for Trustworthy AI, https://oecd.ai/en/catalogue/tools/isoiec-tr-240272021-information-technology-artificial-intelligence-ai-bias-in-ai-systems-and-ai-aided-decision-making

Turkish Standards Institution (Türk Standartları Enstitüsü). (n.d.). *MTC-195 Artificial Intelligence Mirror Committee*. Turkish Standards Institution.

Turkish Statistical Institute (Türkiye İstatistik Kurumu). (n.d.). *National data infrastructure & statistical guidelines*. Turkish Statistical Institute.

USA National Institute of Standards and Technology (NIST), https://www.nist.gov.


Yapay Zekâ
Süreçlerinde
Hukuk

3. YAPAY ZEKÂ SÜREÇLERİNDE HUKUK

Dr. Burak Görentaş[a], Av. Özge Evci Eralp[b]

[a] *Van Yüzüncü Yıl Üniversitesi, Başkale MYO, burakgorentas@yyu.edu.tr*
[b] *Eralp Avukatlık Bürosu ozge@eralp.av.tr*

3.1. YZ Sistemleri Yaşam Döngüsünde Hukuki Sorumluluklar

3.1.1. Anlaşma Süreçleri (Bölüm 6.1.)

YZ sistemlerinin yaşam döngüsünün başlangıcında yer alan anlaşma süreçleri, sistem sağlayıcıları ile veri temin eden taraflar arasında yapılacak tedarik ve satın alma süreçleri ve bu kapsamda yapılacak sözleşmelerin kapsamını belirler. Bu süreçlerde özellikle tedarikçi seçimi, eğitim verilerinin hukuka uygunluğu ve sıhhati ile ilgili hususlar netleşmektedir.

Bu konuda dikkate alınması gereken ilk mevzuat 6698 sayılı Kişisel Verilerin Korunması Kanunu ("KVKK")'dur. KVKK uyarınca, "veri işleyen, veri sorumlusunun verdiği yetkiye dayanarak onun adına kişisel verileri işleyen gerçek veya tüzel kişiyi" ifade eder. Dış hizmet sağlayıcıdan YZ sistemleri tedarik edildiğinde, hizmet sağlayıcıların veri işleyen niteliği dikkate alınarak, bu sistemlerde işlenen eğitim verilerindeki kişisel verilerin ve aynı zamanda işleme süreçlerinin hukuka uygunluğu KVKK kapsamında değerlendirilmelidir.

KVKK'nın 12. maddesi uyarınca, veri sorumlusu ile veri işleyen, kişisel verilerin güvenliğinin sağlanmasında müştereken sorumludur. Kişisel Verileri Koruma Kurumu'nun veri güvenliği rehberine göre, veri sorumlusunun, hizmet aldığı veri işleyenin en az kendi sağladığı güvenlik düzeyine sahip olduğundan emin olması gerekmektedir. Veri işleyen hizmet sağlayıcı ile veri sorumlusu arasındaki ilişkinin yazılı bir sözleşmeyle düzenlenmesi ve bu sözleşmede, veri işleyenin yalnızca veri sorumlusunun talimatları doğrultusunda hareket edeceği, kişisel verileri yalnızca sözleşmede belirtilen amaç ve kapsamla sınırlı olarak işleyeceği ve ilgili mevzuata uygun davranacağına dair hükümlere yer verilmesi önerilmektedir. Ayrıca, sözleşmede veri işleyenin Kişisel Veri Saklama ve İmha Politikasına uygun hareket edeceğine ve işlediği kişisel verilere ilişkin olarak süresiz sır saklama yükümlülüğü altında olacağına dair

açık hükümler bulunmalıdır. Veri güvenliğinin sağlanması bakımından, sözleşmeye veri ihlali durumunda veri işleyenin derhal veri sorumlusunu bilgilendirme yükümlülüğüne yer verilmesi önem arz etmektedir. Bu yükümlülük, veri sorumlusunun, ihlalin gerçekleştiği anda Kurul'a ve ilgili kişilere zamanında bildirimde bulunabilmesi açısından kritik bir işlev görmektedir. Ayrıca, sözleşmenin niteliği elverdiği ölçüde, veri sorumlusundan veri işleyene aktarılan kişisel verilerin kategori ve türlerinin açık biçimde tanımlanması ve veri işleyenin veri güvenliği önlemlerini alma yükümlülüğünün de belirgin şekilde düzenlenmesi gereklidir. Tüm bunların yanında, veri sorumlusunun, kişisel veri içeren sistemler üzerinde yerinde denetim yapma veya yaptırma, denetim sonuçlarını raporlaştırma ve hizmet sağlayıcı faaliyetlerini izleme yetkisi bulunmalıdır. (Kişisel Verileri Koruma Kurumu, Ocak 2018)

Hizmet ilişkisi kapsamında, barındırılan verilerin yurt dışına aktarımı öngörülüyorsa, bu husus da açıkça düzenlenmeli ve KVKK'nın 9. maddesi kapsamında gerekli yasal süreçler yerine getirilmelidir.

Diğer yandan, YZ sistem sağlayıcılarıyla yapılacak anlaşmalarda, Hizmet Seviyesi Anlaşmaları (SLA) ile tarafların teknik yeterlilikleri, gizlilik yükümlülükleri ve güvenlik önlemleri detaylandırılmalıdır. Sözleşmelerde veri ihlali hâlinde sorumluluk dağılımı ve ihlaller halinde işletilecek cezai hükümler açık biçimde ortaya koymalıdır.

YZ sistemlerine yönelik anlaşma süreçlerinde, KVKK'ya ek olarak, 6098 sayılı Türk Borçlar Kanunu (TBK) hükümleri de kritik bir rol oynamaktadır. Özellikle, hizmet sağlayıcılarla yapılan sözleşmelerin geçerliliği, tarafların irade beyanlarının YZ'ye devredilmesi ve bu süreçte doğabilecek sorumluluklar, TBK'nın genel hükümleri çerçevesinde değerlendirilmelidir. YZ sistemlerinin neden olduğu zararlarda tazmin yükümlülüğü, sistemin sağlayıcısı, geliştiricisi veya kullanıcısı arasında nasıl paylaşılacağı sorusu, TBK'daki haksız fiil ve sözleşmeden doğan sorumluluk kuralları ile ele alınmalıdır. Bu bağlamda, YZ'nin neden olduğu zararın öngörülebilirliği, makine öğrenmesi sürecindeki hatalar ve sistemin otonom karar alma kabiliyeti gibi teknik unsurlar, hukuki sorumluluk tespiti için Borçlar Kanunu kapsamında detaylı bir analiz gerektirmektedir.

Karar süreçlerinde ele alınması gereken diğer bir mevzuat 5846 sayılı Fikir ve Sanat Eserleri Kanunu'dur ("FSEK"). FSEK kapsamında da YZ sistemleriyle çalışan dış hizmet sağlayıcılarla yapılan sözleşmelerin içeriksel niteliği önem arz etmektedir. Özellikle YZ sistemlerinin telif hakkına konu olabilecek eser üretim süreçlerinde kullanılması hâlinde,

üretilen içeriğin eser niteliği taşıyıp taşımadığı, bu içeriğin hak sahipliğinin kimde olduğu ve fikri mülkiyet haklarının nasıl paylaşılacağı sözleşme ile açık şekilde düzenlenmelidir.

FSEK'e göre eser sahibinin izni olmaksızın eserin çoğaltılması, yayılması ve işlenmesi yasaktır. Özellikle yazılım, veri seti ve model çıktıları gibi unsurların telif hakkına tabi olup olmadığına ilişkin belirsizlikler, sözleşme düzeyinde netleştirilmelidir. Ayrıca, taraflar arasında meydana gelebilecek olası telif hakkı ihlallerine karşı tazmin sorumluluğu, sorumluluğun sınırları ve telif ihtilaflarının çözüm yolları sözleşme kapsamında düzenlenmelidir. Bu durum, hem YZ destekli yaratıcı süreçlerin şeffaf ve hesap verebilir olmasını hem de hak sahiplerinin korunmasını temin eden bir hukuki çerçeve sunacaktır. Buna ek olarak, YZ sistemlerinin eğitilmesinde kullanılan veri setlerinin FSEK'e uygunluğu da sözleşme kapsamında açıkça düzenlenmelidir. Özellikle eğitim verileri arasında telif hakkıyla korunan eserlerin yer alması hâlinde, bu eserlerin kullanımına dair lisansların varlığı, kapsamı ve süresi açıkça belirtilmelidir. Aksi hâlde, YZ sistemi tarafından üretilen içeriğin hukuka aykırı bir eğitim sürecine dayanması sebebiyle türev eser niteliğinde olsa dahi telif hakkı ihlali gündeme gelebilir. Bu nedenle, dış hizmet sağlayıcı tarafından sağlanan eğitim verilerinin, FSEK m.21 (çoğaltma), m.22 (yayma) ve m.23 (temsil) gibi maddelerde düzenlenen mali haklara aykırılık oluşturmayacak biçimde lisanslandığının taahhüt edilmesi ve bu durumun sözleşmeye bağlanması gerekmektedir. Sözleşmede ayrıca, veri setlerinde yer alan içeriklerin kaynağı, telif statüsü ve lisans durumu hakkında düzenli raporlama yükümlülüğü ile ihlal hâlinde doğacak sorumluluğun kapsamı da belirlenmelidir.

Öte yandan, Avrupa Birliği tarafından 2024 yılında kabul edilen YZ Tüzüğü (AI Act), özellikle yüksek riskli sistemler açısından teknik uygunluk, veri kalitesi, insan gözetimi ve şeffaflık gibi çok boyutlu yükümlülükler öngörmektedir (European Parliament, 2024). Bu düzenleme kapsamında tedarikçilerin sadece teknik yeterlilikleri değil, aynı zamanda veri işleme süreçlerinin etik ve hukuki uygunluğu da değerlendirme konusu yapılmalıdır. Türkiye her ne kadar bu tüzüğe taraf olmasa da, Brüksel Etkisi nedeniyle (Bradford, 2020) uluslararası iş yapan aktörler açısından YZ Tüzüğü ile uyumlu bir sözleşme altyapısının benimsenmesi stratejik bir gereklilik hâlini almaktadır. YZ Tüzüğü kapsamında, yüksek riskli sistemlerde CE uygunluk işareti, teknik dokümantasyon ve veri kalitesine ilişkin denetim yükümlülükleri doğrudan tedarik zincirinde yer alan veri sağlayıcılarını da etkilemektedir. (European Union, 12 Temmuz 2024)

Buna ek olarak, hizmet sağlayıcılarla yapılacak sözleşmelerde, YZ sisteminin etik uyumluluğunun teminat altına alınması açısından, yalnızca teknik yeterlilik değil, aynı zamanda "algoritmik adalet" ilkelerine uygunluk da güvenceye alınmalıdır. Özellikle karar destek sistemlerinde algoritmanın eğitiminde kullanılan verilerde sistematik önyargı riskine karşı, tarafların veri kalitesi, çeşitliliği ve önyargı azaltma politikalarına dair beyan ve taahhütleri sözleşmeye konu edilmelidir. Aksi hâlde, söz konusu sistemlerin ayrımcı sonuçlar üretmesi durumunda, veri sorumlusu yönünden yalnızca KVKK kapsamında değil, aynı zamanda ayrımcılık yasağına dayalı başka sorumluluk riskleri de doğabilir.

3.1.2. Karar Yönetim Süreci (Bölüm 6.3.3.)

YZ sistemlerinin karar yönetimi süreci, yalnızca teknik işlevsellik ile sınırlı kalmayıp, bireylerin temel hak ve özgürlüklerini koruyacak şekilde yapılandırılmalıdır. Bu süreçte, hem ulusal hem de uluslararası düzenlemelerden doğan etik, hukuki ve yönetsel sorumluluklar dikkate alınarak hareket edilmelidir.

KVKK'nın 4. maddesinde yer alan temel ilkeler (hukuka ve dürüstlük kurallarına uygunluk, belirli – açık ve meşru amaçlar için işlenme, işlendikleri amaçla bağlantılı, sınırlı ve ölçülü olma, doğru ve gerektiğinde güncel olma, ilgili mevzuatta öngörülen veya işlendikleri amaç için gerekli olan süre kadar muhafaza edilme), YZ temelli karar süreçlerinde de geçerlidir. Bu bağlamda sistemin karar alma algoritmaları şeffaf, açıklanabilir ve ölçülü bir temele dayandırılmalı, bireyler açısından anlaşılabilir kılınmalıdır. KVKK'nın 4. maddesinde yer alan genel ilkeler (hukuka uygunluk, ölçülülük, veri minimizasyonu, ilgili ve sınırlı işleme gibi) bu ilkenin dolaylı yansımalarını içerse de, veri sorumlusunun bu uygunluğu belgelendirme ve denetim yoluyla ispat etmesi gerektiğine dair açık bir hüküm bulunmamaktadır. GDPR 35. maddedeki veri koruma etki değerlendirmesine (DPIA) karşılık gelen bir zorunluluk KVKK'da doğrudan yer almaz. Ancak, Kişisel Verileri Koruma Kurumu tarafından yayınlanan Kişisel Veri Güvenliği Rehberi'nde "kişisel verilerin güvenliğinin sağlanması için öncelikle veri sorumlusu tarafından işlenen tüm kişisel verilerin neler olduğunun, bu verilerin korunmasına ilişkin ortaya çıkabilecek risklerin gerçekleşme olasılığının ve gerçekleşmesi durumunda yol açacağı kayıpların doğru bir şekilde belirlenerek buna uygun tedbirlerin alınması gerektiği" belirtilmiştir. (Kişisel Verileri Koruma Kurumu, Ocak 2018). Bu ifade, DPIA kadar net bir süreç olmasa da risk analizi yapmamanın en azından rehber düzeyinde tavsiye edildiğini ve risk değerlendirmesine dayalı önleyici tedbirlerin önemsendiğini göstermektedir.

Bu yaklaşım, OECD YZ İlkeleri ile de örtüşmektedir. OECD'ye göre, güvenilir YZ sistemleri, şeffaflık ve açıklanabilirlik, insan hakları ve demokratik değerler, hesap verebilirlik ve sağlamlık ilkeleri çerçevesinde geliştirilmeli ve uygulanmalıdır (OECD, 2024). Özellikle kararların insan müdahalesine açık olması, demokratik denetimi mümkün kılar ve bireyin sisteme karşı haklarını kullanabilmesine olanak tanır.

Ayrıca, YZ Tüzüğü uyarınca yüksek riskli YZ sistemlerinde kararların kullanıcıya açıklanması, ilgili kişinin sistem kararlarına itiraz edebilmesi ve insan denetimi altında kararların uygulanması yükümlülükleri getirilmiştir (AI Act, 2024). Kararın öngörülebilirliği ve denetlenebilirliği, yalnızca etik ve hukuki gereklilik değil, aynı zamanda güvenli bir YZ yaşam döngüsünün temel taşıdır.

Karar yönetim sürecinde, YZ sistemlerinin kamu ve özel sektördeki kullanımı, sadece teknik yeterlilik veya veri mahremiyeti ilkeleriyle sınırlı kalmamalıdır. Bu süreçte, YZ'nin rekabet dinamiklerine etkisi de göz önünde bulundurulmalıdır. 4054 sayılı Rekabetin Korunması Hakkında Kanun, algoritmik sistemlerin özellikle fiyatlandırma, pazar paylaşımı ve tüketici davranışlarını manipüle etme potansiyeli nedeniyle doğrudan uygulama alanı bulmaktadır. Farklı aktörlerin YZ sistemleri aracılığıyla koordineli hareket etmesi ve "algoritmik kartel" oluşturması riski, rekabet hukukunun temel ilkelerine aykırılık teşkil eder. Bu nedenle, YZ destekli karar destek sistemlerinin tasarımı ve denetiminde, rekabetin korunması ilkesi, sistemin şeffaflığı ve hesap verebilirliğiyle birlikte ele alınmalı, potansiyel rekabet karşıtı etkiler önceden analiz edilmelidir.

Bu çerçevede, organizasyonların karar yönetim süreçlerini, etik, şeffaflık ve hesap verebilirlik ilkelerine dayalı olarak belgelemeleri ve sürdürmeleri zorunludur. Kararların yaşam döngüsü boyunca izlenebilir, kayıt altına alınabilir ve gerektiğinde gözden geçirilebilir olması gereklidir. YZ'den kaynaklanan kararlar, özellikle temel hakları etkileyen alanlarda, iç denetim, dış gözetim ve şikâyet mekanizmalarıyla desteklenmelidir.

Karar yönetim sürecine ilişkin olarak, kamu kurumlarının YZ destekli sistemleri kullanırken yalnızca teknik yeterliliğe değil, aynı zamanda bu sistemlerin yargı denetimine açık olup olmadığına da odaklanması gerekmektedir. Özellikle adli karar destek sistemleri, ceza infaz rejiminde kullanılan risk skorlama yazılımları gibi uygulamalarda, algoritmanın kararı ne ölçüde etkilediği ve insan müdahalesine açık olup olmadığı, adil yargılanma hakkı çerçevesinde değerlendirilmelidir. Bu bağlamda, kararın gerekçelendirilmesi yükümlülüğü yalnızca mahkemeleri değil, kamu yararı gözetilerek kullanılan algoritmik sistemleri de kapsamaktadır.

3.1.3. Risk Yönetim Süreci (Bölüm 6.3.4.)

YZ sistemlerinde risk yönetimi süreci, sistemin geliştirilmesi, dağıtılması ve kullanımı boyunca ortaya çıkabilecek her türlü etik, hukuki, sosyal ve teknik riskin önceden tespit edilmesi, değerlendirilmesi ve uygun tedbirlerle kontrol altına alınmasını amaçlayan sistematik bir çerçeveye dayanır. Bu süreç, yalnızca teknik güvenlik risklerini değil; aynı zamanda önyargılı karar alma, açıklanamaz sonuçlar, şeffaflık eksiklikleri, temel haklara yönelik tehditler ve veri mahremiyeti ihlalleri gibi çok katmanlı riskleri kapsamaktadır.

KVKK madde 12 uyarınca, veri sorumlularına kişisel verilerin hukuka aykırı olarak işlenmesini ve erişilmesini önleme, bu verilerin muhafazasını sağlama yükümlülüğü getirilmiştir. Risk yönetimi, bu yükümlülüğün ayrılmaz bir parçası olarak düşünülmeli ve veri güvenliği tehditlerine karşı önleyici mekanizmaların belirlenmesinde temel dayanak olarak ele alınmalıdır (Kişisel Verileri Koruma Kurumu, 2021). Bir önceki başlıkta belirttiğimiz üzere, ilkelere uyumun veri sorumluları tarafından sağlanması tek başına yeterli olmayacak, ispatı da gerekecektir. Bu husus KVKK'da, GDPR'de olduğu gibi açıkça yer almasa da, Kişisel Verileri Koruma Kurulu Kararları ve rehberler ile kabul edilmiştir. (Kişisel Verileri Koruma Kurulu, 2023)(Kişisel Verileri Koruma Kurumu, Ocak 2018)

YZ Tüzüğü çerçevesinde özellikle yüksek riskli YZ sistemlerinin pazara sunulmadan önce detaylı bir risk yönetim çerçevesine tâbi tutulması zorunludur. Bu kapsamda, sistemin eğitim ve test verilerinin nitelik, bütünlük, tarafsızlık ve amaca uygunluk açısından analiz edilmesi gerekmektedir. Ayrıca, önyargıların azaltılması, veri kalitesi, insan gözetimi, siber güvenlik, dayanıklılık, ve doğruluk gibi teknik unsurların belgelenmesi ve izlenmesi gerekmektedir. Risklerin yalnızca teknik değil aynı zamanda temel haklar ve etik boyutlarda da değerlendirilmesi zorunlu tutulmuştur.

Bu bağlamda, son yıllarda Güney Kore'de yürürlüğe giren SKAIA (South Korea Artificial Intelligence Act) kapsamında yer alan ve politika yapıcılar için model teşkil eden bir uygulama, YZ etki değerlendirmesidir (Güney Kore Ulusal Meclisi, 2023). Bu düzenleme ile yüksek etkili YZ sistemlerinin (EU YZ Tüzüğü'ndeki risk değerlendirmesine benzer nitelikte bir kategorizasyon) etik, güvenlik ve insan hakları boyutlarıyla değerlendirilmesi zorunlu kılınmaktadır. SKAIA uyarınca, YZ işletmeleri, yüksek etkili YZ sistemlerini veya bu sistemleri kullanan ürün ve hizmetleri piyasaya sürmeden önce, insan hakları üzerindeki etkilerini değerlendirmek üzere gerekli önlemleri almak zorundadır. Bu değerlendirme süreci, YZ teknolojilerinin etik, güvenlik ve toplumsal etkilerini analiz ederek olası olumsuz etkileri

en aza indirmeyi amaçlamaktadır. Ayrıca, kamu kurumlarının yüksek etkili YZ içeren ürün veya hizmetleri satın alırken, etki değerlendirmesi yapılmış olanları öncelikli olarak tercih etmesi gerekmektedir. Böylece kamusal alanda YZ kullanımının güvenli ve etik standartlara uygun biçimde yaygınlaştırılması mümkün olabilmektedir (Görentaş, 2025).

Yüksek etkili YZ sistemlerinin insan hakları, etik ve güvenlik boyutlarıyla değerlendirilmesini zorunlu kılan bu düzenlemenin, söz konusu sistemlerin toplumsal düzeyde yol açabileceği olumsuz sonuçların önüne geçilmesini hedeflediği açıktır. Türkiye'de de benzer bir politika önerisi olarak, kamu kurumlarının yalnızca etki değerlendirmesine tabi tutulmuş ürün ve hizmetleri tercih etmesi; böylece YZ teknolojilerinin öngörülebilir etkilerinin (risklerinin) önceden analiz edilerek daha sorumlu, güvenli ve etik ilkelere uygun biçimde kamu politikalarına entegre edilmesi önerilmektedir.

Risk yönetim sürecinin etkinliği açısından, yalnızca teknik izleme sistemleri yeterli olmayıp, etik kurul benzeri bağımsız denetim mekanizmalarının oluşturulması önerilmektedir. Özellikle yüksek etkili YZ sistemleri için, sistemin hayata geçirilmesinden önce bir "etik etki değerlendirme raporu" hazırlanması, bu raporun ilgili paydaşlarca incelenmesi ve kamuoyuyla paylaşılması, yönetişim şeffaflığını ve demokratik meşruiyeti güçlendirecektir. Bu yaklaşım, sadece teknik riskleri değil, sistemin sosyal kabul edilebilirliğini de değerlendirme kapsamına alan bütüncül bir çerçeve sunar.

Risk analizi sürecinde göz önüne alınması gereken bir diğer husus da, sistemin birey üzerindeki psikolojik ve davranışsal etkileridir. Özellikle eğitim, sağlık ve kamu hizmetlerinde kullanılan YZ sistemlerinde, sistem çıktılarının bireyde oluşturabileceği algı, yönlendirme ve davranış değişiklikleri dikkate alınmalı; bireylerin sistem karşısındaki konumları, özerkliğe zarar vermeyecek şekilde kurgulanmalıdır. Bu durum, hem etik hem de Anayasa'nın 17. maddesinde yer alan maddi ve manevi varlığı koruma hakkı çerçevesinde önem arz etmektedir.

3.1.4. Kalite Güvence Süreci (Bölüm 6.3.8.)

YZ sistemlerinde kalite güvence süreci; yalnızca teknik doğruluğun sağlanması değil, aynı zamanda sistemin etik, hukuki ve güvenlik çerçevesinde güvenilir, emniyetli ve öngörülebilir biçimde çalışmasını sağlayan çok boyutlu bir yönetişim mekanizması olarak ele alınmalıdır. Bu süreç, YZ sisteminin yaşam döngüsü boyunca performans, güvenlik ve etik uyum kriterlerinin sürekli olarak izlenmesini, değerlendirilmesini ve iyileştirilmesini kapsar.

YZ Tüzüğü'nün 15. maddesi, özellikle yüksek riskli YZ sistemlerinin "dayanıklılık, doğruluk, siber güvenlik ve hatalara karşı güvenlik" yönlerinden gerekli önlemlerle

donatılmasını zorunlu kılar. Bu kapsamda, sistemlerin normal çalışma koşulları dışında da istikrarını koruyabilmesi, dış müdahalelere karşı korunması ve güvenli kalması hedeflenmektedir. Bu sadece algoritmanın teknik yeterliliğini değil, aynı zamanda beklenmeyen yan etkilerin kontrolünü ve insan üzerindeki olası etkilerin öngörülmesini de içerir. Teknik uyumun sağlanmasında Avrupa standart kuruluşları olan CEN ve CENELEC tarafından geliştirilecek uyumlaştırılmış standartlara atıf yapılmakta, bu standartlara uyum sağlayanların "varsayılan uygunluk" elde edeceği kabul edilmektedir (European Union 12 Temmuz 2024)

KVKK açısından da kalite güvencesi süreci, özellikle veri güvenliğine ilişkin teknik ve idari tedbirlerin sürekliliğiyle ilişkilidir. KVKK madde 12 kapsamında veri sorumlularının kişisel verilerin güvenliğini sağlama yükümlülüğü, sistemin yaşam döngüsüne entegre edilmiş kalite kontrol prosedürleriyle desteklenmelidir. Bu nedenle, kalite güvence yalnızca sistem çıktılarının teknik yeterliliğini değil, kişisel verilerin hukuka uygun şekilde işlenmesini de kapsayan bütüncül bir uyum yaklaşımı ile yürütülmelidir

Bu çerçevede, kalite güvence sürecinin etkin biçimde yürütülebilmesi için bilgi güvenliği yönetim sistemi standardı olan ISO/IEC 27001 ile YZ yönetim sistemine özel olarak geliştirilen ISO/IEC 42001 standardı da dikkate alınmalıdır. ISO/IEC 27001, kişisel verilerin gizliliği, bütünlüğü ve erişilebilirliğini sağlamak amacıyla teknik ve idari kontrolleri tanımlarken; ISO/IEC 42001, YZ sistemlerinin etik, güvenilir ve şeffaf biçimde yönetilmesini temin eden kurumsal yapı, risk yönetimi ve sürekli iyileştirme ilkelerini içermektedir. Her iki standart birlikte ele alındığında, YZ sistemlerinin hem güvenli hem de yasal ve etik açıdan uyumlu şekilde işletilmesi için kapsamlı bir yönetişim çerçevesi sunar.

Sonuç olarak, YZ sistemlerinin kamuya açık ya da kurumsal ortamlarda kullanımı açısından kalite güvence süreçleri, sistem performansının ötesine geçerek güvenilirlik, etik uyum, veri güvenliği ve düzenleyici uygunluk gibi birçok bileşeni kapsayan stratejik bir yönetişim alanı haline gelmiştir.

3.1.5. İş veya Görev Analizi Süreci (Bölüm 6.4.1.)

YZ sistemlerinin kurumsal ortamlarda etkin, güvenilir ve uyumlu biçimde kullanılabilmesi için iş veya görev analizinin titizlikle yürütülmesi gerekmektedir. Bu analiz süreci, YZ uygulamasının hangi amaca hizmet edeceği, hangi görevleri otomatikleştireceği ve insan-makine etkileşiminin sınırlarının nasıl belirleneceği gibi temel sorulara yanıt üretir. Etkin bir görev analizi, sistemin hem fonksiyonel hem de etik çerçevede tasarlanmasına yardımcı

olurken, aynı zamanda risk değerlendirmesine ve hukuki sorumluluk sınırlarının netleştirilmesine de hizmet eder.

Türk hukukunda, KVKK madde 4 uyarınca kişisel verilerin işlenmesinde belirli, açık ve meşru amaç ilkesi esastır. Bu ilke kapsamında, YZ sistemleri ile işlenecek verilerin kullanım amaçlarının önceden tanımlanması ve bu amaçlara uygun şekilde sınırlanması gerekmektedir. İş veya görev analizi, bu bağlamda veri işleme faaliyetlerinin hukuka uygunluk denetimini sağlamada temel bir araçtır.

YZ Tüzüğü madde 14 insan gözetimini temel alan sistem tasarımını zorunlu tutmakta; otomatik karar alma süreçlerinde bireylerin karar mekanizmalarına katılımını dışlamayacak biçimde yapılandırma yapılmasını talep etmektedir. Bu çerçevede, görev analizinde yalnızca teknik yeterlilikler değil, aynı zamanda insan müdahalesinin ne zaman ve nasıl devreye gireceği, sistemin kararlarını kimin sorumluluk üstlenerek onaylayacağı, hangi durumlarda otomatik kararın iptal edileceği gibi konular da netleştirilmelidir (EU YZ Tüzüğü, 2024).

YZ uygulamalarının içerik üretim süreçlerinde kullanılması durumunda Fikir ve Sanat Eserleri Kanunu (FSEK) hükümleri de devreye girmektedir. Özellikle FSEK madde 22-25 arasında yer alan eser sahibinin manevi hakları –eser üzerinde adın belirtilmesi, değiştirilmemesi ve kamuya arz yetkisi gibi– YZ tarafından üretilen içeriklerde insan yaratımı unsurlarının korunmasını zorunlu kılar. Bu nedenle iş analiz sürecinde, YZ'nin içerik oluşturma sürecindeki rolü ile insan katkısının oranı net şekilde tanımlanmalı, mülkiyet hakkı açısından yaratıcı katkı sorumluluğu belirlenmelidir.

Sonuç olarak, YZ sistemlerinin hangi görevleri yerine getireceğine ve bu görevlerin hangi insan eylemlerinin yerini alacağına dair analiz yapılmadan bir sistemin devreye alınması; hem hukuki hem teknik hem de etik açıdan ciddi uyumsuzluklara yol açabilir. Bu nedenle görev analizi yalnızca bir iş planı değil, regülasyonlarla bütünleşik bir sorumluluk matrisidir.

3.1.6. Paydaş İhtiyaçları ve Gereksinimleri Tanımlama Süreci (Bölüm 6.4.2.)

Paydaş ihtiyaçlarının ve gereksinimlerinin tanımlanması, YZ sistemlerinin tasarımı ve uygulanması sürecinin merkezinde yer alır. Bu süreçte, sistemin etkileşimde bulunacağı tüm tarafların -kullanıcılar, geliştiriciler, denetleyiciler ve etkilenen bireylerin- hakları ve beklentileri dikkate alınmalıdır. KVKK madde 10, ilgili kişilerin aydınlatılmasını ve şeffaflık esasına göre bilgilendirilmesini zorunlu kılmaktadır. Bu yükümlülük, yalnızca yasal bir

zorunluluk değil, aynı zamanda etik ve sosyal sorumluluk açısından da temel bir ilkedir. Benzer şekilde, YZ Tüzüğü madde 13'te, kullanıcıya sistemin doğası, yetenekleri ve sınırları hakkında anlaşılır ve erişilebilir bilgi verilmesi gerektiği düzenlenmiştir. Bu sayede, bireylerin sistemle kuracağı etkileşim daha bilinçli ve öngörülebilir hâle gelmektedir.

Paydaş beklentilerinin doğru belirlenememesi, sistem çıktılarının hem hukuken hem de toplumsal olarak sorunlu olmasına yol açabilir. Özellikle sağlık, eğitim, adalet gibi kamuya dokunan sektörlerde, sistemin tasarımı sırasında paydaş katılımı sağlanmadığı takdirde, sistem çıktılarının önyargılı, dışlayıcı veya hukuka aykırı olması riski artar. Bu nedenle, YZ sistemlerinin sorumluluk çerçevesinde tasarlanması, yalnızca teknik değil, çok paydaşlı yönetişim ilkeleri çerçevesinde de ele alınmalıdır.

Ayrıca, paydaş analizinin yalnızca sistem tasarım sürecinde değil, sistemin yaşam döngüsü boyunca dinamik biçimde sürdürülmesi gerekmektedir. Bu bağlamda, geri bildirim mekanizmalarının kurulması, şikâyet ve öneri kanallarının açık tutulması ve bu bildirimlerin periyodik olarak analiz edilerek sistemin gelişimine katkı sağlaması önerilir. Bu yaklaşım, yalnızca kullanıcı memnuniyeti değil, etik hesap verebilirliğin kurumsallaşması açısından da önemlidir.

3.1.7. Bilgi Edinme Süreci (Bölüm 6.4.7.)

YZ sistemlerinin güvenli, şeffaf ve etik kullanımı açısından bilgi edinme süreci yalnızca teknik bilgi aktarımından ibaret olmayıp; kurum içi kültürün, paydaş farkındalığının ve hukuki uyumun geliştirilmesine hizmet eden çok boyutlu bir yapıdır. Bu süreç, sistemin hem geliştiricileri hem de kullanıcıları açısından sürekli öğrenme, güncelleme ve farkındalık oluşturma amacını taşır. Özellikle yüksek riskli sistemlerde bilgi edinmenin sürekliliği ve kapsamlılığı, hem bireysel hakların korunması hem de toplumsal güvenin inşası açısından kritik öneme sahiptir.

KVKK madde 10 uyarınca veri sorumlusu, kişisel verilerin elde edilmesi sırasında ilgili kişileri bilgilendirme yükümlülüğü altındadır. Aydınlatma yükümlülüğü yalnızca teknik detayları açıklamakla sınırlı değildir; veri işleme amaçları, hukuki sebepler, veri aktarımı, ilgili kişinin hakları ve başvuru yolları gibi unsurlar hakkında açık, sade ve erişilebilir bilgi sağlanmasını da içerir. Bu yükümlülük, YZ sistemlerinin kullanıcı etkileşimli çalıştığı durumlarda özellikle önem kazanır.

YZ Tüzüğü kapsamında, özellikle yüksek riskli sistemlerin üreticilerine yönelik olarak sistemin yapısı, kullanım sınırları, potansiyel riskleri ve gözetim gerekliliklerini içeren eğitim materyalleri, kullanım kılavuzları ve karar destek belgelerinin hazırlanması zorunlu kılınmaktadır. Bu yükümlülük, yalnızca ürün güvenliği açısından değil, kullanıcıların sistemin kararlarını sorgulayıp müdahale edebilmeleri için de önemlidir. Kullanıcı, sistemin çalışma prensibini anlamadığı takdirde etik ihlallerin ya da önyargıların farkında olmayabilir.

OECD'nin 2024 yılında güncellediği YZ İlkeleri de bu bağlamda bilgi edinmeyi, hesap verebilirlik, şeffaflık ve insan merkezli sistem tasarımı açısından temel bir ilke olarak değerlendirmektedir (OECD, 2024). Böylece hem kamu hem de özel sektör için bilgi edinme süreçlerinin sadece iç düzenleme değil aynı zamanda uluslararası iyi yönetişim standartlarına uyum aracı olduğu vurgulanmaktadır.

3.1.8. İmha Süreci (Bölüm 6.4.17.)

YZ sistemlerinin yaşam döngüsünde kişisel veri içeren içeriklerin silinmesi, yok edilmesi veya anonim hale getirilmesi, hem etik hem de hukuki açıdan büyük önem arz etmektedir. KVKK'nın 7. maddesi, kişisel verilerin işlenmesini gerektiren sebeplerin ortadan kalkması hâlinde, veri sorumlusu tarafından söz konusu verilerin resen ya da ilgili kişinin talebi üzerine silinmesini, yok edilmesini veya anonim hâle getirilmesini hükme bağlamaktadır .

YZ sistemleri, yüksek hacimli ve hassas verilerle çalıştığı için, özellikle bu sistemlerde kullanılan verilerin süresi dolduğunda uygun yöntemlerle imha edilmesi gerekmektedir. YZ Tüzüğü, genel anlamda veri minimizasyonu ve amaçla sınırlılık ilkeleri kapsamında veri saklama sürelerinin net olarak belirlenmesini ve kullanım amacı ortadan kalktığında verilerin sistematik biçimde silinmesini teşvik etmektedir.

Ayrıca FSEK kapsamında, veri niteliğinde olmasa da sistemin yarattığı içeriklerin telif hakkı ihlali riski taşıması durumunda, bu içeriklerin de erişilemez ve kullanılmaz hâle getirilmesi gerekebilir. Örneğin, eğitim seti olarak kullanılan telifli eserlerin sözleşme süresi sonunda sistemden kaldırılması gerekebilir.

İmha süreci, yalnızca hukuki yükümlülüklerin yerine getirilmesi açısından değil, aynı zamanda kullanıcıların sisteme duyduğu güvenin korunması ve YZ etiği bakımından da kritik önemdedir. Özellikle yüksek riskli YZ sistemlerinde, bu süreçlerin yazılı politikalarla desteklenmesi, düzenli olarak test edilmesi ve denetlenmesi beklenmektedir. İmha işlemleri sırasında yapılan her adımın loglanması, işlem geçmişinin tutulması ve gerektiğinde denetim

raporlarına konu edilmesi, hem şeffaflık hem de hesap verebilirlik ilkeleriyle uyumluluk sağlar (OECD, 2024).

Bu kapsamda, YZ sistemlerinde kişisel veri içeren kayıtların imhası sürecinde log kayıtlarının yalnızca teknik anlamda değil, hukuki bağlamda da denetlenebilir olması sağlanmalıdır. Özellikle kamu kurumlarında kullanılan sistemlerde, imha işlemlerine ilişkin kararların belgeye dayalı biçimde alınması ve bu belgelerin denetime açık tutulması, hem KVKK'nın şeffaflık ilkesiyle hem de 5018 sayılı Kamu Mali Yönetimi ve Kontrol Kanunu çerçevesinde hesap verebilirlik ilkesiyle uyumluluk sağlar.

3.1.9. Özet Tablo

Yaşam Döngüsü Aşaması	Açıklama	İlgili Mevzuat	Temel Hukuki ve Etik Yükümlülükler
1. Anlaşma Süreci	YZ sisteminin tedariki ve taraflar arası sözleşmeler	KVKK m.12, TBK, FSEK, YZ Tüzüğü m.10	Veri işleyenin sorumluluğu, lisans ve telif düzenlemeleri, veri ihlali bildirimi, SLA hükümleri
2. Karar Yönetimi	YZ sisteminin karar süreçlerine etkisi ve insan müdahalesi	KVKK m.4, YZ Tüzüğü m.14, GDPR m.22, OECD İlkeleri	Açıklanabilirlik, insan denetimi, ayrımcılık yasağı, şeffaflık
3. Risk Yönetimi	Etik, hukuki ve teknik risklerin önlenmesi	KVKK m.12, YZ Tüzüğü m.9–10, SKAIA m.35	Etki değerlendirmesi, veri kalitesi, güvenlik önlemleri, insan hakları koruması
4. Kalite Güvence	Performans, güvenlik ve etik uyumun izlenmesi	AI Act m.15, ISO/IEC 27001, ISO/IEC 42001	Dayanıklılık, siber güvenlik, hata toleransı, teknik dokümantasyon
5. Görev Analizi	YZ'nin hangi görevleri üstleneceği ve insanla etkileşim sınırları	KVKK m.4, YZ Tüzüğü m.14, FSEK m.22–25	Otomasyon sınırı, yaratıcı katkı, mülkiyet sorumluluğu

Yaşam Döngüsü Aşaması	Açıklama	İlgili Mevzuat	Temel Hukuki ve Etik Yükümlülükler
6. Paydaş Analizi	Sistemden etkilenen tüm tarafların ihtiyaçlarının belirlenmesi	KVKK m.10, YZ Tüzüğü m.13, OECD İlkeleri	Aydınlatma yükümlülüğü, sistemin yapısı ve sınırları hakkında bilgilendirme
7. Bilgi Edinme	Kullanıcı ve geliştiricilere bilgi sunumu ve eğitim	KVKK m.10, YZ Tüzüğü, OECD İlkeleri	Kullanım kılavuzu, risk bilgilendirmesi, dokümantasyon gereklilikleri
8. İmha Süreci	Veri süresi dolduğunda silme, yok etme veya anonimleştirme	KVKK m.7, YZ Tüzüğü, FSEK	Veri işleme amacının ortadan kalkması halinde, kişisel verilerin uygun yöntemlerle imha edilmesi; eğitim verileri arasında telif koruması olan içeriklerin sözleşme süresi sonunda sistemden çıkarılması

3.1.10. Kaynakça

Kişisel Verileri Koruma Kurumu [KVKK]. (2016). *6698 sayılı Kişisel Verilerin Korunması Kanunu*. Resmî Gazete, 29677.

Kişisel Verileri Koruma Kurumu. (Ocak 2018). *Veri Güvenliği Rehberi*. Kişisel Verileri Koruma Kurumu. Erişim tarihi 31.07.2024, https://www.kvkk.gov.tr/yayinlar/veri_guvenligi_rehberi.pdf

Kişisel Verileri Koruma Kurumu (Ocak 2018). *Kişisel Veri Güvenliği Rehberi (Teknik ve İdari Tedbirler)*. Kişisel Verileri Koruma Kurumu. Erişim 8 Temmuz 2025 https://www.kvkk.gov.tr/yayinlar/veri_guvenligi_rehberi.pdf

Kişisel Verileri Koruma Kurulu. (2023). *Tüketici Finansman Kredisiyle Alışveriş İmkânı Sunan Şirket Tarafından İlgili Kişilerden e-Devlet Şifrelerinin Talep Edilmesi hakkında 22/03/2023 tarihli ve 2023/426 sayılı Karar Özeti*. Erişim 8 Temmuz 2025, https://www.kvkk.gov.tr/Icerik/7599/2023-426

Bradford, A. (2020). The Brussels effect: How the European Union rules the world. Oxford University Press. https://scholarship.law.columbia.edu/books/232/

European Union (12 Temmuz 2024). *Regulation (EU) 2024/1689 of the European Parliament and of the Council of 13 June 2024 laying down harmonised rules on artificial intelligence and amending Regulations (EC) No 300/2008, (EU) No 167/2013, (EU) No 168/2013, (EU) 2018/858, (EU) 2018/1139 and (EU) 2019/2144 and Directives 2014/90/EU, (EU) 2016/797 and (EU) 2020/1828 (Artificial Intelligence Act)*, Erişim tarihi 8 Temmuz 2025. https://eur-lex.europa.eu/eli/reg/2024/1689/oj

Güney Kore Ulusal Meclisi. (2023). *YZ Temel Yasası (SKAIA)*. Erişim tarihi: 1 Ağustos 2025, https://likms.assembly.go.kr/bill/billDetail.do?billId=PRC_R2V4H1W1T2K5M1O6E4Q9T0V7Q9S0U0

Görentaş, M. B. (2025). Güney Kore YZ Yasası'nın Analizi ve Türk Hukukuna Bir Model Olarak Değerlendirilmesi. *Çukurova Üniversitesi Hukuk Araştırmaları Dergisi*, 7, 413–441.

OECD. (2024). *OECD YZ Principles (Updated 2024)*. Retrieved from https://oecd.ai/en/dashboards/ai-principle

European Union. (2024). *Regulation (EU) 2024/1689 of the European Parliament and of the Council laying down harmonised rules on artificial intelligence (AI Act)*. Official Journal of the European Union.

3.2. KVKK, GDPR ve AB YZ Yasası'nın (AI Act) YZ Uygulamaları Kapsamında Karşılaştırılması

Küresel düzeyde dijitalleşme ve veri ekonomisinin yükseldiği bir çağda, bireylerin kişisel verilerinin korunması gerek temel bir insan hakkı gerekse de bilişim temelli hizmetlerin güvenliği için merkezi bir konuma oturmuştur. Teknolojik gelişmeler, özellikle büyük veri analitiği, nesnelerin interneti ve YZ sistemleri, kişisel veri işleme faaliyetlerini dönüştürmüş, bu alanı yalnızca veri güvenliği değil aynı zamanda etik, şeffaflık ve algoritmik hesap verebilirlik bağlamında da yeniden ele almayı gerekli kılmıştır.

Bu kapsamda Türkiye'de 7 Nisan 2016 tarihinde yürürlüğe giren 6698 sayılı Kişisel Verilerin Korunması Kanunu (KVKK), Avrupa Birliği'nde ise 4 Mayıs 2016 tarihinde Avrupa Birliği (AB) Resmi Gazetesi'nde yayımlanan ve 25 Mayıs 2018 tarihinde uygulanmaya başlayan Genel Veri Koruma Tüzüğü (GDPR) ile temel veri koruma normları oluşturulmuştur. Ancak bu iki düzenleme de doğrudan YZ sistemlerinin regülasyonuna odaklanmamakta olup veri koruma alanında düzenlemeler yapmaktadır.

YZ teknolojilerinin öngörülemez karar alma yapıları, veri temelli öğrenme modelleri ve otomatikleştirilmiş işlem kapasitesi karşısında Avrupa Birliği'nde 2021 yılından itibaren yeni bir normatif çerçeve hazırlıkları başlamıştır. Avrupa Komisyonu, Nisan 2021'den itibaren yaptığı çalışmalarda Avrupa Birliği'nin ilk YZ yasasını önererek risk temelli bir sınıflandırma sistemi öngörmüştür. Bu sistemde YZ uygulamaları kullanıcıya oluşturduğu risk seviyesine göre analiz edilmekte ve uyulması gereken yükümlülükler bu risk seviyesine göre farklılık göstermektedir. Bu kapsamda, Avrupa Birliği tarafından hazırlanan ve 12 Temmuz 2024 tarihinde Avrupa Resmi Gazetesi'nde yayımlanan YZ Tüzüğü (AI Act), dünya çapında YZ sistemlerine ilişkin ilk kapsamlı hukuki düzenleme olarak dikkat çekmektedir. (European Parliament, 2023)

3.2.1. Türkiye'de Kişisel Verilerin Korunması Kanunu ve Uygulaması

KVKK, 7 Nisan 2016 tarihinde Resmî Gazete'de yayımlanarak yürürlüğe girmiştir. KVKK, 95/46/EC sayılı AB Veri Koruma Direktifi esas alınarak hazırlanmış (Kişisel Verileri Koruma Kurumu, 2020) ve "hukuka ve dürüstlük kurallarına uygun olma, doğru ve gerektiğinde güncel olma, belirli, açık ve meşru amaçlar için işlenme, işlendikleri amaçla bağlantılı, sınırlı ve ölçülü olma, İlgili mevzuatta öngörülen veya işlendikleri amaç için gerekli olan süre kadar muhafaza edilme" temel ilkeleri benimsemiştir. Kanun kapsamında verilen yetkileri kullanmak üzere Kişisel Verileri Koruma Kurumu ve Kişisel Verileri Koruma Kurulu kurulmuştur.

KVKK, "sadece Türkiye'de gerçekleşen veri işleme faaliyetlerini düzenleyip düzenlenmediğine", yani "mülkilik ilkesine" ilişkin bir hüküm içermemektedir. Ancak, Kişisel Verileri Koruma Kurulu'nun uygulamaları, sınırların "etki ilkesine" göre genişletildiğini göstermektedir. Nitekim, Kişisel Verileri Koruma Kurulu'nun 24.01.2019 tarihli ve 2019/10 sayılı "Kişisel Veri İhlali Bildirim Usul ve Esaslarına İlişkin Kararı"nda (Kişisel Verileri Koruma Kurumu, 29.01.2019), bu yaklaşım açık biçimde ortaya konmuştur. Bu kararda, "veri ihlalinin yurtdışında yerleşik veri sorumlusu nezdinde yaşanması halinde, bu ihlalin

sonuçlarının Türkiye'de yerleşik ilgili kişileri etkilemesi ve ilgili kişilerin sunulan ürün ve hizmetlerden Türkiye'de faydalanmaları durumunda, bu veri sorumlusu tarafından da aynı esaslar çerçevesinde Kurul'a bildirimde bulunulmasına" karar verilmiştir. Söz konusu karar, KVKK'nın yalnızca Türkiye'de faaliyet gösteren veri sorumlularını değil, aynı zamanda yurtdışında yerleşik veri sorumlularının ihlallerini de kapsayabildiğini ortaya koymaktadır. Zira, yurtdışında gerçekleşen bir veri ihlalinin Türkiye'de yerleşik ilgili kişileri etkilemesi halinde, bu ihlalin de Kurul'a bildirilmesi gerektiği açıkça belirtilmiştir. Bu durum, KVKK uygulamasında yalnızca mülkilik ilkesine değil, aynı zamanda etki ilkesine de yer verildiğini göstermektedir. (Kişisel Verileri Koruma Kurumu, Ocak 2025)

3.2.2. Avrupa Birliği Veri Koruma Tüzüğü (GDPR) ve Uygulaması

GDPR, 95/46/EC sayılı AB Veri Koruma Direktifi'ni yürürlükten kaldırarak 27 Nisan 2016 tarihinde kabul edilmiş ve 25 Mayıs 2018 tarihinde tüm AB üyelerinde doğrudan uygulanabilir hale gelmiştir (European Data Protection Supervisor, t.y.). GDPR, "veri taşınabilirliği", "unutulma hakkı", "profil oluşturma", "otomatik karar verme" gibi yenilikçi kavramlar getirmiştir. GDPR'yle birlikte Avrupa Veri Koruma Kurulu kurulmuş, cezalar yıllık cironun %4'üne kadar belirlenmiştir.

Avrupa Birliği Genel Veri Koruma Tüzüğü (GDPR), yer bakımından uygulanabilirliği, GDPR'nin 3. maddesi ile açık bir şekilde düzenlenmiştir. Söz konusu maddeye göre, GDPR yalnızca AB içinde kurulu veri sorumlularının veya veri işleyenlerin faaliyetlerine değil, aynı zamanda AB dışındaki veri sorumlularının, eğer AB'de bulunan gerçek kişilere mal veya hizmet sunuyorsa ya da bu kişilerin davranışlarını izliyorsa, bu işleme faaliyetlerine de uygulanmaktadır.

3.2.3. Avrupa Birliği Yapay Zekâ Yasası (AI Act)

YZ Yasası / Tüzüğü, Avrupa Komisyonu tarafından 21 Nisan 2021 tarihinde taslak olarak sunulmuş, 13 Mart 2024 tarihinde Avrupa Parlamentosu tarafından kabul edilmiş ve kademeli yürürlük tarihleri belirlenmiştir. YZ Tüzüğü, GDPR gibi, Avrupa Parlamentosu ve Konseyi tarafından çıkarılmış bir tüzük (regulation) niteliğindedir. Bu düzenleme, tüm AB üye devletlerinde doğrudan ve bağlayıcı olarak uygulanmaktadır (European Union, 12 Temmuz 2024). YZ Tüzüğü, veri koruma ile sınırlı olmayıp, YZ sistemlerinin risk esaslı sınıflandırılmasını öngörmekte ve teknik sistemlere doğrudan müdahale etmektedir. YZ Tüzüğü'nün temel amacı, insan merkezli, güvenilir ve sürdürülebilir YZ'yi teşvik etmek; bireylerin temel hak ve özgürlüklerine, özellikle de kişisel verilerin korunması hakkına saygı

göstermektir. Bu düzenlemenin ana hedeflerinden biri, "yüksek riskli YZ sistemlerinin" geliştirilmesi, uygulanması ve kullanımı sürecinde ayrımcılık ve önyargının önlenmesidir. (Stefano De Luca, Marina Federico, 2025)

YZ Tüzüğü'nün 2. Maddesi, tüzüğün yer bakımından uygulanma alanını ("territorial scope") belirlemektedir. Madde 2/1 uyarınca YZ Tüzüğü "AB içinde pazara sunulan YZ sistemleri, AB içinde kullanılacak olan YZ sistemleri, AB dışında yerleşik olmalarına rağmen, AB pazarına YZ ürünü sunan ya da bu pazarda YZ sistemini kullanan gerçek/tüzel kişiler" için uygulanır.

Bu düzenleme ile, YZ Tüzüğü, tıpkı GDPR gibi, sadece mülkilik ilkesine değil aynı zamanda etki ilkesine dayanmaktadır. Yani, AB dışında geliştirilmiş bir YZ sistemi, eğer AB vatandaşlarını etkiliyor ya da AB pazarında kullanılıyorsa, YZ Tüzüğü hükümlerine tabidir. Bu sayede düzenleme, sınır ötesi etkisi olan YZ faaliyetlerini denetim altına alma amacı taşır.

Bu düzenleme, YZ sistemlerini kullanıcıya yönelik oluşturduğu risk seviyesine göre sınıflandırmakta ve risk düzeyine göre farklı hukuki yükümlülükler öngörmektedir.

Kabul edilemez risk kategorisinde değerlendirilen YZ uygulamaları (örneğin sosyal skorlama sistemleri, çocukları tehlikeli davranışlara yönlendiren sesli oyuncaklar, kamuya açık alanlarda gerçek zamanlı yüz tanıma) tamamen yasaklanmış; ancak ciddi suçların soruşturulması gibi durumlarda gerçek zamanlı uzaktan biyometrik tanımlamaya sınırlı istisnalar tanınmıştır.

YZ Tüzüğü, diğer bir risk kategorisinde olan yüksek riskli YZ sistemlerini iki ana gruba ayrılmıştır: AB ürün güvenliği mevzuatına tabi sistemler (örneğin tıbbi cihazlar, otomotiv sistemleri) ve kamu yararına ilişkin özel kullanım alanlarındaki sistemler (örneğin kritik altyapı, eğitim, istihdam, sınır güvenliği, hukuk uygulamaları). Tüzük uyarınca, bu sistemlerin piyasaya sürülmeden önce ve kullanım süresi boyunca uygunluk değerlendirmeleri ile denetlenmesi gerekmektedir. Üretken YZ sistemleri tek başına model çeşidine göre yüksek riskli kategoride değerlendirilmemektedir; ancak yine de şeffaflık yükümlülükleri kapsamındadır. Bu kapsamda, içeriklerin YZ tarafından üretildiğinin açıkça belirtilmesi, yasa dışı içerik üretimini önlemeye yönelik teknik önlemlerin alınması ve kullanılan telifli eğitim verilerinin özetinin yayımlanması zorunludur. Bununla birlikte, GPT-4 gibi sistemik risk oluşturabilecek düzeyde gelişmiş genel amaçlı YZ modelleri, ayrı ve daha katı bir denetim sürecine tabi tutulmaktadır. Bu modellerin daha kapsamlı teknik değerlendirmelerden geçmesi ve ciddi olayların Avrupa Komisyonu'na bildirilmesi gerekmektedir.

Bunun yanında, YZ Tüzüğü, belirlediği sınırlı riskli YZ sistemleri kategorisi için özellikle şeffaflık yükümlülükleri öngörmektedir. Bu kapsamda, kullanıcıların bir YZ sistemi ile etkileşimde bulunduklarını anlayabilmeleri için gerekli bilgilendirmenin yapılması zorunludur. Örneğin, sohbet robotları (chatbots) veya YZ ile üretilmiş görsel ve işitsel içeriklerde (deepfake) kullanıcılara, içeriklerin YZ tarafından oluşturulduğunun açıkça belirtilmesi gerekmektedir. Böylece, bireylerin yanıltılmasının önüne geçilmekte ve temel hakların korunması sağlanmaktadır.

Diğer taraftan, minimal riskli YZ sistemleri, insan hakları veya güvenlik bakımından ihmal edilebilir düzeyde tehdit oluşturan uygulamalardır. Bu nedenle YZ Tüzüğü, bu kategorideki sistemlere herhangi bir özel hukuki yükümlülük getirmemektedir. E-posta hizmetlerinde spam filtreleme mekanizmaları bu gruba örnek olarak gösterilebilir. Bu yaklaşım, düzenleyicinin inovasyonu engellemeden yalnızca temel haklar ve güvenlik açısından riskli alanlara odaklandığını ortaya koymaktadır

Parlamentonun bir diğer hedefi ise Avrupa'daki YZ inovasyonunu desteklemek olup, bu çerçevede genç girişimciler (start-up'lar) ve KOBİ'ler için gerçek dünyaya yakın test ortamları sağlanması düzenlenmiştir. YZ Tüzüğü'nün bazı bölümleri (örneğin yasaklı sistemlere ilişkin hükümler) 2 Şubat 2025'te yürürlüğe girmiş olup, üretken YZ sistemleri için öngörülen şeffaflık yükümlülükleri yürürlük tarihinden 12 ay sonra, yüksek riskli sistemlere ilişkin yükümlülükler ise 36 ay sonra uygulanacaktır (European Parliament, 08.06.2023).

3.2.4. KVKK ve GDPR'nin YZ'ye Uygulanabilirliği Açısından Karşılaştırması

YZ teknolojileri, otomatik karar alma, profil çıkarma ve büyük veri analizleri gibi bireylerin temel hak ve özgürlüklerini doğrudan etkileyen uygulamalar içerdiğinden, kişisel verilerin korunması mevzuatlarıyla doğrudan ilişkilidir. Bu bağlamda, KVKK ve GDPR'nin YZ sistemlerine uygulanabilirliği, her iki düzenlemenin kapsamı, ilkeleri ve denetim mekanizmaları üzerinden karşılaştırmalı olarak değerlendirilmelidir.

1. **Profil Oluşturma Tanımı Açısından Karşılaştırma**

GDPR Madde 4, kişisel veri, işleme ve özellikle profil oluşturma kavramlarını ayrıntılı biçimde tanımlamaktadır. Özellikle Madde 4/4 uyarınca, profil oluşturmayı; "bir bireye ilişkin kişisel verilerin analiz edilmesi yoluyla bireyin davranış, ilgi alanı veya performansı gibi yönlerinin değerlendirilmesi" olarak tanımlanır. YZ kavramı, GDPR'de doğrudan

tanımlanmasa da, mevcut profil oluşturma tanımı üzerinden YZ sistemlerini kapsaması mümkündür.

KVKK'da GDPR'ye benzer bir "profil oluşturma" tanımı yer almasa da, Kişisel Verileri Koruma Kurulu, kararları aracılığıyla uygulamaya yön vermektedir. Kişisel Verileri Koruma Kurulu'nun 23/12/2021 tarihli ve 2021/1303 sayılı kararı, özellikle profilleme niteliğindeki veri işleme faaliyetleri bakımından dikkat çekici bir değerlendirme sunmaktadır. Kurul, söz konusu kararda, araç kiralama sektöründeki bir uygulamanın profilleme niteliğinde olduğunu tespit etmiştir. Karara konu uygulamada, müşterilerin teslim süresi, ödeme davranışı ve hasar geçmişi gibi verileri analiz edilerek "kara liste" adı altında bir sınıflandırma yapılmakta ve bu liste diğer firmalarla paylaşılmaktadır. Bu durum, geçmiş davranışlara dayalı otomatik karar alma süreçleri oluşturulması anlamına gelmektedir. Kurul, profillemeyi "bireyin davranışlarının analiz edilerek bunu izleyen davranışları hakkında tahminlerde bulunmak adına kişisel verilerin otomatik olarak işlenmesi" olarak tanımlamıştır. Kurul, profillemenin her zaman olumsuz bir eylem olmadığını belirtmekle birlikte, bu süreçte bireylerin temel hak ve özgürlüklerinin tehdit edilmemesi gerektiğini vurgulamıştır. Kurul kararına göre, veri sorumluları, KVKK'nın 4. maddesinde yer alan genel ilkelere (hukuka ve dürüstlük kurallarına uygunluk, doğru ve güncel olma, belirli ve meşru amaçlar için işlenme, amaçla bağlantılı, sınırlı ve ölçülü olma, saklama süresi) uygun hareket ettiği ve Kanun'un 5. maddesinde sayılan hukuka uygunluk şartlarından en az birine dayandığı sürece profilleme ve otomatik karar verme sistemlerini kullanma hakkına sahiptir. Bu bağlamda, "kara liste" gibi profillemeye dayalı uygulamalar dahi, söz konusu ilkelere ve hukuka uygunluk şartlarına dayandığı sürece mümkün olabilir. Özellikle açık rıza dışındaki veri işleme şartlarından biri olan "veri sorumlusunun meşru menfaati"ne dayanılması durumunda, veri sorumlularının dikkatli bir değerlendirme yapması gerekmektedir. Bu süreçte şeffaflık, öngörülebilirlik ve hesap verebilirlik ilkelerine uyulmalı, ilgili kişinin hak ve özgürlüklerinin zarar görmemesi için gerekli teknik ve idari tedbirler alınmalı, işleme faaliyeti veri minimizasyonu ve amaçla sınırlılık ilkelerine uygun yürütülmelidir. (Kişisel Verileri Koruma Kurulu, 2021) Kişisel Verileri Koruma Kurulu'nun bu yaklaşımı, Avrupa Veri Koruma Kurulu'nun (EDPB) profilleme ve otomatik karar alma süreçlerine ilişkin yorumlarıyla da uyumludur. (European Data Protection Board, 17 Aralık 2024).

2. Temel İlkelere Uyum ve Etki Değerlendirmesi Açısından Karşılaştırma

Kişisel verilerin korunmasına ilişkin düzenlemeler hem GDPR hem de Türkiye'de yürürlükte olan KVKK kapsamında belirli ilkelere dayandırılmaktadır. Her iki düzenleme de veri işleme faaliyetlerinin şeffaflık, veri minimizasyonu, amaca bağlılık, sınırlılık ve ölçülülük

gibi temel ilkelere uygun yürütülmesini zorunlu kılar. Ancak GDPR m.5/2'de açıkça yer alan "hesap verilebilirlik" ilkesi, KVKK'da doğrudan yer almamaktadır. GDPR m.5/2'ye göre, veri sorumlusu yalnızca kişisel verileri ilkelere uygun şekilde işlemekle değil, aynı zamanda bu uygunluğu kanıtlayabilmekle yükümlüdür. GDPR m.35'e göre, özellikle yeni teknolojiler kullanılırken ve işlemenin mahiyeti, kapsamı, bağlamı ve amaçları dikkate alındığında bir işleme türünün gerçek kişilerin hakları ve özgürlükleri açısından yüksek bir riske sebebiyet vermesinin muhtemel olduğu hâllerde, veri sorumlusu, işleme öncesinde, öngörülen işleme faaliyetlerinin kişisel verilerin korunması üzerindeki etkisine ilişkin bir veri koruma etki değerlendirmesi (DPIA) yapar. Bu değerlendirme, planlanan işlemenin niteliğini, amacını, risklerini ve bu risklere karşı alınacak önlemleri içermelidir. EDPB'nin 28/2024 sayılı görüşü de bu çerçevede, YZ sistemlerinin eğitimi ve kullanımı süreçlerinde DPIA'nın sadece bir yükümlülük değil, aynı zamanda hesap verilebilirliğin temel bileşeni olduğunu vurgulamıştır. (European Data Protection Board, 17 Aralık 2024).

KVKK'da hesap verilebilirlik ilkesi, GDPR'deki kadar açık ve doğrudan düzenlenmemiştir. KVKK'nın 4. maddesinde yer alan genel ilkeler (hukuka uygunluk, ölçülülük, veri minimizasyonu, ilgili ve sınırlı işleme gibi) bu ilkenin dolaylı yansımalarını içerse de, veri sorumlusunun bu uygunluğu belgelendirme ve denetim yoluyla ispat etmesi gerektiğine dair açık bir hüküm bulunmamaktadır. Aynı şekilde, GDPR madde 35'teki DPIA'ya karşılık gelen, yüksek riskli veri işleme faaliyetleri için önleyici risk değerlendirme zorunluluğu da KVKK'da doğrudan yer almaz. Ancak, Kurum tarafından yayınlanan Kişisel Veri Güvenliği Rehberi'nde DPIA düzeyinde olmasa da bi risk analizi yapılmasını "kişisel verilerin güvenliğinin sağlanması için öncelikle veri sorumlusu tarafından işlenen tüm kişisel verilerin neler olduğunun, bu verilerin korunmasına ilişkin ortaya çıkabilecek risklerin gerçekleşme olasılığının ve gerçekleşmesi durumunda yol açacağı kayıpların doğru bir şekilde belirlenerek buna uygun tedbirlerin alınması gerektiği" ifadesiyle önermiştir. (Kişisel Verileri Koruma Kurumu, Ocak 2018).

Kişisel Verileri Koruma Kurulu da, doğrudan "hesap verilebilirlik" terimini kullanmasa da bazı kararlarında belgelendirme eksikliği ve önleyici tedbir alınmaması gerekçeleriyle veri sorumlularını sorumlu tutmuştur. Örneğin, 22/03/2023 tarihli ve 2023/426 sayılı Kurul Kararında bir tüketici finansman şirketinin, müşterilerden e-Devlet şifresi talep ederek çeşitli verilere erişim sağladığı şikâyeti incelenmiş, Kurul, "kişisel verilerin güvenliğinin sağlanması için öncelikle veri sorumlusu tarafından işlenen tüm kişisel verilerin neler olduğunun, bu verilerin korunmasına ilişkin ortaya çıkabilecek risklerin gerçekleşme olasılığının ve

gerçekleşmesi durumunda yol açacağı kayıpların doğru bir şekilde belirlenerek buna uygun tedbirlerin alınması gerektiği, bu riskler belirlenirken kişisel verilerin özel nitelikli kişisel veri olup olmadığı, mahiyeti gereği hangi derecede gizlilik seviyesi gerektirdiği, güvenlik ihlali halinde ilgili kişi bakımından ortaya çıkabilecek zararın niteliği ve niceliğinin dikkate alınması gerektiği, bu risklerin tanımlanması ve önceliğinin belirlenmesinden sonra söz konusu risklerin azaltılması ya da ortadan kaldırılmasına yönelik kontrol ve çözüm alternatiflerinin maliyet, uygulanabilirlik ve yararlılık ilkeleri doğrultusunda değerlendirilmesi ve gerekli teknik ve idari tedbirlerin planlanarak uygulamaya konulması gerektiği" yönünde karar tesis edilmiştir. (Kişisel Verileri Koruma Kurulu, 2023)

Sonuç olarak, GDPR'de hesap verilebilirlik, sistematik belgelendirme ve DPIA gibi mekanizmalarla desteklenen normatif bir yükümlülük iken, KVKK'da bu yükümlülük doğrudan düzenlenmemiş, ancak rehberler ve özellikle Kurul kararları yoluyla uygulamaya dönük olarak tesis edilmiştir.

3. Veri Sahibinin Hakları Açısından Karşılaştırma

GDPR, doğrudan "yapay zekâ" ifadesini içeren düzenlemeler içermese de, 22. maddesiyle "yalnızca otomatik işleme süreçlerine dayalı karar alma" mekanizmalarını düzenlemektedir. Bu kapsamda, GDPR madde 22/1'e göre "ilgili kişi, kendisi ile ilgili hukuki sonuçlar doğuran veya kendisini benzer şekilde önemli ölçüde etkileyen, profilleme dâhil, yalnızca otomatik işleme dayalı bir karara tabi tutulmama hakkına" sahiptir. Aynı şekilde 22/3 maddesinde veri sahibinin hakları olarak "veri sorumlusu tarafından insan müdahalesi elde etme, kendi görüşünü ifade etme ve karara itiraz etme hakkı" sayılmıştır.

KVKK düzenlemesi de doğrudan "yapay zekâ" terimini içermemekle birlikte, kişisel verilerin otomatik sistemlerle işlenmesi ve bu işleme dayalı karar alma süreçlerine dair hüküm içermektedir. KVKK'nın 11. maddesinin 1. fıkrasının (g) bendi, ilgili kişiye "kişisel verilerin münhasıran otomatik sistemler vasıtasıyla analiz edilmesi suretiyle aleyhe bir sonucun ortaya çıkmasına itiraz etme" hakkını tanımaktadır. Bu düzenleme, GDPR'deki 22. maddeye benzer şekilde, yalnızca otomatik işleme dayalı ve kişi aleyhine sonuç doğuran kararların sınırlandırılmasını amaçlamaktadır. Ancak KVKK'da, GDPR'deki gibi açıkça "hukuki sonuç doğuran ya da kişiyi benzer şekilde önemli ölçüde etkileyen kararlar" şeklinde bir tanımlama yapılmamış, daha dar ve "aleyhe sonuç" çerçevesi tercih edilmiştir. Ayrıca, KVKK'da GDPR madde 22/3'te olduğu gibi insan müdahalesi talep etme, görüş bildirme veya itiraz hakkı gibi detaylı güvencelere açıkça yer verilmemiştir.

Kişisel Verileri Koruma Kurumu tarafından yayınlanan YZ Alanında Kişisel Verilerin Korunmasına Dair Tavsiyeler Rehberinde, YZ sistemlerinin geliştirilmesi ve kullanımı sürecinde veri sorumlularının uyması gereken temel ilkeler sistematik şekilde ele alınmış; bu sistemlerin şeffaflık, hesap verebilirlik, veri minimizasyonu, ayrımcılığın önlenmesi ve insan müdahalesi hakkı çerçevesinde hukuka uygun olması gerektiği vurgulanmıştır. Rehber, yalnızca teknik uyumluluğu değil, etik sorumluluğu da merkezine alarak bireylerin temel haklarının YZ karşısında zayıflamaması gerektiğine dikkat çekmiş ve özellikle çocuklar ile hassas gruplara yönelik YZ uygulamalarında daha yüksek koruma önermiştir (Kişisel Verileri Koruma Kurumu, Nisan 2025).

4. Veri Kümesi Oluşturma, Anonimleştirme ve Takma Ad (Pseudonymisation) Açısından karşılaştırma

GDPR, anonimleştirme ve takma ad kullanımını (Pseudonymisation) açıkça birbirinden ayırır. Anonimleştirme, kişisel verilerin "artık kimliği belirli veya belirlenebilir bir gerçek kişiyle ilişkilendirilememesi" hâlidir ve bu durumda GDPR hükümleri uygulanmaz. Ancak bu anonimlik mutlak olmalıdır: veri, makul yollarla başka verilerle eşleştirilse bile kimlik tespiti mümkün olmamalıdır.

Buna karşılık, takma ad kullanımı (pseudonymisation) GDPR'de özel olarak "kişisel verilerin, ek bilgi kullanılmaksızın belirli bir veri sahibine atfedilemeyecek şekilde işlenmesi anlamına gelir, ancak söz konusu ek bilginin ayrı tutulması ve kişisel verilerin belirli veya belirlenebilir bir gerçek kişiye atfedilmesini önlemek için teknik ve organizasyonel tedbirlere tabi olması gerekir" şeklinde tanımlanmıştır. Bu durumda veri hâlâ kişisel veri sayılır ve GDPR uygulanmaya devam eder, ancak risk azaltıcı bir önlem olarak değerlendirilir.

EDPB'nin 2024 tarihli görüşü, YZ sistemlerinin eğitimi sürecinde kullanılan veri kümelerinin takma adla işlense dahi hâlâ kişisel veri kapsamında olduğunu ve bu veriler için Veri Koruma Etki Değerlendirmesi (DPIA) yapılması gerektiğini belirtmektedir (European Data Protection Board, 17 Aralık 2024)

KVKK, anonimleştirmeyi 3. maddesinde "kişisel verilerin başka verilerle eşleştirilse dahi hiçbir şekilde bir gerçek kişiyle ilişkilendirilemeyecek hâle getirilmesi" şeklinde tanımlar. KVKK'da pseudonymisation kavramı açıkça tanımlanmamıştır, ancak bazı Kurul kararları ve rehberlerde dolaylı biçimde bu tekniğe değinildiği görülmektedir. Nitekim, Kişisel Verileri Koruma Kurulu, araştırma şirketleri tarafından yapılan telefonla mülakat süreçlerinde, açık rıza verilmemesi hâlinde dahi, görüşmenin tarih ve saati, arayan ve aranan numaralar gibi log

verilerinin ve ses kaydının, takma adlandırma (Pseudonymisation yöntemiyle iki yıl süreyle saklandığını tespit etmiştir. Kurul, bu verilerin anonim hâle getirilmediği için kişisel veri niteliğini koruduğunu ve bu nedenle 6698 sayılı Kanun hükümlerinin uygulanması gerektiğini belirtmiştir. Ayrıca bu tür işlemlerin, Kanun'un 28/1-b maddesinde yer alan "anonim hâle getirilmek suretiyle yapılan istatistiksel veri işleme" istisnası kapsamında değerlendirilemeyeceğini açıkça ortaya koymuştur (Kişisel Verileri Koruma Kurumu, 26.08.2024)

5. Veri Taşınabilirliği Hakkı Bakımından Karşılaştırma

GDPR'nin 20. maddesiyle düzenlenen veri taşınabilirliği hakkı, veri koruma rejiminde dönüştürücü bir yenilik olarak öne çıkmaktadır. Anılan düzenleme, ilgili kişilere kişisel verilerini yapılandırılmış, yaygın olarak kullanılan ve makine tarafından okunabilir bir formatta elde etme ve bu verileri herhangi bir engelle karşılaşmadan başka bir veri sorumlusuna aktarma imkânı tanımaktadır. Bu hak, yalnızca bireylerin verileri üzerindeki özerkliklerini artırmakla kalmamakta; aynı zamanda dijital piyasalarda rekabetin güçlenmesine ve hizmet sağlayıcılar arasında dengeleyici bir ortamın oluşmasına da hizmet etmektedir.

Veri taşınabilirliğinin önemi, özellikle YZ uygulamaları bağlamında daha belirgin hale gelmektedir. Zira, veri taşınabilirliği, YZ ekosisteminde rekabetin korunmasında kritik bir araç olarak öne çıkmaktadır. Bu hak, özellikle büyük teknoloji platformlarının veri siloları yaratarak kullanıcıları bu ekosistemlere hapsolmuş halde bırakması riskine karşı bir denge mekanizması sunar. Veri taşınabilirliği, dijital hizmetler arasında bilgi akışını destekleyerek kullanıcıları güçlendirir ve geçiş maliyetlerini düşürerek rekabeti teşvik eder (Reimsbach-Kounatze, C. and A. Molnar, 2024) Dolayısıyla, veri taşınabilirliği hakkı, yalnızca bireysel özerklik ve tüketici hakları bakımından değil; aynı zamanda dijital pazarlarda hak temelli rekabetin tesisi açısından da stratejik bir rol oynamaktadır.

KVKK'da ise, veri taşınabilirliği hakkı yer almamaktadır. Bu durum, özellikle YZ ekosisteminde yenilikçiliği ve rekabeti sınırlayan bir boşluk yaratmaktadır. Avrupa Birliği'nde GDPR madde 20 ile güvence altına alınan bu hak, KVKK'da yer almadığından, bireylerin veri üzerindeki tasarruf yetkisi kısıtlı kalmaktadır.

3.2.5. Yapay Zekâ Uygulamaları Bağlamında GDPR, KVKK ve YZ Tüzüğü Karşılaştırması

YZ uygulamaları bağlamında kişisel veri koruma rejimi, hâlihazırda hem Türkiye'de yürürlükte olan KVKK hem de Avrupa'da uygulanan GDPR ile belirli ölçüde düzenlenmektedir. Her iki metin de kişisel verilerin işlenmesi bakımından şeffaflık, veri minimizasyonu, amaca uygunluk, sınırlılık ve güvenlik gibi temel ilkelere dayanır (KVKK, m.4; GDPR, m.5). Ancak bu normatif altyapı, YZ'ye özgü ürün güvenliği, model şeffaflığı veya veri yönetişimi gibi teknik ve yönetsel boyutları doğrudan ele almaz. Ancak, GDPR ve YZ Tüzüğü farklı amaç ve kapsamlarla ortaya çıkmış olsalar da bazı temel ilkeleri paylaşmaktadır. Bunlar arasında şeffaflık, hesap verilebilirlik, güvenlik ve insan haklarının korunması yer alır. Ayrıca, her iki mevzuatta da bireyin temel haklarının korunması önceliklidir.

1. **Amaç ve Kapsam Bakımından Farklılık**

GDPR ve KVKK'nın temel amacı, bireylerin kişisel verilerinin işlenmesi süreçlerinde temel hak ve özgürlüklerinin korunmasıdır. GDPR madde 1/2 ve KVKK madde 1, bu odağı doğrudan ortaya koyar.

YZ Tüzüğü ise, sistem odaklı bir düzenlemedir. YZ Tüzüğü madde 2 ve madde 3'e göre, tüzük; YZ sistemlerinin tasarımı, geliştirilmesi, piyasaya sunulması ve kullanımını kapsar. Odak, sistemin kendisinden kaynaklanabilecek risklerdir. (European Union 12 Temmuz 2024). Bu nedenle YZ Tüzüğü, verinin niteliğinden bağımsız olarak, tüm YZ sistemlerini kapsayabilir. Örneğin, YZ Tüzüğü uyarınca sistemler risk düzeyine göre "yasaklı", "yüksek riskli", "sınırlı riskli" veya "minimal/düşük riskli" olarak sınıflandırılır. Buna göre yüksek riskli sistemler için yalnızca kişisel veri değil; eğitim verisi kalitesi, sistem mimarisi, belgelendirme ve insan gözetimi gibi çok daha geniş yükümlülükler söz konusudur (European Union 12 Temmuz 2024).

GDPR kişisel verilerin korunmasına odaklanan, hak temelli ve yatay bir düzenleme iken; YZ Tüzüğü, YZ sistemlerinin teknik işleyişi ve risk düzeylerine göre sınıflandırılmasına dayalı sistem odaklı ve dikey bir yaklaşımla yapılandırılmıştır. YZ Tüzüğü, AB ürün güvenliği hukukundan esinlenerek, sistemlerin piyasaya sunulmadan önce risk temelli uygunluk değerlendirmelerine tabi tutulmasını öngörürken; GDPR bireyin verisi üzerindeki kontrolünü esas almaktadır. Bu durum, GDPR'nin veri odaklı yaklaşımı ile YZ Tüzüğü'nün sistem ve işlev odaklı yaklaşımı arasındaki temel farkı net biçimde yansıtmaktadır. (Veale & Zuiderveen Borgesius, 2021)

2. Teknik Standartlar ve Uyumluluk Gereklilikleri

GDPR ve KVKK, veri işleme ilkeleri ile sınırlı olup, herhangi bir model eğitimi, veri kümesi kalitesi, genel amaçlı YZ sistemleri (GPAI) gibi kavramları tanımlamaz. Buna karşın YZ Tüzüğü, temel modellerin (foundation models) eğitimi, açıklanabilirlik, çekişmeli (adversarial) testler ve sistem belgeleri gibi çok daha derin teknik yükümlülükler içerir (AI Act, m.50–56).

Örneğin, büyük dil modelleri artık "genel amaçlı YZ sistemi" sayılmakta olup bu sistemler için belge sunumu, açıklanabilirlik raporları ve telif hakkı riski değerlendirmesi zorunludur (AI Act, m.53). YZ Tüzüğü kapsamında, sistem geliştiricilerinin risk yönetimi, veri kalitesi, teknik dokümantasyon, şeffaflık, insan gözetimi, dayanıklılık ve doğruluk gibi çok boyutlu gerekliliklere uyması zorunludur. Özellikle Madde 10 ve Ek IV'te yer alan hükümler, eğitim ve test veri kümelerinin tarafsız, eksiksiz ve amaca uygun olması gerektiğini belirtirken, Madde 13 ve 14 şeffaflık ve insan gözetimini zorunlu kılar. Ayrıca YZ sistemlerinin dayanıklılığı ve siber güvenliği Madde 15 ile güvence altına alınmakta; tüm yüksek riskli sistemler için CE uygunluk işareti ve önceden yapılmış uygunluk değerlendirmesi şart koşulmaktadır. Teknik uyumun sağlanmasında Avrupa standart kuruluşları olan CEN ve CENELEC tarafından geliştirilecek uyumlaştırılmış standartlara atıf yapılmakta, bu standartlara uyum sağlayanların "varsayılan uygunluk" elde edeceği kabul edilmektedir (European Union 12 Temmuz 2024).

3. Etki Değerlendirmesi ve Risk Yönetimi Yaklaşımı

GDPR madde 35 kapsamında yer alan Veri Koruma Etki Değerlendirmesi (DPIA) yalnızca kişisel verilerin yüksek riskli işlenmesinde öngörülür. KVKK'da DPIA benzeri bir sistem mevzuatta yer almamaktadır ancak rehberlerde veri sorumlusunun riskleri değerlendirmesi gerektiği belirtilmiştir (Kişisel Verileri Koruma Kurumu, Ocak 2018; bakınız yukarıda 5.2). Buna karşın, YZ Tüzüğü, DPIA'ya ek olarak, sistemlerin kendisine ilişkin riskleri tespit eden ve teknik dokümantasyon ile desteklenen kapsamlı değerlendirmeler öngörür.

GDPR Madde 32, veri işleme faaliyetlerinde yer alan veri sorumluları ve işleyenlerin, kişisel verilerin gizliliğini, bütünlüğünü ve erişilebilirliğini sağlamak, sistem sürekliliğini güvence altına almak amacıyla, uygun teknik ve organizasyonel önlemleri (şifreleme, takma adlandırma, sızma testleri vb.) almakla yükümlü olduğunu belirtir. Benzer şekilde, KVKK Madde 12, veri sorumlularının kişisel verilerin hukuka aykırı işlenmesini, erişilmesini ve

muhafazasını engelleyecek "uygun güvenlik düzeyine ilişkin her türlü teknik ve idari tedbir" almasını zorunlu kılar. Ancak bu hükümler, büyük ölçüde veri odaklı olup, YZ sistemlerinin teknik mimarisine, eğitim sürecine, algoritmik açıklanabilirliğe veya sistemsel risk testlerine yönelik somut mühendislik standartları sunmamaktadır. Buna karşılık, YZ Tüzüğü (özellikle Madde 50–56) temel modeller (foundation models) için adversarial testler, model belgeleri, sistem performans raporları, açıklanabilirlik yükümlülükleri ve telif hakkı risk değerlendirmesi gibi teknik uyumluluk kriterlerini açık biçimde düzenlemektedir.

4. **Telif Hakları ve Eğitim Verisi Kullanımı**

YZ Tüzüğü, büyük dil modelleri gibi temel modellerin eğitimi sırasında kullanılan verilerin telif haklarına tabi olabileceğini kabul etmekte ve bu hususta şeffaflık sağlamayı zorunlu kılmaktadır. YZ Tüzüğü madde 53'e göre, genel amaçlı yapay zekâ (GPAI) sağlayıcıları, eğitim veri setlerinde telif hakkına konu olabilecek içeriklerin kullanımına dair açıklamalar sunmakla yükümlüdür. Buna karşılık, GDPR ve KVKK, telif hakkı gibi fikrî mülkiyet meselelerini veri koruma hukukunun bir parçası olarak değerlendirmemekte, yalnızca kişisel veri niteliği taşıyan içerikler bakımından düzenleme getirmektedir.

3.2.6. Sonuç

YZ teknolojilerinin hızla gelişmesi, yalnızca veri güvenliği meselesini değil, aynı zamanda algoritmik karar süreçlerinde şeffaflık, hesap verebilirlik ve ayrımcılığın önlenmesi gerekliliğini de gündeme getirmektedir. Bu bağlamda incelenen üç düzenleme farklı yönelimler sunmaktadır. KVKK, etki ilkesi sayesinde sınır ötesi veri işleme faaliyetlerini kapsayarak Türkiye'de ulusal ölçekte bireylerin verilerini koruma işlevi görmektedir. GDPR, küresel ölçekte hak temelli yaklaşımıyla bireylere güçlü güvenceler sağlamakta, şeffaflık ve bireysel kontrol ilkelerini ön plana çıkarmaktadır. YZ Tüzüğü ise teknolojik risk temelli yapısıyla, özellikle yüksek riskli sistemlerde güvenlik, ayrımcılığın önlenmesi ve şeffaflık ekseninde sistematik bir uygunluk çerçevesi ortaya koymaktadır.

Türkiye bakımından bu karşılaştırma, KVKK'nın özellikle YZ uygulamaları açısından GDPR uygulamalarını da gözeterek güncellenmesinin önemine işaret etmektedir. Özellikle GDPR'de düzenlenmiş olan ilgili kişinin hakları, profil oluşturma, etki değerlendirmesi, veri taşınabilirliği hakkı gibi kavramlar YZ uygulamalarında yüksek önem taşımaktadır.

Diğer yandan, mevcut düzenlemelerin YZ alanında uygulanması, rehberler ve sektörel çalışmalar ile detaylandırılmalı ve YZ Tüzüğü örneği de dikkate alınarak küresel düzenlemelere uyum sağlanmalıdır.

Sonuç olarak, KVKK, GDPR ve YZ Tüzüğü'nün farklı perspektifleri birbirini tamamlayan bir bütünlük sunmakta; gelecekte daha bütünleşik, hak odaklı ve teknolojiye uyumlu bir veri koruma rejiminin inşası için temel taşları oluşturmaktadır. Bu bütünleşme, yalnızca Türkiye ve Avrupa Birliği için değil, küresel ölçekte de YZ hukukunun gelişiminde belirleyici rol oynayacaktır.

3.3. Yapay Zekâ'nın Hukuki ve Etik Açıdan Denetlenmesine Yönelik Politika Önerileri

YZ sistemlerinin hukuki ve etik denetlenebilirliğini sağlamak, sadece teknik güvenlik önlemleriyle sınırlı kalmayan; aynı zamanda bireyin temel hak ve özgürlüklerini güvence altına alan çok katmanlı bir yönetişim yaklaşımını gerektirir. Türkiye'de 6698 sayılı Kişisel Verilerin Korunması Kanunu (KVKK), temel veri koruma ilkelerini belirlemekle birlikte, Avrupa Birliği Genel Veri Koruma Tüzüğü (GDPR) ile uyumun sağlanması, ilgili kişilerin profil çıkarma, otomatik karar verme ve algoritmik ayrımcılığa karşı korunması açısından elzemdir. KVKK'nın 4. maddesi uyarınca veri işleme faaliyetlerinin şeffaf, amaçla sınırlı ve KVKK'nın 10. maddesi uyarınca ilgili kişiyi bilgilendirici nitelikte olması gerekirken; bu ilkelerin YZ sistemlerinin tüm yaşam döngüsüne entegre edilmesi, etkili bir yasal uyum mekanizması açısından zorunludur. Buna karşılık, Avrupa Birliği'nin YZ Tüzüğü (AI Act) düzenlemesi, YZ sistemleri için risk temelli yükümlülükler getirerek algoritmik hesap verebilirliği kurumsallaştırmayı hedeflemektedir. Güney Kore ise YZ Tüzüğü'nden farklı olarak, öncelikle kamu otoriteleri arasındaki görev dağılımını ve yönetişim yapısını kurumsallaştıran, stratejik ve yönlendirici nitelikte bir temel yasa (YZ Temel Yasası-SKAIA) benimsemiş; böylece düzenleyici altyapıyı şekillendirmeden önce kamu kapasitesini güçlendirmeyi amaçlamıştır. (Güney Kore Ulusal Meclisi, 2023)

ABD'de ise bağlayıcı bir federal yasa yerine, sektör bazlı gönüllü standartlar (örneğin NIST YZ Risk Framework) ve etik ilkeler rehberliği benimsenerek, kurumların hesap verebilirlik ve iç denetim kapasitesi artırılmaya çalışılmaktadır. Birleşik Krallık ise "pro-innovation" yaklaşımı çerçevesinde, regülasyonlardan çok rehber dokümanlarla yönlendirme yapan esnek bir model inşa etmiş, düzenleyici kurumların kendi sektörlerine özgü çerçeveler geliştirmesine olanak tanımıştır. Bu yaklaşımlar birlikte değerlendirildiğinde, Türkiye'nin de YZ sistemlerinin denetlenebilirliğini artırmak için yalnızca teknik değil, kurumsal, hukuki ve etik katmanları birlikte ele alan çok paydaşlı ve aşamalı bir yönetişim mimarisi kurması gerektiği açıktır.

Dolayısıyla Türkiye'de, yalnızca KVKK'nın revizyonu değil; aynı zamanda YZ'ye özgü etik kurallar, rehberler, ikincil düzenlemeler ve teknik yeterlilik çerçevelerinin hayata geçirilmesi gerekmektedir. Bu bağlamda, gerek ISO/IEC 42001 gibi YZ yönetim sistemlerine ilişkin standartlar gerekse OECD YZ İlkeleri doğrultusunda hesap verebilirlik, izlenebilirlik ve insan merkezlilik temelli bir denetim kültürü inşa edilmelidir.

3.3.1. KVKK ve GDPR Uyumlaştırması

1. İlgili Kişi Haklarının Güçlendirilmesi

GDPR madde 22/1, ilgili kişilere "kendisi ile ilgili hukuki sonuçlar doğuran veya kendisini benzer şekilde önemli ölçüde etkileyen, profilleme dâhil, yalnızca otomatik işleme dayalı bir karara tabi tutulmama hakkı" tanımaktadır. GDPR'nin 22/3 maddesi ise veri sahibinin haklarını detaylandırarak, "veri sorumlusu tarafından insan müdahalesi elde etme, kendi görüşünü ifade etme ve karara itiraz etme hakkı" gibi somut güvenceler sunar. Bu maddeler, otomatik sistemlerin şeffaflığını ve hesap verebilirliğini artırmayı hedeflerken, bireylerin algoritmik kararlar karşısında savunmasız kalmasını önleyici mekanizmalar sunmaktadır.

KVKK'nın 11. maddesinin 1. fıkrasının (g) bendi, ilgili kişiye "kişisel verilerin münhasıran otomatik sistemler vasıtasıyla analiz edilmesi suretiyle aleyhe bir sonucun ortaya çıkmasına itiraz etme" hakkını tanır. Bu düzenleme, GDPR'deki 22. maddeye benzer bir amaç gütmekle birlikte, kapsamı itibarıyla daha dar bir çerçeve çizmektedir. Özellikle GDPR'deki "hukuki sonuç doğuran ya da kişiyi benzer şekilde önemli ölçüde etkileyen kararlar" gibi geniş bir tanımlama yerine, KVKK'da "aleyhe sonuç" kavramına odaklanılmıştır. Ayrıca, KVKK'da GDPR madde 22/3'te yer alan insan müdahalesi talep etme, görüş bildirme veya karara itiraz etme gibi detaylı güvencelere açıkça yer verilmemesi, Türk mevzuatının bu alandaki koruma düzeyinde bir farklılık yaratmaktadır. Bu durum, YZ temelli otomatik karar alma süreçlerinde ilgili kişilerin haklarının KVKK kapsamında GDPR'ye kıyasla daha sınırlı yorumlanabileceği bir potansiyel doğurmaktadır. Bu güvencelerin KVKK değişiklikleri ile tesisi gerekmektedir.

2. Veri Taşınabilirliği Hakkı

Sayısallaşan dünyada bireylerin kişisel verileri üzerindeki kontrolü ve bu verilerin serbestçe dolaşımı giderek daha fazla önem kazanmaktadır. GDPR, bu konuda önemli bir adım atarak, Madde 20 ile bireylere veri taşınabilirliği hakkını tanımıştır. Bu hak, ilgili kişilere kendilerine ait kişisel verileri, yapılandırılmış, yaygın olarak kullanılan ve makine tarafından okunabilir bir formatta alma ve bu verileri başka bir veri sorumlusuna engelsiz bir şekilde aktarma olanağı sunar.

YZ sistemlerinin giderek yaygınlaştığı günümüzde bu hak çok daha stratejik bir boyut kazanmıştır. Zira YZ modelleri, büyük ölçüde bireylerin sanal dünyadaki sayısal etkileşimlerinden, tercihlerinden, sağlık ve finansal geçmişlerinden elde edilen verilerle beslenmektedir. Eğer bireyler bu verilerini alternatif hizmet sağlayıcılarına taşıyamazlarsa, mevcut sağlayıcıların veri üzerindeki hâkimiyetleri artar ve kullanıcıların daha yenilikçi ya da kişiselleştirilmiş hizmetlerden yararlanma imkânı ciddi biçimde sınırlanır. Bu durum yalnızca bireysel özerklik bakımından değil, aynı zamanda rekabet hukuku ve sayısal pazarların sağlıklı işleyişi açısından da risk taşır (Reimsbach-Kounatze & Molnar, 2024).

Veri taşınabilirliği hakkı, YZ çağında şu açılardan kritik rol oynamaktadır:

- ***Algoritmik Hesap Verebilirlik ve Şeffaflık:*** Bireylerin, YZ sistemleri tarafından kendileri hakkında kullanılan verileri alma ve inceleme yeteneği, algoritmaların nasıl çalıştığına dair daha fazla şeffaflık sağlar. Veri sahipleri, verilerinin nasıl işlendiğini ve kararların nasıl alındığını daha iyi anlayabilirler.
- ***Hizmet Değişikliği ve Rekabet*:** Veri taşınabilirliği, bireylerin bir YZ hizmetinden diğerine kolayca geçiş yapmasını mümkün kılar. Bu, sağlayıcılar arasında rekabeti teşvik eder, inovasyonu artırır ve bireylerin kendi ihtiyaçlarına en uygun hizmeti seçmelerine olanak tanır.
- ***Veri Özerkliği ve Kontrol:*** YZ sistemleri tarafından oluşturulan "profil" veya "skor" gibi sonuçlar da veri sahibinin kişisel verisi olarak kabul edilebilir. Veri taşınabilirliği, bireylerin bu tür türetilmiş verileri de kontrol etme ve taşıma yeteneğini güçlendirir.
- ***Yenilikçiliğin Teşviki:*** Yeni YZ girişimleri ve *start-up*'lar için, veri havuzlarına erişim genellikle bir engel teşkil eder. Veri taşınabilirliği, bireylerin verilerini yeni ve yenilikçi hizmetlere aktarmasına izin vererek bu engeli hafifletebilir.

KVKK ise, GDPR'de açıkça düzenlenen veri taşınabilirliği hakkına ilişkin doğrudan bir hüküm içermemektedir. Bu eksiklik, Türkiye'deki bireylerin dijital haklarının korunması ve hizmet sağlayıcılar arası geçiş kolaylığı açısından önemli bir boşluk yaratmaktadır. Özellikle YZ uygulamalarının günlük yaşamda giderek daha fazla yer almasıyla, veri taşınabilirliği hakkının KVKK'ya entegre edilmesi zaruri hale gelmiştir. Bu entegrasyon, sadece bireylerin dijital özerkliğini güçlendirmekle kalmayacak, aynı zamanda Türkiye'nin dijital ekonomisinde sağlıklı rekabetin ve inovasyonun desteklenmesine de katkıda bulunacaktır. KVKK'da

yapılacak bir düzenleme, bireylerin kendi sayısal ayak izlerini yönetme kapasitesini artırarak, YZ destekli sistemlerin daha adil ve şeffaf bir şekilde işlemesine zemin hazırlayacaktır.

3. **Veri Koruma Etki Değerlendirmesi (DPIA)**

GDPR madde 35 uyarınca, kişisel veri işleme faaliyetinin bireylerin hak ve özgürlükleri üzerinde yüksek risk oluşturması hâlinde veri sorumlusunun, işlemeye başlamadan önce "Veri Koruma Etki Değerlendirmesi" (Data Protection Impact Assessment – DPIA) yapması zorunludur. Bu değerlendirme, sistemin doğurabileceği olumsuz sonuçların önceden belirlenmesini, gerekli teknik ve idari tedbirlerin alınmasını ve gerekirse denetim otoritesine danışılmasını içerir. YZ sistemleri ise doğaları gereği otomatik karar alma, ayrımcılık, şeffaflık eksikliği, denetlenemezlik ve mahremiyet riskleri içerdiğinden, bu tür sistemlerin uygulamaya konulmadan önce DPIA süreçlerinden geçmesi, GDPR kapsamında zorunludur.

KVKK'da ise DPIA'ya ilişkin açık bir düzenleme bulunmamaktadır. Her ne kadar Kişisel Verileri Koruma Kurumu'nun çeşitli rehber dokümanlarında (örneğin Veri Güvenliği rehberi, Kişisel Verileri Koruma Kurumu, Ocak 2018) ve kararlarında (örneğin 22/03/2023 tarihli ve 2023/426 sayılı Kurul Kararı) DPIA benzeri risk değerlendirmelerine atıf yapılsa da, yasal düzeyde bağlayıcı bir düzenleme bulunmamaktadır. Bu eksiklik, yüksek riskli YZ sistemlerinin birey hakları üzerindeki etkilerini önceden değerlendirmeyi engellemekte ve veri güvenliği açısından önemli bir zafiyet oluşturmaktadır.

Bu nedenle KVKK'ya, GDPR madde 35'e benzer şekilde, yüksek riskli veri işleme faaliyetleri için DPIA yükümlülüğünü açıkça düzenleyen bir madde eklenmeli; YZ uygulamaları da bu kapsamda özellikle tanımlanmalıdır. Ek olarak, kamu otoritelerinin onay süreçlerine tabi kritik sistemlerde, DPIA'nın bir uzmana veya bağımsız kuruma sunulması zorunlu kılınabilir.

3.3.2. AB YZ Tüzüğü ile Uyumlaştırma, Kademeli Geçiş ve Rehber Politikalar

YZ Tüzüğü, YZ sistemlerinin risk derecelerine göre, özellikle yüksek riskli YZ sistemleri için sıkı kurallar getirirken, sistem sağlayıcılarına şeffaflık, veri kalitesi, dokümantasyon, insan gözetimi ve siber güvenlik alanlarında yükümlülükler getirmektedir. Ancak Türkiye'de, doğrudan YZ Tüzüğü'ne benzer bir yasal düzenlemeye geçmeden önce, sektörel bazda rehberler ve izleme sistemleri geliştirilerek kademeli geçiş yapılmalıdır. Bu konuda öneriler aşağıda verilmektedir.

1. Düzenleyici Korumalı Alan (Regulatory Sandbox) Uygulamaları

YZ teknolojilerinin hızla ilerlemesi, beraberinde çeşitli hukuki ve etik riskleri de getirmektedir. YZ Tüzüğü, bu riskleri yönetmek için yenilikçi bir araç olarak "Regulatory Sandbox" (düzenleyici korumalı alan) uygulamasını hayata geçirmeyi öngörmektedir. Bu uygulama, YZ sistemlerinin gerçek dünya koşullarına yakın ama kontrollü ve denetlenebilir ortamlarda test edilmesine olanak tanır. Böylece, potansiyel hukuki ve etik riskler, sistemler geniş çapta kullanılmaya başlamadan önce belirlenebilir, değerlendirilebilir ve giderilebilir. Bu korumalı alanlar, geliştiricilere yasal sınırlamalar içinde yenilik yapma esnekliği sunarken, regülatörlere de YZ'nin karmaşık doğasını daha iyi anlama ve uygun düzenleyici çerçeveler oluşturma fırsatı tanır.

Türkiye'de de YZ uygulamalarının sağlıklı ve güvenli bir şekilde gelişmesi için benzer korumalı alan (“sandbox”) modellerinin oluşturulması büyük önem taşımaktadır. Bu süreçte, Bilgi Teknolojileri ve İletişim Kurumu, Siber Güvenlik Başkanlığı, Kişisel Verileri Koruma Kurumu ve Sanayi ve Teknoloji Bakanlığı gibi kilit kurumların koordinasyonu kritik rol oynayacaktır. Bu kurumların işbirliğiyle kurulacak YZ korumalı alanları, özellikle yüksek riskli YZ sistemlerinin test edilmesine odaklanmalıdır. Test süreçlerinde, aşağıdakiler gibi temel risk alanlarına yoğunlaşılmalıdır:

- Ayrımcılık (Bias)
- Telif Hakları
- Güvenlik
- Veri Gizliliği ve Mahremiyet
- Şeffaflık ve Açıklanabilirlik

Bu korumalı alan modelleri, yalnızca teknik test alanları olmanın ötesinde, YZ etiği ve hukuk alanındaki tartışmalar için de bir laboratuvar görevi görecektir. Elde edilecek bulgular, Türkiye'nin YZ stratejilerinin ve yasal düzenlemelerinin daha sağlam temellere oturmasına yardımcı olacak, böylece hem inovasyon desteklenecek hem de bireylerin hakları etkili bir şekilde korunacaktır.

Ulusal fonlar artırılarak bağımsız bir finansman modeli oluşturulmalı, çip ve yarı iletken teknolojilerinde dışa bağımlılığı azaltacak yatırımlar yapılmalı ve kamu-özel sektör işbirlikleri derinleştirilmelidir.

2. Sektörel Rehberler

YZ teknolojilerinin hızla ilerlemesi, beraberinde getirdiği potansiyel hukuki ve etik risklerin etkin yönetilmesini zorunlu kılmaktadır. YZ Tüzüğü, bu risklere bir yanıt olarak, "Code of Practice" (Uygulama Esasları veya İyi Uygulama Kodu) gibi yol gösterici belgeler yayımlamayı öngörmektedir. (Avrupa Komisyonu, 2023) Bu tür rehberler, YZ geliştiricileri ve kullanıcıları için bir yol haritası sunarak, gelecekteki olası yasal yükümlülüklere hazırlanmalarını sağlamakta ve inovasyonu etik sınırlar içinde teşvik etmektedir. Türkiye'nin de YZ ekosisteminin sağlıklı ve sorumlu bir şekilde gelişimi için benzer bir yaklaşımla sektörel rehberler yayımlaması, stratejik bir önem arz etmektedir.

Türkiye'de yayımlanacak sektörel YZ rehberleri, AB'deki "code of practice" modeline paralel olarak, yasal olarak bağlayıcı olmamalı ancak ilgili paydaşlara yol gösteren destekleyici ve tamamlayıcı araçlar olarak işlev görmelidir. Bu rehberler, YZ uygulamalarının tüm yaşam döngüsü boyunca dikkat edilmesi gereken etik ilkeleri, en iyi uygulamaları ve risk yönetim stratejilerini detaylandırmalıdır. Bu sayede, hem YZ teknolojilerini geliştiren aktörler hem de bu teknolojileri kullanan kurum ve kuruluşlar, gelecekte yürürlüğe girebilecek daha kapsamlı düzenlemelere uyum sağlayabilirler.

Rehberlerin sektörel şekilde ayrıştırılarak kapsaması gereken temel başlıklar aşağıdaki gibi sıralanabilir:

- İnsan Denetimi (Human Oversight)
- Algoritmik Açıklanabilirlik (Explainability)
- Kayıt Tutma Yükümlülüğü (Record Keeping)
- Eğitim Verilerinin Şeffaflığı ve Kalitesi (Data Transparency and Quality)
- Risk Değerlendirmesi ve Azaltımı
- Şikayet Mekanizmaları ve Hakların Kullanımı
- Yayımlanma Süreci ve Etkinliği İçin Öneriler
- YZ geliştirme ve temin süreçlerinde yerine getirilmesi gerekenler

Bu sektörel rehberlerin yayımlanma süreci, ilgili kamu kurumlarının yanı sıra, üniversiteler, teknoloji şirketleri, sivil toplum kuruluşları ve sektör dernekleri gibi tüm ilgili paydaşların katılımıyla şeffaf ve kapsayıcı bir şekilde yürütülmelidir. Bu, rehberlerin hem teorik olarak sağlam temellere oturmasını hem de pratik olarak uygulanabilirliğini sağlamak açısından elzemdir. YZ düzenleyici sürecin etkinliği için sürekli yasal araştırmalar yürütülmeli

ve kapsayıcı politikalar geliştirmek amacıyla kamu, özel sektör, akademi ve sivil toplumdan görüşler alınmalıdır.

Bu rehberler, YZ teknolojilerinin Türkiye'de sorumlu bir şekilde geliştirilmesi ve dağıtılması için temel bir yol haritası sunacaktır. Yasal bağlayıcılığı olmasa da, sektörde "iyi uygulamalar" standartlarını belirleyerek, şirketlerin uyum süreçlerini hızlandıracak ve Türkiye'yi YZ çağında daha rekabetçi ve güvenli bir konuma taşıyacaktır.

3.3.3. Yüksek Riskli YZ Sistemlerinin Tespiti ve Sorumluluklarının Belirlenmesi

İnsan hayatı, temel haklar ve güvenlik üzerinde ciddi etkisi olabilecek enerji, sağlık, adalet, biyometrik veri kullanımı gibi alanlarda kullanılan YZ sistemleri, yüksek riskli veya yüksek etkili YZ sistemleri olarak tanımlanmalıdır. Bu sistemleri kullanan işletmelere güvenlik ve güvenilirlik standartlarına uyma, düzenli testler ve güncellemeler yapma yükümlülüğü getirilmelidir. Ancak bu tür sistemlere yönelik sorumluluk çerçevesi oluşturulurken, Avrupa Birliği'nin YZ Tüzüğü düzenlemesindeki gibi teknik ayrıntılarla tanımlanmış, doğrudan bağlayıcı ve katı bir risk sınıflandırmasından ziyade; Güney Kore'nin SKAIA modelinde olduğu gibi daha esnek, önleyici ve yönlendirici bir yaklaşım benimsenmelidir. SKAIA'daki "yüksek etkili (high impact)" sistem tanımı, YZ Tüzüğü'ndeki risk temelli sınıflandırmadan kavramsal ve yapısal olarak ayrılmakta; teknik mevzuattan çok, sektörel rehberlik ve idari yönlendirme mekanizmalarını esas almaktadır. Bu bağlamda, yüksek etkili sistemlerin tespiti ve izlenmesinde teknik sınıflandırmalar yerine yönetişim temelli ilkelerin esas alınması; sektöre özgü rehberlerin geliştirilerek uyumlaştırıcı süreçlerin desteklenmesi önerilmektedir. Böylece hem düzenleyici esneklik korunmakta hem de yüksek etki potansiyeli taşıyan sistemlerin hesap verebilirliği sağlanabilmektedir.

3.4. Telif Hakkı Rejiminin Güncellenmesi

3.4.1. Yapay Zekâ Tarafından Üretilen İçeriklerin Statüsü

Fikir ve Sanat Eserleri Kanunu (FSEK) madde 1/B'de yer alan eser tanımı, özü itibariyle insan yaratıcılığını temel almaktadır. Bu tanım, bir eserin "fikir ve sanat eseri" sayılabilmesi için sahibinin hususiyetini taşıması ve ilim ve edebiyat, musiki, güzel sanatlar veya sinema eserleri gibi belirli kategorilerden birine girmesi gerektiğini belirtir. Ancak, YZ sistemlerinin gelişim hızı ve bu sistemler tarafından oluşturulan görsel, metin, ses, video ve benzeri içeriklerin nicelik ve nitelik olarak artmasıyla birlikte, mevcut yasal çerçeve yetersiz kalmaya

başlamıştır. Hem ABD hem AB'de insan katkısı telif koruması için olmazsa olmaz bir unsur olarak görülmektedir. (U.S. Copyright Office, 2023), (U.S. Court of Appeals for the D.C. Circuit, 2024), (Court of Justice of the European Union, 2009) YZ'nin kendi başına eser sahibi olması şu anki mevzuat ve içtihatlara göre mümkün değildir. YZ destekli üretimlerin giderek karmaşıklaşması ve yaratıcılık unsurları barındırması, bu içeriklerin FSEK kapsamında eser sayılıp sayılmayacağı konusunda ciddi hukuki belirsizlikler yaratmıştır.

YZ ve telif hakkı sahipliği ile ilgili olarak özellikle ele alınması gereken başlıklar şunlardır:

- ***Eser Niteliği Kazanma Kriterleri:*** YZ destekli içeriklerin "eser" olarak kabul edilip edilmeyeceği, eğer edilecekse hangi koşullarda bu niteliği kazanacağı netleştirilmelidir. Bu noktada, insan müdahalesinin derecesi, YZ'nin otonomisi ve çıktıdaki "özgünlük" veya "hususiyet" kriterlerinin YZ bağlamında nasıl yorumlanacağı belirleyici olacaktır.
- ***Hak Sahipliği:*** YZ tarafından üretilen bir eserin hak sahipliğinin kime ait olacağı, mevcut hukukun en karmaşık sorularından biridir. YZ'nin kendisi bir hukuk süjesi olmadığı için eser sahibi olamaz. Bu durumda hak sahipliği;
    - YZ sistemini geliştiren kişiye/kuruluşa mı,
    - YZ sistemini kullanan kişiye/kuruluşa mı,
    - YZ sistemini eğiten veri setini sağlayan kişilere/kuruluşlara mı,
    - Yoksa bu aktörlerin bir kombinasyonuna mı ait olacaktır? Bu soruya verilecek yanıt, YZ tabanlı yaratıcı endüstrilerin gelişimini doğrudan etkileyecektir.
- ***Koruma Süresi:*** Eğer YZ destekli içerikler eser korumasına tabi tutulursa, bu korumanın süresi mevcut eserlerden farklı mı olacaktır? FSEK kapsamında insan eseri için "eser sahibinin hayatı boyunca ve ölümünden sonra 70 yıl" gibi belirlenmiş bir süre varken, YZ için bu sürenin başlangıç noktası ne olacaktır? YZ'nin "ömrü" gibi bir kavramdan bahsedilemeyeceği için, eserin kamuya açıklanma tarihi veya YZ sisteminin ilk kullanıma sunulduğu tarih gibi alternatif başlangıç noktaları değerlendirilmelidir. (Benzer şekilde değerlendirme, UK Intellectual Property Office, 2022)
- ***Sorumluluk ve Yükümlülükler:*** YZ tarafından üretilen içeriklerin telif hakkı ihlali, itibar zedeleme veya yanıltıcı bilgi yayma gibi hukuka aykırı sonuçlar doğurması durumunda sorumluluğun kime ait olacağı da netleştirilmelidir. Bu, hak sahipliği tartışmalarıyla doğrudan ilişkili olup, YZ'nin öngörülemeyen çıktıları için bir sorumluluk matrisi oluşturulmalıdır.

- ***Dokümantasyon Standardizasyonu:*** Model kartları, veri seti kartları ve etki değerlendirmeleri gibi YZ sistemlerine ilişkin dokümantasyonların standardizasyonu teşvik edilmeli ve düzenleyici kurumlarca benimsenmelidir.

ABD Telif Hakkı Ofisi'nin YZ ve telif hakkı ilişkisine dair görüşü, telif hakkı korumasının temelinde insan yaratıcılığının vazgeçilmez bir koşul olduğu prensibine dayanmaktadır. Ofis, tamamen YZ tarafından, insan müdahalesi olmadan oluşturulan çıktıların eser niteliği taşımadığı ve dolayısıyla telif hakkı korumasından faydalanamayacağı kanaatindedir. Bir YZ çıktısının telif hakkı koruması elde edebilmesi için, bir insan yazarın esere yeterli düzeyde ifade edici ve yaratıcı katkı sağlaması gerekmektedir. Bu durum, insanın YZ çıktısında algılanabilir bir yaratıcı eser bırakması veya çıktıyı seçerek, düzenleyerek ya da üzerinde yaratıcı değişiklikler yaparak dönüştürmesi gibi halleri kapsar; ancak, yalnızca YZ sistemine istem (prompt) sağlamak, tek başına yeterli yaratıcı katkı olarak değerlendirilmemektedir. Ofis, YZ'nin yaratım sürecinde sadece bir araç olarak kullanılması veya YZ tarafından üretilen materyalin daha büyük, insan tarafından oluşturulmuş bir esere dahil edilmesinin telif hakkına engel teşkil etmediğini açıkça belirtir. (U.S. Copyright Office, 2025)

Birleşik Krallık Fikri Mülkiyet Ofisi (UK IPO), YZ teknolojilerinin hızla yayılmasıyla ortaya çıkan telif hakkı sorunlarına ilişkin dengeli bir politika izlemeyi hedeflemektedir. 17 Aralık 2024'te başlatılan ve 25 Şubat 2025'te tamamlanan son istişare süreci, bu alandaki güncel zorlukları ve çözüm önerilerini belirlemede kilit rol oynamıştır. UK IPO'nun temel amacı, hem Birleşik Krallık'ın yaratıcı endüstrilerinin haklarını korumayı hem de ülkenin YZ sektöründe inovasyonu teşvik etmeyi kapsayan çifte bir hedef gütmektedir. İstişare kapsamında, özellikle metin ve veri madenciliği faaliyetleri için telif hakkı yasasına ticari kullanımları da içerecek yeni bir istisna getirilmesi önerilmiştir. Ancak bu istisna, telif hakkı sahiplerine materyallerinin YZ eğitimi için kullanılmasını engelleme ("opt-out") hakkı tanıyarak, Avrupa Birliği'nin yaklaşımına benzer şekilde bir denge mekanizması sunmaktadır. Bu model, YZ geliştiricilerine yasal risk olmaksızın yüksek kaliteli verilere erişim sağlarken, aynı zamanda hak sahiplerinin eserleri üzerindeki kontrolünü ve potansiyel ücretlendirme haklarını güçlendirmeyi amaçlamaktadır. Bu bağlamda, YZ geliştiricilerinin kullandıkları belirli veri setlerini açıklamaları istenebilir; ancak, büyük veri setlerinin karmaşıklığı nedeniyle etkili şeffaflık araçlarının geliştirilmesi için ek destek ihtiyacı da dile getirilmiştir. İstişare çıktısına göre yaratıcı endüstrilerden, YZ'nin telif hakkı materyallerini ancak açık izinle kullanması gerektiği ve "opt-out" mekanizmasının sonraki kopyalar için yetersiz kalabileceği yönünde eleştiriler

gelmektedir. Bu endişeleri gidermek amacıyla Birleşik Krallık hükümeti, etkili bir "hakları saklı tutma" sistemi kurulmadan istişaredeki tekliflere devam edilmeyeceğini beyan etmiştir. Bu süreç, ulusal ve uluslararası hukuki normları entegre etme hedefiyle devam etmektedir. (UK Intellectual Property Office, 2025).

Avrupa Birliği Fikri Mülkiyet Ofisi'nin (EUIPO) *"Telif Hakkı Perspektifinden Üretken Yapay Zekâ'nın Gelişimi"* başlıklı raporuna göre, Üretken YZ (GenAI) sistemlerinin gelişimi, telif hakkı hukukunda yeni zorluklar yaratmaktadır. Başlıca sorunlar, telif hakkıyla korunan içeriğin GenAI eğitiminde kullanımı, hak sahiplerinin eserleri üzerindeki kontrol kaybı riski, YZ tarafından üretilen içeriklerin telif hakkı statüsünün belirsizliği ve şeffaflık eksikliğidir. Bu problemlere yönelik çözüm önerileri, çok katmanlı bir yaklaşımı gerektirmektedir. EUIPO, telif hakkı sahiplerine eserlerini YZ eğitimi için kullanıma kapatma ("opt-out") yeteneğinin etkin şekilde sağlanmasını temel bir gereklilik olarak vurgulamaktadır. Ayrıca, YZ tarafından üretilen çıktılarda şeffaflığın artırılması, sentetik içeriğin tespiti için teknik çözümlerin geliştirilmesi ve kamuoyu farkındalığının artırılması önerilmektedir. Hak sahipleri ile GenAI geliştiricileri arasında doğrudan lisanslama anlaşmalarının kolaylaştırılması ve telif ihlali riskini azaltıcı dahili araçların geliştirilmesi de önemlidir. Son olarak, EUIPO gibi kamu otoritelerinin, metin ve veri madenciliği kullanımına ret hakkı bildirimi yönetimi için teknik destek sağlaması ve Telif Hakkı Bilgi Merkezi gibi platformlar aracılığıyla paydaşlar için bilgi ve işbirliği ortamları yaratması gerekmektedir. Bu adımlar, inovasyonu desteklerken telif hakkı korumasını sağlayan dengeli bir ekosistemin inşasında hayati rol oynayacaktır. (European Union Intellectual Property Office, 2025)

Bu kapsamda ülke olarak uluslararası çalışmaları takip ederek rekabet edilebilirliği artıracak aynı zamanda telif haklarını koruyacak bir model üzerinde düzenlemeler yapmak gerekecektir. Dünya Fikri Mülkiyet Örgütü (WIPO), 2023'ten bu yana "AI and IP" başlıklı kamuya açık istişare süreçleri yürütmektedir. YZ'nin telif haklarına etkisi, sınai mülkiyet (patent, tasarım) alanlarındaki etkisi ve yapay yaratımın kim tarafından sahiplenileceği tartışmaları çok paydaşlı şekilde değerlendirilmektedir.

3.4.2. Eğitim Verilerinde Telif İhlalinin Önlenmesi

YZ Tüzüğü madde 53/1-d uyarınca, genel amaçlı YZ sağlayıcıları, eğitim verileri hakkında yeterince detaylı özet bilgi yayımlamakla yükümlüdür. Bu kapsamda Avrupa Yapay Zekâ Ofisi, özet bilgi şablonu ile ilgili çalışmalarını sürdürmektedir. (European Commission, 2025)

Aynı zamanda, Avrupa Komisyonu tarafından yayımlanan genel amaçlı YZ modellerine ilişkin kılavuzda, YZ sağlayıcılarının telif hakkı yükümlülüklerine dair sorumluluklarına açıkça yer verilmiştir. YZ Tüzüğü madde 53/1-cuyarınca, sağlayıcıların Avrupa Birliği telif hakkı mevzuatına uyumu sağlamak amacıyla bir politika geliştirmeleri gerekmektedir. Buna ek olarak, aynı maddenin (d) bendine göre, sağlayıcıların modelin eğitimi sırasında kullanılan içeriklere ilişkin kamuya açık bir özet sunmaları zorunludur. Bu yükümlülük, açık kaynaklı lisansla sunulan modeller bakımından da geçerlidir ve sadece sistemik risk taşımayan modellerde belirli sınırlı muafiyetler uygulanabilir. Telif hakkı uyum politikalarının, özellikle 2019/790 sayılı AB Telif Hakları Direktifi'nin 4(3) maddesi kapsamında öngörülen "haklar saklıdır" bildirimlerini de içerecek şekilde yapılandırılması beklenmektedir. (European Comission 2025)

YZ Tüzüğü gelişmelerinin ülkemizde de özellikle AB pazarına ürün sunan geliştirmeciler için ciddi etki doğuracağı muhakkak olduğundan, Türkiye'de de özellikle AB'de gerçekleşen çalışmalar incelenerek, telif hakları kapsamında sektörel çalışmalar yapılmalı ve telif ihlallerinin önlenmesi için açıklayıcı rehberler yayınlanmalıdır.

3.4.3. Açık Veri Bankası ve Şeffaflık Mekanizmaları

Önceki bölümde açıklandığı üzere, YZ Tüzüğü madde 53(1)(d) doğrultusunda, genel amaçlı YZ sistemlerinin eğitiminde kullanılan verilerin niteliği kamuya açık bir şekilde özetlenmelidir. Türkiye'de de benzer biçimde, kamu ve özel sektör paydaşlarının sunduğu veri setlerinin merkezi ve erişilebilir bir "Ulusal Açık Veri Bankası"nda toplanması önerilmektedir. Bu bankada, veri setinin kaynağı, anonimleştirilme durumu, telif durumu ve kullanım koşulları gibi bilgiler yer almalıdır. Bu şeffaflık, hem hesap verebilirliği artırır hem de etik ve hukuki denetimi kolaylaştırır.

3.4.4. Güvenli YZ ve Açık Veri Sağlayıcılarına Teşvik Mekanizmaları

YZ sistemleri için güvenilir ürünler sunanlar ile nitelikli açık veri sağlayıcılarının desteklenmesi amacıyla çeşitli teşvik mekanizmaları geliştirilmelidir. TÜBİTAK, KOSGEB ve Kalkınma Ajansları gibi kurumlar aracılığıyla veri kalitesini artırmaya yönelik projelere fon sağlanabilir. Ayrıca, bu kuruluşlara ulusal açık veri stratejisine katkıda bulunmaları durumunda vergi avantajları, öncelikli kamu alımı ve AR-GE destekleri gibi teşvikler sağlanması önerilmektedir.

3.4.5. Eğitim, Sertifikasyon ve Farkındalık

Hukuk, mühendislik ve sosyal bilimler fakültelerinde YZ hukuk ve etiği dersleri zorunlu hâle getirilmeli; kamu görevlileri ve sektör profesyonellerine yönelik sertifikalı eğitim programları hazırlanmalıdır. Bu kapsamda, ülkemizdeki YZ geliştiricilerini de etkilemesi beklenen YZ Tüzüğü konusunda eğitim ve bilinçlendirme çalışmaları yürütülmelidir.

3.5. Kamu Yapılanmaları için Öneriler

3.5.1. Ulusal Yapay Zekâ Komitesi Kurulması

YZ politikalarını yönlendirecek, stratejiler belirleyecek, yatırımları teşvik edecek, veri altyapısını geliştirecek ve risk derecelerine göre YZ sistemlerinin hukuki düzenlemesine yönelik kamu, özel sektör ve sivil toplum kuruluşlarından uzmanları barındıran bir "Ulusal YZ Komitesi" kurulması önemlidir. Bu komite, etik kurallara uyumu artırmak ve YZ sistemlerinin toplumsal etkilerini değerlendirmek için tavsiyelerde bulunmalıdır.

3.5.2. Yapay Zekâ Politika ve Güvenlik Araştırma Merkezleri

Sanayi ve Teknoloji Bakanlığı veya Siber Güvenlik Başkanlığı bünyesinde ve ilgili kamu kurumları ile istişare halinde çalışacak şekilde, YZ politikalarının geliştirilmesi, uygulanması ve düzenlenmesi için "Yapay Zekâ Politika Merkezi" ile YZ güvenlik risklerini analiz etmek ve standartları belirlemek için "Yapay Zekâ Güvenlik Araştırma Merkezi" gibi yapılar kurulması önerilmektedir. Bu merkezlerin yetki alanları net çizilmeli ve bürokratik engeller önlenmelidir.

3.5.3. Yapay Zekâ İşletmeleri için Temsilci Atama Yükümlülüğü

Güney Kore'de olduğu gibi, Türkiye'de adresi veya işyeri bulunmayan ve belirli kullanıcı sayısı veya gelir eşiğini aşan yabancı YZ işletmelerinin ülkede bir temsilci ataması zorunlu hale getirilmelidir. Bu temsilci, güvenlik yükümlülükleri, yüksek etkili YZ sistemlerinin doğrulanması ve güvenlik önlemlerinin denetimi gibi görevleri üstlenmelidir. (Görentaş, 2025)

3.5.4. *Deepfake*, İstismar ve Kişilik Hakkı İhlallerinin Engellenmesi

Üretken YZ tarafından oluşturulan içeriklerin (ses, görüntü, yazı, video) YZ tarafından üretildiği açıkça belirtilmelidir. Siyasi reklamlarda YZ kullanımının açıklanması zorunlu kılınmalı ve yanıltıcı veya manipülatif içeriklerin (deepfake) yayılmasını önlemek için sıkı

düzenleyici önlemler alınmalıdır. Sosyal medya şirketlerine deepfake içerikleri tespit etme, etiketleme ve bildirme yükümlülüğü getirilmelidir.

Aynı zamanda YZ tarafından üretilen çocuk cinsel istismarı materyallerinin oluşturulması, bulundurulması ve dağıtılmasının açıkça yasaklanması ve bu tür suçlara karşı ağır cezai yaptırımlar öngörülmesi gerekmektedir.

3.5.5. Kapsamlı Yasal ve Yönetişim Çerçeveleri Oluşturulması

AB YZ Yasası'ndan farklı olarak, Güney Kore'nin SKAIA modelinde kamu otoriteleri arasında görev dağılımı, stratejik planlamaya dayalı gelişim hedefleri ve düzenleyici altyapının inşasını öncelemektedir. Bu nedenle öneri, bağlayıcılığı YZ Tüzüğü gibi doğrudan maddi yükümlülük getiren bir yasa değil; kurumsal kapasiteyi artıran, strateji-temelli ve yönetişimi yapılandırıcı bir "çerçeve yasa"dır. Nitekim SKAIA da, YZ Tüzüğü'nden farklı olarak özel sektöre değil, devlete görev yüklemektedir ve bu yönüyle politika hukuku niteliğindedir.

Türkiye, Güney Kore'deki SKAIA gibi, kurumsal koordinasyon ve kamu otoritesinin görev dağılımına odaklanan kapsamlı bir YZ yasası çıkarmayı düşünebilir. Bu yasa, YZ ekosisteminin sağlıklı gelişimi ve güven temelli bir yapı için ulusal düzeyde stratejik bir yönlendirme sunmalıdır.

3.5.1. Sayısal İçeriğin Kökeni ve Kimlik Doğrulaması

AB YZ Tüzüğü'nün 50. maddesi, YZ sistemleri tarafından üretilen içeriklerin etiketlenmesine ilişkin şeffaflık yükümlülüklerini düzenlemektedir. Madde 50/1 uyarınca, metin, ses, görsel veya video gibi içeriklerin YZ sistemleri tarafından üretildiği ve bu içeriklerin insanlar tarafından gerçekmiş gibi algılanabileceği durumlarda, kullanıcıların bu içeriğin YZ tarafından üretildiği konusunda bilgilendirilmesi zorunludur. Bildirim, açık, anlaşılır ve zamanında olmalıdır. Madde 50/2, bu yükümlülüğün kullanıcıya sunulan içeriğin niteliğine göre yerine getirilmesi gerektiğini belirtmektedir.

Sayısal içeriğin kökenini ve YZ ile üretilip üretilmediğini tespit edebilmek için hem teknik hem de yönetişim temelli çözümler geliştirilmelidir. Bu kapsamda, ABD'de yaygınlaşan dijital filigranlama (watermarking) yöntemleriyle, YZ tarafından üretilen içeriklere görünmez dijital işaretler eklenmesi ve bu içeriklerin sonradan ayırt edilebilir hâle getirilmesi mümkündür. Ayrıca Adobe'nin "Content Credentials" gibi girişimleri, içeriğin kim tarafından, ne zaman ve nasıl üretildiğini belgeleyen meta veri sistemleri sunarak içerik güvenliğine katkı sağlamaktadır. Türkiye'de de özellikle haber, sosyal medya ve kamu iletişimi gibi alanlarda bu

tür sistemlerin teşvik edilmesi; dezenformasyonla mücadele, yapay içeriklerin şeffaf biçimde etiketlenmesi ve kamu ihalelerinde içerik kökeni takibinin standart hâline getirilmesi gibi uygulamalarla desteklenebilir. Bu çözümler, hem kullanıcı güvenini artıracak hem de yapay içeriklerin etik ve sorumlu kullanımına katkı sağlayacaktır.

3.5.2. İnsan Kaynağı ve İşbirliğinin Güçlendirilmesi

Kamu sektöründe YZ teknolojilerini tedarik eden, kullanan ve yöneten personelin YZ yetkinlikleri, potansiyel riskleri ve etik boyutları konusunda eğitilmesi önemlidir. YZ alanında uzmanlaşmış insan kaynağının yetiştirilmesi ve ülkeye kazandırılması için uluslararası işbirlikleri, yetenek çekme teşvikleri ve araştırma merkezleri desteklenmelidir.

3.5.3. Risk Yönetimi ve Etik Standartların Belirlenmesi

1. **Yapay Zekâ Etik İlkeleri ve Uygulama Stratejileri**

YZ'nin insan hayatına zarar vermemesi, güvenilir olması, fiziksel ve zihinsel sağlığı koruması ve topluma fayda sağlaması gibi temel etik prensipler belirlenmeli ve yaygınlaştırılmalıdır. Özel sektör bünyesinde "YZ Etik Kurulları'nın oluşturulması teşvik edilmelidir. YZ'nin insan hayatına zarar vermemesi, güvenilir olması, bireylerin fiziksel ve zihinsel sağlığını koruması ve toplumsal faydayı öncelemesi, yalnızca etik bir ideal değil, aynı zamanda kurumsal yönetişim için stratejik bir gerekliliktir. ISO/IEC 42001:2023 standardı, YZ sistemlerinin yönetimi için ilk defa oluşturulan uluslararası bir yönetim sistemi standardı olup, kuruluşların YZ uygulamalarını risk temelli, güvenli, şeffaf ve insan merkezli biçimde yönetmelerini hedefler.

2. **Yapay Zekâ Güvenlik ve Güvenilirlik Doğrulama ve Sertifikasyon Desteği**

YZ sistemlerinin güvenliğini ve güvenilirliğini sağlamak amacıyla kuruluşların kendi kendine gerçekleştirdiği doğrulama ve sertifikasyon faaliyetleri desteklenmeli, yüksek riskli YZ sistemlerinde sertifikalı ürünlerin öncelikli kullanımı teşvik edilmelidir.

3.5.4. Şeffaflık ve Hesap Verebilirlik Mekanizmalarının Geliştirilmesi

YZ sistemlerindeki önyargıları tespit etmek ve ortadan kaldırmak için bağımsız denetim mekanizmaları oluşturulmalı ve özellikle kritik alanlarda (istihdam, finans, sağlık, eğitim, konut) YZ tabanlı karar alma süreçlerinin adil ve tarafsız işleyişi sağlanmalıdır. Tüketiciler üzerindeki etkileri değerlendirilmeli ve raporlanmalıdır.

Bu bağlamda, algoritmik ayrımcılığın önlenmesine yönelik önerilen yaklaşım, doğrudan yeni bir yasal zorunluluk getirmekten ziyade, öncelikle şeffaflık, kurumsal sorumluluk ve iç denetim mekanizmalarının güçlendirilmesi ekseninde yapılandırılmalıdır. Nitekim ABD'deki düzenleme deneyimlerinde de görüldüğü üzere, şirketlerin YZ sistemlerini kendi iç süreçleri içinde düzenli olarak değerlendirmesi, bu sistemlerin bireyler üzerinde yaratabileceği olumsuz etkileri tespit etmesi ve bu etkileri kamuya açık biçimde raporlaması beklenmektedir. Özellikle işe alım, kredi değerlendirmesi ve eğitim gibi sosyal eşitlik açısından kritik alanlarda bu tür uygulamalar hayati önem taşımaktadır.

Bu doğrultuda önerilen yaklaşım, ilk aşamada bağlayıcı düzenlemelere dayanmayan, ancak zaman içinde kurumsallaşarak yasal nitelik kazanabilecek bir "denetim kültürü"nün geliştirilmesini amaçlamaktadır. Söz konusu sürecin başlangıçta, sektörel rehberler, gönüllü uyum mekanizmaları ve kurumsal etik çerçeveler aracılığıyla yönlendirilmesi, hem YZ sistemlerinin esnek gelişimini destekleyecek hem de kamusal denetimin meşru zeminini güçlendirecektir. Uzun vadede bu uygulamaların, sistematik biçimde yasal düzenlemelere dönüşmesi kaçınılmaz olmakla birlikte, ilk adımda "rehberlik temelli ve hesap verebilirliği teşvik eden bir yönetişim modeli" benimsenmelidir.

3.5.5. Kapsamlı Yasal ve Yönetişim Çerçeveleri Oluşturulması

Türkiye, SKAIA gibi, kurumsal koordinasyon ve kamu otoritesinin görev dağılımına odaklanan kapsamlı bir YZ yasası çıkarmayı düşünebilir. Bu yasa, YZ ekosisteminin sağlıklı gelişimi ve güven temelli bir yapı için ulusal düzeyde stratejik bir yönlendirme sunmalıdır.

YZ sistemlerine yönelik sorumluluk çerçevesi oluşturulurken, Avrupa Birliği'nin YZ Tüzüğü'ndeki katı ve teknik risk sınıflandırmaları yerine, Güney Kore'nin SKAIA modelindeki gibi daha esnek, önleyici ve rehberlik odaklı bir yaklaşım tercih edilmelidir. SKAIA'da "yüksek etkili sistem" tanımı, teknik detaylardan çok sektörel yönetişim ve rehberlerle şekillenmekte; bu da hem düzenleyici esneklik sağlamakta hem de hesap verebilirliği artırmaktadır.

3.5.6. Yapay Zekâ'nın Hukuki Statüsü ve Sorumluluk Rejimi

YZ sistemlerinin makine öğrenimi yoluyla kendi kendine öğrenebilen, karar alabilen ve öngörülemeyen çıktılar üretebilen yapısı, bu sistemlerin davranışlarının geliştiricileri tarafından dahi tam olarak öngörülememesine yol açmaktadır. Bu durum, özellikle sorumluluk hukukunun temel ilkeleri açısından ciddi tartışmaları beraberinde getirmektedir. Türk Borçlar Kanunu'nda sorumluluk genel olarak kusura dayalı ve kusursuz sorumluluk olmak üzere ikiye ayrılmakta;

ancak YZ sistemlerinin insan müdahalesi olmaksızın karar verebilmesi bu rejimi yetersiz kılmaktadır.

Avrupa Birliği, bu sorunu çözmek amacıyla iki temel düzenleme üzerinde çalışmıştır. Başlangıçta Avrupa Komisyonu, YZ kaynaklı zararların tazmininde karşılaşılan ispat zorluklarını gidermek amacıyla özel bir düzenleme olarak YZ *Liability Directive* taslağını sunmuş ve yüksek riskli YZ sistemlerine ilişkin sorumluluk rejiminde mağdur lehine kanıt yükünün ters çevrilmesini ve bilgi edinme kolaylığı sağlanmasını hedeflemiştir Ancak Avrupa Parlamentosu Hukuk Komitesi ve üye devletlerin hukuk birimleri, halihazırda yürürlükte olan *Product Liability Directive (Ürün Sorumluluğu Direktifi)* ve YZ Tüzüğü kapsamındaki önleyici düzenlemelerin zaten sorumluluk çerçevesini yeterince güçlendirdiğini savunmuş, bu nedenle ikili düzenleme riski, hukuki belirsizlik yaratma potansiyeli ve aşırı düzenleme endişesi gibi gerekçelerle "AI Liability Directive" taslağını 11 Şubat 2025'te geri çekmiştir. Bunun yerine, ürün sorumluluğu direktifinin YZ'yi kapsayacak şekilde revize edilmesi ve YZ Tüzüğü'nün teknik düzenlemeleriyle birlikte uygulanması yönünde bir bütüncül yaklaşım benimsenmiştir. Böylece, daha yalın, uygulanabilir ve teknik–hukuki uyumu yüksek bir çerçeve oluşturulması amaçlanmıştır.

Avrupa Komisyonu'nun YZ Sorumluluk Direktifi'ni geri çekme kararı, Türkiye için önemli düzenleyici dersler içermektedir. AB'deki bu gelişme, dijital teknolojilerde aşırı detaylı, teknik ve katı kuralların hem uygulamada zorluklara yol açtığını hem de inovasyonu sekteye uğrattığını göstermektedir. Türkiye, bu noktada Avrupa Birliği'nin YZ Tüzüğü gibi temel güvenlik ve etik kuralları içeren düzenlemelerinden ilham alabilir; ancak YZ Liability Directive gibi uygulaması zor ve yüksek uyum maliyetine neden olan sorumluluk modellerini doğrudan benimsemek yerine, daha esnek ve sektörel rehberlik temelli bir yaklaşım benimsemelidir. Dünyadaki bu gelişmeler dikkate alınarak, YZ'nin öngörülemezliği ve otonomi kazanan sistemlerinin, klasik sorumluluk hukukunu zorladığı dikkate alınarak (örn. Türk Borçlar Kanunu'nun kusur ve kusursuz sorumluluk ilkeleri) hukuki ve cezai sorumluluk konuları yeniden değerlendirilmelidir.

3.5.7. Türkiye İçin Ek Değerlendirmeler ve Sonuç

YZ'nin hukuki ve etik açıdan denetlenebilir hale getirilmesi, YZ teknolojilerinin hızla gelişen doğası göz önüne alındığında kritik bir öneme sahiptir. Bu, yalnızca teknolojik ilerlemeyi sağlamakla kalmayıp, aynı zamanda bireysel hakları korumak, toplumsal güveni inşa etmek ve olası riskleri en aza indirmek için gereklidir

Nitekim Türkiye'de YZ uygulamaları hızla yaygınlaşmakta; ancak bu gelişime karşılık, denetim altyapısı ile kurumsal farkındalık düzeyinin henüz yeterince olgunlaşmadığı bir geçiş süreci yaşanmaktadır. Bu nedenle, düzenleyici çerçevenin oluşturulmasında Avrupa Birliği'nin katı ve teknik ağırlıklı yaklaşımı yerine; ABD ve Güney Kore örneklerinde olduğu gibi, esnek, yönetişim temelli ve aşamalı olarak gelişebilen modellerin benimsenmesi daha uygun bir strateji olacaktır.

Her iki model de, YZ'nin gelişim sürecinde doğrudan yaptırım uygulamak yerine, sistemin yönlendirilmesini ve kamu-özel sektör uyumunun güçlendirilmesini öncelemektedir. Türkiye açısından da benzer şekilde, öncelikli olarak çerçeve yasalar, sektörel rehberler ve kurumsal sorumluluğa dayalı ilkeler aracılığıyla bir denetim ve yönetişim kültürünün oluşturulması; uzun vadede daha kapsayıcı ve sürdürülebilir bir düzenleme ortamının inşasına katkı sağlayacaktır.

3.5.8. Kaynakça

Avrupa Komisyonu. (2023). YZ *Code of Practice*. Erişim tarihi: 11 Temmuz 2025, Erişim adresi: https://digital-strategy.ec.europa.eu/en/policies/ai-code-practice

Court of Justice of the European Union. (2009, 16 Temmuz). *Infopaq International A/S v. Danske Dagblades Forening*, Case C-5/08. Erişim tarihi: 11 Temmuz 2025 Erişim adresi: https://eur-lex.europa.eu/legal-content/EN/TXT/PDF/?uri=CELEX:62008CJ0005

European Commission. (2025, 10 Temmuz). *Drawing-up a General-Purpose YZ Code of Practice*. Shaping Europe's Digital Future. Erişim tarihi: 22 Temmuz 2025 Erişim adresi: https://digital-strategy.ec.europa.eu/en/policies/ai-code-practice

European Commission. (2025). *Guidelines on the scope of obligations for providers of general-purpose YZ models under the YZ Tüzüğü*. Erişim tarihi: 22 Temmuz 2025, Erişim adresi: https://digital-strategy.ec.europa.eu/en/library/guidelines-scope-obligations-providers-general-purpose-ai-models-under-ai-act

European Union Intellectual Property Office. (2025). *Generative YZ from a copyright perspective: Executive brief*. Erişim tarihi: 22 Temmuz 2025 Erişim adresi: https://euipo.europa.eu/tunnel-web/secure/webdav/guest/document_library/observatory/documents/reports/2025

_GenAI_from_copyright_perspective/2025_GenAI_from_copyright_perspective_executive_brief_en.pdf

Görentaş, M. B. (2025). Güney Kore YZ Yasası'nın Analizi ve Türk Hukukuna Bir Model Olarak Değerlendirilmesi. *Çukurova Üniversitesi Hukuk Araştırmaları Dergisi*, 7, 413–441.

Görentaş, M. B. (2025). *ABD YZ düzenlemelerinin YZ yöntemleriyle analizi* (Yayımlanmak üzere)

Güney Kore Ulusal Meclisi. (2023). *YZ Temel Yasası (SKAIA)*. Erişim tarihi: 1 Ağustos 2025, Erişim tarihi: 22 Temmuz 2025 Erişim adresi: https://likms.assembly.go.kr/bill/billDetail.do?billId=PRC_R2V4H1W1T2K5M1O6E4Q9T0V7Q9S0U0

International Association of Privacy Professionals. (2025, 12 Şubat). *European Commission withdraws YZ Liability Directive from consideration*. IAPP. Erişim tarihi: 30 Nisan 2025, Erişim adresi: https://iapp.org/news/a/european-commission-withdraws-ai-liability-directive-from-consideration

International Association of Privacy Professionals. (12 Şubat 2025). *European Commission withdraws YZ Liability Directive from consideration*. IAPP. Retrieved April 30, 2025, from https://iapp.org/news/a/european-commission-withdraws-ai-liability-directive-from-consideration

Kişisel Verileri Koruma Kurumu. (Ocak 2018). *Veri Güvenliği Rehberi*. Kişisel Verileri Koruma Kurumu. Erişim tarihi 31.07.2024, https://www.kvkk.gov.tr/yayinlar/veri_guvenligi_rehberi.pdf

Kişisel Verileri Koruma Kurulu. (2023). *Tüketici Finansman Kredisiyle Alışveriş İmkânı Sunan Şirket Tarafından İlgili Kişilerden e-Devlet Şifrelerinin Talep Edilmesi hakkında 22/03/2023 tarihli ve 2023/426 sayılı Karar Özeti*. Erişim 8 Temmuz 2025, https://www.kvkk.gov.tr/Icerik/7599/2023-426

Reimsbach-Kounatze, C. and A. Molnar (2024), "The impact of data portability on user empowerment, innovation, and competition", OECD Going Digital Toolkit Notes, No. 25, OECD Publishing, Paris, Erişim Tarihi: 16.08.2026, https://doi.org/10.1787/319f420f-en.

UK Intellectual Property Office. (2022). *Artificial intelligence and intellectual property: Copyright and patents – Government response to consultation*. GOV.UK. Erişim tarihi: 22 Temmuz 2025 Erişim adresi: https://www.gov.uk/government/consultations/artificial-intelligence-and-ip-copyright-and-patents/outcome/artificial-intelligence-and-intellectual-property-copyright-and-patents-government-response-to-consultation

UK Intellectual Property Office. (2025). *Consultation on copyright and artificial intelligence*. UK Government. Erişim tarihi: 22 Temmuz 2025 Erişim adresi: https://ipoconsultations.citizenspace.com/ipo/consultation-on-copyright-and-ai/

U.S. Copyright Office. (2023, 21 Şubat). *Letter regarding copyright registration for "Zarya of the Dawn" by Kris Kashtanova* [PDF]. Erişim tarihi: 22 Temmuz 2025 Erişim adresi: https://www.copyright.gov/docs/zarya-of-the-dawn.pdf

U.S. Copyright Office. (2025, 3 Temmuz). *Copyright Office Releases Part 2 of Artificial Intelligence Report*. Erişim tarihi: 22 Temmuz 2025 Erişim adresi: https://www.copyright.gov/newsnet/2025/1060.html

U.S. Court of Appeals for the District of Columbia Circuit. (2024, 13 Mart). *Stephen Thaler v. Shira Perlmutter, Register of Copyrights and Director of the United States Copyright Office, et al.* (No. 23-5233). Erişim tarihi: 22 Temmuz 2025 Erişim adresi: https://media.cadc.uscourts.gov/opinions/docs/2025/03/23-5233.pdf

World Intellectual Property Organization. (2023). *WIPO Conversation on Intellectual Property and Artificial Intelligence*. Geneva. Erişim tarihi: 22 Temmuz 2025 Erişim adresi: https://www.wipo.int/about-ip/en/artificial_intelligence/

Yapay Zekâ Süreçlerinde İnsan Faktörü

4. YAPAY ZEKÂ SÜREÇLERİNDE İNSAN FAKTÖRÜ

Ülkü Bayer[a], Prof. Dr. Serkan Güneş[b], Dr. Serkan Alkan[c]

[a] *TBD Ankara Şubesi, Sistekno Bilişim Ltd., ubayer@koplan.com.tr*
[b] *Gazi Üniversitesi, Endüstriyel Tasarım Bölümü, serkangunes@gazi.edu.tr*
[c] *ODTÜ Tasarım Fabrikası Yardımcı Direktörü, serkanalkan@live.com*

Bu bölümün amacı YZ standartları ve düzenlemelerinin insan faktörleri perspektifinden değerlendirilmesidir. Üç bölümden oluşan çalışma, (1) birinci bölümde YZ ve İnsan-Sistem standartlarını yaşam döngüsü çerçevesinde teknik olarak ele alır. İnsan-Faktörü uzmanlarının YZ süreçlerinin geliştirme, kullanma ve imha süreçlerinde oynayacağı rol ile ilgili değerlendirmelerde bulunur. (2) İkinci bölüm ise YZ sistemlerinde tasarım disiplininin değişen rolünü, temel tasarım kavramlarını kuramsal açıdan tartışmakta, günümüzde yaygın olan uygulamaları örnekleyerek standartlar açısından insan-sistem etkileşimini incelemektedir. Son bölüm (3) YZ standartlarının, düzenlemelerinin ve YZ-İnsan etkileşiminin yakın gelecekte nasıl şekillenebileceği konusunda öngörüler içermektedir.

4.1. İnsan Faktörünün Yapay Zekâ Yaşam Döngüsündeki Rolü ve ISO 9241-210 Standardı

YZ sistemlerine özgü süreçlere ilişkin gereksinimler ve örnekler sunan ISO/IEC 5338 teknik raporu, insan merkezli tasarım süreçleriyle uyumlu ve birbirini tamamlayıcı bir sıralama yapısına sahiptir. Kullanıcı merkezli tasarım standardı ISO 9241-210'un dört aşamalı çerçevesi, ISO/IEC 5338'de tanımlanan yaşam döngüsü süreçleriyle uyumludur.[1]

Bu bölümün temel amacı, söz konusu uyumu temel alarak insan faktörü uzmanlarının YZ yaşam döngüsündeki rollerini tanımlamak, bu roller kapsamında işbirliği yaptıkları diğer uzmanlık alanlarını belirlemek ve ilgili standartların gerekliliklerini kurumsal açıdan değerlendirmektir.

4.1.1. Kullanım Bağlamını Anlamak ve Tanımlamak

İlk aşama olan *bağlam araştırması* sürecinde, sistemin hangi ortamda, kimler tarafından ve hangi amaçlarla kullanılacağı tanımlanır. Bu aşamada standartlar kurallar koymaz; ancak insan faktörü uzmanının nelere odaklanması gerektiği konusunda rehberlik eder. Bağlam araştırması, sistem ortamının, kullanıcı özelliklerinin ve hedeflerinin, ayrıca sistemin etkileşimde bulunacağı diğer sistemlerin ve destekleyici unsurların sistematik biçimde analiz edilmesini kapsar. ISO 9241-210 çerçevesi, bağlam anlaşılmadan tasarım sürecine ilerlenemeyeceğini açıkça vurgular.[1]

ISO/IEC 5338'de tanımlanan ***6.4.1 iş veya görev analizi süreci*** *[*2], ISO 9241-210 standardında yer alan 'Kullanım Bağlamı' (context of use) kavramıyla benzerdir. İş veya görev analizi süreci, sistemin hangi ortamda, kimler tarafından, hangi görevler için ve hangi araçlarla kullanılacağını tanımlayan temel bölümdür. Bu bölüm yalnızca kullanıcıların rolleri, görevleri, kullandıkları araçlar ve fiziksel-sosyal çevreyi değil, aynı zamanda YZ sistemlerine özgü ek boyutları da kapsayacak şekilde genişletilir.[1]

İnsan faktörü uzmanının, bu aşamada sistem çıktılarının değer zinciri üzerindeki stratejik etkilerini ve iş akışlarını incelemesi beklenir; çünkü YZ projelerinde tek bir çıktının farklı zincirleme kararlar doğurabildiği açıktır. Bu analiz, model çıktısının bölümler arası bağımlılığı nasıl artırdığını ve stratejik dönüşümün nasıl yönlendirdiğini ortaya koyar.

Ayrıca kullanıcıların veriyle olan ilişkisi incelenerek verinin kaynağı, erişim biçimleri ve güvenlik-gizlilik kaygıları bağlam dokümanına işlenir. YZ'nin dinamik doğası gereği öğrenme ve uyarlanabilirlik boyutları da dikkate alınarak yalnızca "sistem ne yapacak?" sorusuna değil, aynı zamanda "sistem değiştikçe kullanıcı bunu nasıl algılar, iş süreçleri ve stratejik hedefler nasıl etkilenir?" sorusuna da yanıt aranır. İnsan faktörü uzmanının bu noktada değişen sistem davranışının iş akışlarına etkilerini kullanıcı senaryoları, görüşmeler ve pilot testlerle analiz etmesi, bulguları iş analizi raporuna aktarması beklenir.

Bununla birlikte etik, risk ve regülasyon boyutları bağlama entegre edilir ve bunlar doğrudan ölçülebilir başarı kriterlerine dönüştürülür. Geleneksel yatırım getirisi (ROI) veya verimlilik gibi ölçütlerin ötesine geçilerek doğruluk, güvenilirlik, açıklanabilirlik, kullanıcı güveni, şeffaflık, tarafsızlık, gizlilik ve toplumsal etki gibi metrikler tanımlanmalıdır.

Bu yaklaşım, YZ sistemlerinin kalite modelleri ile güven, dayanıklılık, doğruluk, güvenlik ve gizlilik gibi unsurlar üzerine odaklanan uluslararası çerçevelerle uyumludur. İnsan

faktörü uzmanı, kalite güvence uzmanları, güvenlik ve gizlilik profesyonelleri, etik danışmanları, hukuk uzmanları ve teknik ekiplerle işbirliği yapar; kullanılabilirlik ve kullanıcı güveniyle ilgili kalite kriterleri, güvenlik açıkları ve gizlilik riskleri, etik değerler ve tarafsızlık ilkeleri, yasal uyumluluk gereklilikleri ile teknik performans sınırlamaları konusunda geri bildirim alır.[19]

Böylece iş analizi süreci, ISO 9241-210'un kullanım bağlamı yaklaşımıyla bütünleşerek yalnızca kullanıcı görevlerini değil, aynı zamanda stratejik dönüşüm, iş akışları, veri ilişkileri, sistemin öğrenebilirliği ve insan odaklı başarı metriklerini kapsayan genişletilmiş bir yapıya kavuşur. Ayrıca bu geri bildirimler, ilgili uluslararası standartlara başvurularak sistematik biçimde doğrulandığında, bağlam analizi ve iş analizi süreçlerinin iyileştirilmesine, risklerin daha erken görünür hâle getirilmesine ve başarı metriklerinin daha sağlam bir temele oturtulmasına katkı sağlar.[1][2][3]

4.1.2. Kullanıcı Gereksinimlerini Tanımlamak

YZ sistemlerinin yaşam döngüsünde, paydaş ihtiyaçlarının yakalanması ve bu ihtiyaçların sistem gereksinimlerine dönüştürülmesi kritik bir aşamadır. ISO/IEC 5338'de 6.4.2 ve 6.4.3 süreçleri bu amaca hizmet eder ve ISO 9241 serisi ile bağlantılıdır. İnsan faktörü uzmanı, bu süreçlerde paydaşların farklı bakış açılarını dengeleyerek, kullanıcı deneyimi ve insan merkezli ilkelerin teknik gereksinimlere yansıtılmasını sağlar.[1][2]

ISO/IEC 5338'deki ***6.4.2 paydaş ihtiyaçları ve gereksinimleri tanımlama sürecinde*** *[2*] insan faktörü uzmanı, kullanıcıların yanı sıra geliştiriciler, düzenleyiciler, veri sahipleri ve iş ortakları gibi tüm paydaşların gereksinimlerini yakalayarak görünür hâle getirir. Bu aşamada yalnızca açıkça ifade edilen ihtiyaçlar değil, adalet, şeffaflık, erişilebilirlik, kullanıcı güveni, etik hassasiyetler ve gizlilik kaygıları gibi beklentiler de analiz edilir, yalnızca ilkeler düzeyinde kalmamasını, sistemin kullanıcı arayüzü, dokümantasyonu ve kullanım senaryolarına yön verecek biçimde gereksinim setine yansıtılması sağlanır. Uzman, farklı paydaşlar arasındaki çatışmaları uzlaştırarak gereksinimler üzerinde açık bir mutabakat sağlanmasına katkıda bulunur. Bu gereksinimler, kalite güvence süreçlerinde doğrulanabilecek şekilde izlenebilir kriterlere dönüştürülür.[1][2][3]

ISO/IEC 5338'deki ***6.4.3 sistem gereksinimlerinin tanımlanması sürecinde*** [2] ise insan faktörü uzmanı, kullanıcılardan ve paydaşlardan elde edilen ihtiyaçları teknik gereksinimlere dönüştürürken geliştirilen sistemin özelliklerine göre ilgili standart ailesinden yararlanır.[1]

Kullanıcı gruplarının ihtiyaçlarını dikkate alır, etkileşim ilkelerini sistem tasarımına taşır. Bu süreçte hem genel ergonomi ilkelerini ortaya koyan standartlar, hem de bağlama özgü farklı gereksinimleri dikkate almak amacıyla sağlık, savunma, ulaşım, enerji gibi sektörlere özgü ergonomi standartları kullanılır. Örneğin sağlık alanında *ISO 13485* Tıbbi Cihazlar standardı insan faktörleri uzmanına yol gösterici olacaktır. Fiziksel-Siber Sistemler, hem *fiziksel dünyadan veri toplayan sensörler ve cihazları hem de siber tarafta bu veriyi işleyen, karar alan ve sonuçları tekrar fiziksel ortama aktaran yazılım-bilişim katmanlarını* kapsar. Bu yapının karmaşık ve çok katmanlı olması, her sektörün kendi risklerini ve insan–sistem etkileşim özelliklerini dikkate alan özel standartları gerekli kılar.[2][3][14]

YZ yalnızca arka planda çalışan bir teknoloji değil hem iş hayatında hem de günlük yaşamda insanın doğrudan kullandığı, işini yaparken ya da günlük ihtiyaçlarını karşılarken etkileşim kurduğu bir araç haline gelmiştir. İş aracı olarak YZ sistemleri, bir makine, yazılım ya da üretim hattında kullanılan teknik bir ekipman gibi, çalışanın görevini yerine getirmesine yardımcı olan bir üründür. Örneğin, bir doktorun teşhis destek sistemi kullanması, bir mühendisin CAD yazılımındaki YZ modülünden yararlanması veya bir işçinin üretim bandındaki robotik kolu YZ algoritmalarıyla yönetmesi bu kapsama girer.[15]

Kullanım nesnesi olarak ise YZ sistemleri, iş dışında ya da günlük yaşamda doğrudan kullanıcıyla etkileşim kurulan bir ürün haline gelmiştir; örneğin bir müşterinin çağrı merkezinde sohbet botu ile konuşması, bir öğrencinin YZ destekli eğitim uygulamasını kullanması veya bir vatandaşın belediye hizmetlerinde YZ tabanlı bir portaldan yararlanması gibi. YZ sistemlerinin iş aracı veya kullanım nesnesi olarak tasarımında, ergonomik ilkeler yalnızca kullanılabilirliği değil, aynı zamanda güvenliği, verimliliği ve insan refahını doğrudan etkileyen temel kalite özellikleridir. *Ergonomi standartları, kullanıcıların fiziksel, bilişsel ve örgütsel gereksinimlerinin tasarım sürecine aktarılmasını sağlayarak, insan-merkezli tasarım yaklaşımını sistematik hale getirir.* Geçmişte ergonomi kavramları çoğunlukla statik teknik sistemlere (ör. sabit makineler, arayüz tasarımı) odaklanırken, YZ dinamik yapısı (nedensellik ilişkileri, öngörülemeyen etkileşimler, geri alınamaz kararlar) nedeniyle bu çerçeveyi genişletmeyi zorunlu kılmakta, kullanıcı–teknoloji–organizasyon etkileşimini de kapsayacak şekilde ele alınmalıdır. *ISO 6385:2016*, ergonominin temel ilkelerini tanımlar ve iş sistemlerinin insanın yetenekleri ile sınırlılıklarına uyarlanmasını hedefler. *ISO 26800:2011* ise ergonomik prensiplerin genel çerçevesini ortaya koyar ve bu prensiplerin farklı alanlardaki uygulamaları için temel bir referans sağlar. *ISO 10075 serisi*, zihinsel iş yükü ve psikososyal faktörleri ele alırken, *ISO 11064:2011* kontrol merkezlerinin ergonomik tasarımına odaklanır.

YZ sistem tasarımı ve kullanımına ilişkin kontrol edilebilirliğin gerçekleştirilmesi ve artırılması hususunda rehberlik sağlamak için *ISO/IEC 8200 (2024***)** standardından destek alınabilir.[15][17][14]

Temel olarak ISO 9241-11:2018 etkileşimli sistemlerde kullanılabilirlik kavramını anlamak ve uygulamak için bir çerçeve sunar; kullanılabilirliği etkinlik (doğruluk ve tamlık), verimlilik (hedefe ulaşırken harcanan kaynaklar) ve memnuniyet boyutlarıyla tanımlar. Diğer standartlar (110, 112, 171), tasarım ilkeleri veya erişim rehberliği sağlarken*, ISO 9241-11 bu ilkeleri "ölçülebilir kalite hedeflerine" çevirir.* Kullanılabilirlik ile ilgili tüm standartlar, doğrudan ya da dolaylı olarak *ISO 9241-11*'de yer alan tanımlara dayanır. [4][22]

Diğer yandan, ISO 9241-110:2020 insan–sistem etkileşiminin ergonomisine odaklanır ve tüm etkileşimli sistem türleri için geçerli olan yedi diyalog ilkesini tanımlayarak kullanıcı merkezli, erişilebilir ve sezgisel arayüzler geliştirilmesine katkı sağlar. Bu etkileşim ilkelerinin uygulanması için bir çerçeve ve etkileşimli sistemler için genel tasarım tavsiyeleri sunar. *ISO 9241-171*, erişilebilirlik rehberliği ile engelli, yaşlı, sınırlı motor becerisine sahip ya da görme/işitme zorluğu yaşayan bireyler dâhil olmak üzere farklı kullanıcı gruplarının ihtiyaçlarının sistemin tasarım aşamasında dikkate alınmasını sağlar ve böylece herkesin erişebileceği bir sistem hedeflenir. *ISO 9241-112* kullanıcı arayüzleri tarafından yazılım kontrollü bilgi sunumuna ilişkin etkileşimli sistemler için ergonomik tasarım ilkelerini ortaya koyar. Bu ilkeler, bilgi ve iletişim teknolojilerinde tipik olarak kullanılan üç temel modaliteye (görsel, işitsel, dokunsal) uygulanır. Bu ilkeler, etkileşimli sistemlerin analizi, tasarımı ve değerlendirilmesinde kullanılabilir.[5][6][7][22]

Bu aşamada insan faktörü uzmanı, kullanıcıların hızlı erişim, adil kararlar veya güvenli kullanım gibi beklentilerini ölçülebilir gereksinimlere dönüştürür. Örneğin "işlem en fazla üç adımda tamamlanmalı" ya da "farklı kullanıcı gruplarında model doğruluk oranı ±%5'i aşmamalı" gibi somut ölçütler belirler. Uzman ayrıca YZ'nin özerklik düzeyini insan-merkezli biçimde tanımlar, hangi karar noktalarında insan kontrolünün zorunlu olduğunu netleştirir ve beklenmeyen davranışlara karşı kullanıcıya müdahale imkânı tanıyacak önleyici gereksinimler ekler. Bu yaklaşım sayesinde etik, yasal ve toplumsal beklentiler, yalnızca soyut ilkeler düzeyinde değil, sistem mimarisi içinde güvence altına alınmış olur. Aşağıdaki tablo 1 ISO 9241 ergonomi standardı ailesinin kapsamını özetleyen bir sınıflandırmadır.

Tablo 1: ISO 9241 Yazılım–Donanım–Çevre–Duyusal Etkileşim–Tasarım Süreci boyutlarının tamamını kapsayan ergonomi standardı ailesi (Not: Tablodaki "xx" ibaresi, ISO 9241 standardı ailesinde aynı konuyu kapsayan alt standartların seri numarası aralığını ifade eder.) [8]

Tür	Seri	Kapsam
Yazılım / Etkileşim	1xx + eski 13–16	Menü, komut, bilgi sunumu, etkileşim ilkeleri
Donanım	3xx / 4xx	Ekranlar, klavye, fare, donanım ergonomisi
Çevresel Faktörler	15xx	Işık, gürültü, titreşim, termal konfor
Dokunsal / Multimodal	9xx	Haptik geri bildirim, VR/AR ergonomisi
Süreç / Çerçeve	2xx	İnsan merkezli tasarım süreçleri, çerçeveler

İnsan faktörleri uzmanı, bu süreçlerde tüm sistem paydaşlarıyla işbirliği yaparak kullanıcı ihtiyaçlarının doğru biçimde yakalanmasını ve sistem gereksinimlerine dönüştürülmesine destek olur. Özellikle kalite güvence ve risk yönetimi uzmanlarının tanımladığı gereksinimlerin yalnızca teorik düzeyde kalmamasını, ölçülebilir ve doğrulanabilir kriterlere dönüştürülmesi konusunda diğer uzmanlar ile iş bilirliği yapar. Bu noktada ilgili standartların varlığı, sürecin verimini artırır; çünkü ortak bir referans çerçevesi sağlayarak paydaşlar arasındaki mutabakatı kolaylaştırır ve gereksinimlerin tutarlı biçimde uygulanmasını destekler.

4.1.3. Tasarım Çözümleri Üretmek

Genel olarak tasarım sürecinin amacı, kullanıcı gereksinimlerini karşılayan, kullanıcı deneyimini gözeten ve uygulanabilir çözümler geliştirmektir. Bu kapsamda kullanıcı görevleri, etkileşimler ve arayüzler tasarlanır; çözümler senaryolar, prototipler veya simülasyonlarla somutlaştırılır; kullanıcı geri bildirimleri doğrultusunda yenilenir ve son olarak uygulamadan sorumlu ekiplere net bir biçimde aktarılır.[1]

ISO/IEC 5338'deki ***6.4.5 tasarım tanımlama sürecinde*** [2], insan faktörleri uzmanları sistem mimarisinden gelen gereksinimleri ayrıntılı, uygulanabilir tasarım çözümlerine dönüştürür. Bu süreçte teknolojiler belirlenir, tasarım özellikleri netleştirilir, alternatifler değerlendirilir ve sistem öğeleri için çizimler, prototipler ve belgeler gibi somut çıktılar üretilir. Ayrıca tasarımın yapı ile uyumu sağlanır, paydaşların beklentileri gözetilir ve izlenebilirlik garanti edilir. [1]

Sistem mimarisi aşamasından sonra "kullanılabilirlik, erişilebilirlik, güvenlik" gibi özellikler, tasarım sürecinde arayüz gereksinimlerine, kullanıcı görev akışlarına ve etkileşim

prototiplerine dönüştürülür; bu noktada özellikle *ISO 9241-110* (Etkileşim İlkeleri) ve *ISO 9241-112* (Bilgi Sunum Ergonomisi) kritik rol oynar, çünkü tasarımın kullanıcıya bilgi aktarma biçimi, görev destek düzeyi ve insan-makine etkileşim kalitesi burada belirlenir. İnsan faktörü veya kullanıcı deneyimi uzmanı ise bu süreçte kullanıcı ihtiyaçlarını görev analizleri, senaryolar ve prototiplerle somutlaştırır; ergonomi standartları uygulayarak arayüz tasarımı, bilgi sunumu, iş yükü dengesi ve duyusal modaliteleri gözetir; kullanıcı güvenliği ve risk azaltma açısından hata risklerini ve aşırı bilişsel yükü minimize edecek gereksinimleri ekler.[5][7][22]

Tasarım alternatiflerini gerçek kullanıcılarla test ederek geri bildirimleri sürece dahil eder; kullanılabilirlik gereksinimlerinin izlenebilirliğini sağlayıp her ergonomi kriterini ilgili tasarım unsurlarıyla eşleştirir; son olarak etik, güvenlik ve toplumsal beklentileri entegre ederek insan kontrolünün zorunlu olduğu noktaları tanımlar.

İnsan faktörü uzmanı, tasarım sürecinde birçok paydaşla yakın işbirliği yapar. *Sistem mühendisleri ve uzmanlarla* çalışarak kullanıcı gereksinimlerini sistem özelliklerle uyumlu hale getirir*, yazılım ve donanım geliştiricilerle* ergonomi ilkelerinin arayüz ve donanım bileşenlerine doğru aktarılmasını sağlar. *Kalite güvence uzmanlarıyla* kullanılabilirlik kriterlerinin ölçülebilir ve izlenebilir olmasını denetler, risk yönetimi uzmanlarıyla insan hatası, bilişsel yük ve güvenlik risklerini azaltmaya yönelik önleyici gereksinimler belirler. Ayrıca *etik ve uyum, hukuk uzmanlarıyla* birlikte çalışarak yasal düzenlemelerin, toplumsal beklentilerin ve etik standartların tasarıma doğru şekilde entegre edilmesini güvence altına alır.

İnsan faktörü uzmanının diğer bir sorumluluğu da ISO/IEC 5338'deki ***6.4.7 bilgi edinme süreci*** [2] bağlamında, teknik ekibe sadece "veri toplama" açısından değil, bilginin *kullanıcı deneyimi, ergonomi, etik ve güvenlik boyutlarıyla nasıl temsil edilmesi gerektiğini* de göstermektir. İnsan faktörü uzmanı bu süreçte hangi bilginin gerekli olduğu, nasıl toplanacağı ve nasıl temsil edilmesi gerektiğini kullanıcı ve insan bağlamında yönlendiren konumundadır.[23]

İnsan faktörü uzmanı, bilgi edinme sürecinde bilginin kapsamını kullanıcı bağlamıyla ilişkilendirebilir; bu sayede hangi verilerin görevler, bilişsel süreçler ve etkileşimler açısından kritik olabileceğini tanımlayabilir. Sürece yalnızca kullanıma yönelik teknik literatür ve sistemden elde edilen veriler değil, aynı zamanda kullanıcı gözlemleri, saha/vaka çalışmaları, etnografik veriler ve mülakatların da dahil edilebileceği öngörülür. İnsan faktörü uzmanı, bilgi edinme sürecinde *ISO 16982'*yi uygulayabilir. ISO 16982, kullanıcı araştırmalarında (gözlem, anket, mülakat, görev analizi) hangi yöntemlerin seçileceğini ve nasıl uygulanacağını

tanımlayarak tasarım sürecinde kullanılabilirlik testleri, senaryo çalışmaları ve prototip değerlendirmeleri gibi yöntemlerin seçiminde rehberlik eder.[9]

İnsan faktörü uzmanı, bilgi edinme sürecinde "Bu bilgi modeli eğitmek için doğru mu?", "Bilginin sunuluş biçimi hataya sebep olur mu?", "Etiketler insan algısıyla uyumlu mu?" gibi sorular sorarak sürecin insan davranışı, karar seçenekler ve duygularıyla uyumlu olması için yönlendirici olur; örneğin kullanıcı davranışlarını etiketlerken "başarısızlık/başarı" yerine "görev tamamlandı/yarım kaldı" gibi kullanıcı açısından daha anlamlı ve tutarlı etiketler önerebilir. Böylece insan kaynaklı hata riskini azaltır ve bilginin yalnızca algoritmayla uyumlu değil, aynı zamanda kullanıcı bakış açısını doğru biçimde temsil eden bir formatta sisteme aktarılmasını sağlar. Son olarak, toplanan bilginin hangi kullanıcı ihtiyaçlarına dayandığını belgeleyerek izlenebilirliği güvence altına alır ve gelecekteki projelerde bilgi tekrar kullanımına katkıda bulunur.

İnsan faktörü uzmanı bu süreçte doğrudan bir "girdi sağlayıcı" olarak konumlanır; bilgi mühendisliği, kalite ve risk yönetimi uzmanlarıyla işbirliği içinde sağladığı bilgi doğru formatta, algoritmalara aktarılabilir.

4.1.4. Tasarımı Değerlendirmek

ISO 9241-210 kullanıcı merkezli değerlendirmenin insan merkezli tasarımın zorunlu bir parçası olduğunu ve en erken aşamalardan itibaren tasarım kavramlarının kullanıcı ihtiyaçlarını anlamak için değerlendirilmesi gerektiğini vurgular.

Bu süreçte yeni kullanıcı ihtiyaçlarının belirlenmesi, tasarımın güçlü ve zayıf yönleri hakkında geri bildirim alınması, kullanıcı gereksinimlerinin karşılanıp karşılanmadığının ölçülmesi ve tasarımlar arasında kıyaslama yapılması hedeflenir. Kullanıcı tabanlı testler her aşamada yapılabilir; ilk aşamalarda eskiz, senaryo A/B testleri veya prototiplerle geri bildirim alınırken, ilerleyen aşamalarda kullanılabilirlik hedefleri ve performans ölçütleri test edilir; beta testleri ve sahada kullanım testleri, uzaktan ölçüm teknikleri de bu kapsama girer. İncelemeye dayalı değerlendirme, kullanıcı testlerinden önce maliyet-etkin bir yöntem olarak uzmanların deneyim ve ergonomi rehberlerine dayanarak sorunları belirlemesini sağlar; kontrol listeleri, standartlar ve kılavuzlarla desteklenebilir, ancak her zaman kullanıcı testindeki sorunları ortaya çıkaramayabilir. Bu bölüm, tasarımların yalnızca ilk geliştirme aşamasında değil, tüm yaşam döngüsü boyunca kullanıcı merkezli yöntemlerle sürekli değerlendirilmesi gerektiğini ve bunun kullanıcı testleri, uzman incelemeleri ve uzun vadeli izlemeden oluştuğunu belirtir.[1]

6.4.13 Geçerleme ve 6.4.14 sürekli geçerleme süreçlerinde [2] insan faktörü uzmanı sistemin yalnızca teknik olarak değil, kullanıcı ihtiyaçları, güven, etik ve kullanım bağlamı açısından da geçerli olup olmadığını değerlendirir. Ayrıca veri/model güncellemelerinin kullanıcıya etkisini takip eder, kullanım kalitesi "quality in use" ölçütlerini tanımlar ve uzun vadeli izleme mekanizmalarını kurar. Yani teknik ekibin model performansına baktığı yerde, insan faktörü uzmanı kullanım performansına odaklanır. İnsan faktörleri uzmanı için nicel kullanım verileri dışında nitel olarak elde edilen verilerde tasarımın dengeli değerlendirilmesi açısından değerlidir.[1]

Sistemin operasyonel bağlamının kullanıcıyla uyumunu doğrulayarak kullanım senaryolarında gerçekten ihtiyaçları karşılayıp karşılamadığını, görev hedefleriyle çıktılarının örtüşüp örtüşmediğini ve başarı ölçütlerinin (hata oranı, iş yükü, güven, tatmin) kullanıcı açısından anlamlı olup olmadığını değerlendirir; veri veya kavram kayması durumunda bunun kullanıcıya etkisini izler, modelde yapılacak güncellemelerde zamanlamanın etkileşim kalitesini bozup bozmadığını kontrol eder.

İnsan faktörü uzmanları aynı zamanda, güvenlik protokollerinin kullanıcı davranışlarıyla uyumunu gözetir; izleme ölçütlerini insan merkezli biçimde tasarlayarak teknik doğruluğun yanı sıra bilişsel yük, etkililik, verimlilik, memnuniyet, güven, sağlık ve güvenlik etkilerini sürekli izler; uzun vadeli kullanıcı deneyimi ve güvenin korunması için geri bildirim mekanizmaları kurar, kullanıcı güveninin zamanla azalıp azalmadığını değerlendirir; son olarak kalite yönetim süreçlerine insan boyutunu da dahil ederek yalnızca teknik ölçütleri değil, kullanım kalitesi temelli göstergeleri de dikkate alır ve kullanım miktarındaki değişimlerin nedenleri insan faktörleri açısından analiz eder. [1][2]

İnsan faktörü uzmanı farklı uzmanlıklardan (bilgi mühendisliği, kalite–risk, etik–güvenlik, hukuk, sistem mühendisliği) gelen girdileri kullanıcı ihtiyaçları, görev başarısı, güven ve memnuniyet açısından değerlendirir; böylece sistemin yalnızca teknik değil, aynı zamanda insan merkezli geçerliliğini sağlar. Bu süreçte en çok *ISO 9241-210* (insan merkezli tasarım süreçleri), *ISO 9241-110* (etkileşim ilkeleri), *ISO 9241-11* (Kullanılabilirlik Tanımı ve Ölçütleri) ve *ISO/TR 16982* (Kullanılabilirlik Yöntemleri) standartlarını uygular; ayrıca kalite ve ölçüm boyutu için *ISO/IEC 25059 serisi* (YZ Sistemleri için Kalite Modeli), risk/etik uyum için *ISO/IEC 23894* (YZ Risk Yönetimi) ve regülasyonlara uyum için *ISO/IEC 42001:2023* (Yönetim Sistemi Standardı) ve *ISO/IEC 38507* (YZ Yönetişim) gibi standartlardan geri bildirim alır.[1][5][4][18]

ISO 9241-210'da kullanıcı eğitimine ilişkin doğrudan bir madde bulunmazken, *ISO/IEC TR 5338* YZ'ye özgü yaşam döngüsünde kullanıcıların doğru bilgilendirilmesi ve eğitilmesini gerekli bir süreç unsuru olarak tanımlar.[2]

ISO/IEC TR 5338:2022 YZ'ye özgü yaşam döngüsü süreçlerinde kullanıcıların bilgilendirilmesi ve eğitilmesine özel olarak değinmekte; özellikle ***6.4.15 işletim (operasyon) süreci*** [2] kullanıcıların sistemin sınırları, riskleri, güvenlik kısıtları ve doğru kullanım biçimleri konusunda bilgilendirilmesi gerektiğini vurgulamaktadır [2][3]. Bu çerçevede insan faktörü uzmanının rolü yalnızca tasarım aşamasıyla sınırlı olmayıp, kullanıcıların sisteme uyumu, eğitim içeriklerinin hazırlanması ve farkındalığın artırılması gibi görevleri de kapsamaktadır. Eğitim ve bilgilendirme faaliyetleri, yanlış kullanım, güven kaybı veya etik risklerin engellenmesini sağlar.

YZ'ye özgü bir standart olmasa da YZ sistemlerinde kullanıcı eğitiminin kurumsal düzeyde yönetilmesi, planlanması, uygulanması ve değerlendirilmesi için standartlar uygun çerçeveler oluşturabilir. Stanford Üniversitesinin 2024 yılında yayınladığı İnsan Merkezli YZ raporunda belirtildiği gibi toplumun YZ'nin potansiyel etkisi hakkında farkındalığı arttıkça, bu konudaki olumlu beklentiler dışında endişe düzeyi de artmaktadır. Bu nedenle YZ işletilme süreçlerinde, toplumdaki endişe düzeyinin dikkate alınması ve insan merkezli standartlar ile uyumun sağlanması, YZ'nin toplum tarafından benimsenmesine katkı sağlayacak, aynı zamanda endişe düzeyinin benimsenme hızını aşarak artmasının önüne geçecektir. [16]

6.4.17 İmha sürecinde [2] doğrudan olmasa da insan faktörü uzmanı, yalnızca teknik imhayı değil aynı zamanda kullanıcı, kurum ve paydaşlar açısından doğabilecek riskleri gözetebilir; bu kapsamda araştırma ve kullanıcılar ile çalışma deneyimine dayanarak sistemin veya verinin ortadan kaldırılmasının kullanıcı alışkanlıkları, güven beklentileri ve iş süreçleri üzerindeki olası etkilerini değerlendirir. İnsan faktörü uzmanı daha çok kullanıcı deneyimi, iş süreçlerine etkiler ve davranışsal boyut üzerinde yoğunlaşır. Kullanıcı alışkanlıklarını ve iş süreçlerini analiz eder. Bunun için sistem kullanım verilerini inceler, anket ve mülakatlarla geri bildirim toplar, ayrıca olasılık senaryoları üzerinden imhanın olası etkilerini değerlendirir. *ISO/IEC 42001* süreçleri kapsamında, geri bildirimlere göre saklama ve imha gerekliliklerini dikkate alır.[12]

4.1.5. Yapay Zekâ Sistemleri ve Tasarım

YZ sistemleri günümüzde yalnızca teknik doğruluk, algoritmik başarı veya veri işleme kapasitesiyle sınırlı değildir. Bu sistemler, bireylerin bilişsel ve duygusal süreçlerine, toplumsal

karar mekanizmalarına ve kurumsal iş akışlarına nüfuz eden çok boyutlu deneyimsel ekosistemlere dönüşmüştür. Kullanıcı, YZ ile etkileşime girdiğinde yalnızca bilgi edinmez; aynı zamanda bu sistemi bilişsel, duygusal ve toplumsal düzeylerde anlamlandırır, güven duyar ya da güvenini kaybeder. Çoğu zaman YZ sistemlerini yalnızca araç değil, karar süreçlerinde bir "ortak" olarak görür. Bu durum, klasik mühendislik yaklaşımını aşan, insan faktörlerini stratejik düzeyde öne çıkaran yeni bir paradigma gerektirir. YZ yaşam döngüsünde insan faktörleri uzmanları; kullanım bağlamını anlamak, disiplinler arası ve demografik çeşitliliği gözetmek, danışma süreçlerine katılmak, kullanıcı deneyimi tasarlamak ve test etmek, insan merkezli değerlendirme yapmak ve etki analizlerine katkı sunmak gibi çok disiplinli becerilerle tüm yaşam döngüsünde rol oynar.[18]

YZ ile insanın bilişsel, duygusal ve toplumsal bütünlüğünü dikkate alması üç seviyede ele alınmalıdır:

- Büyük (Makro) ölçek – hukuk, regülasyon, standartlar
- Orta (Mezzo) ölçek – yaşam döngüsü boyunca insan faktörleri entegrasyonu, organizasyonel ve stratejik yönetişim boyutu
- Küçük (Mikro) ölçek – ergonomi, arayüz tasarımı, kişiselleştirilmiş kullanıcı deneyimi

Bu bölümün temel amacı, mezzo ve mikro seviyede, YZ sistemlerinin yaşam döngüsünde insan faktörleri uzmanlarının rollerini bütünleşik biçimde ele almak; uluslararası standartlar ışığında kullanılabilecek bir çerçeve sunmaktır.

1. İnsan Faktörleri Disiplini: Tasarımda Güvenlik ve Kullanılabilirlik Boyutu

İnsan faktörleri, tasarım, mühendislik ve beşerî bilimlerin kesişiminde yer alan disiplinler arası bir alandır. Amacı, insan-makine etkileşimlerini güvenli, verimli ve kullanıcı dostu biçimde düzenlemektir. Bu alan özellikle havacılık, sağlık hizmetleri, savunma ve otonom araçlar gibi hata toleransının düşük olduğu yüksek riskli sektörlerde kritik bir rol üstlenir.

Tasarımcıyı "ayrı" bir kategori olarak değil, "insan faktörleri şemsiyesi altında konumlanan bir uzmanlık kolu" olarak görmek hem uluslararası standartlarla uyumlu hem de tasarım disiplininin epistemolojik derinliğini ve bütüncüllüğünü koruyan daha isabetli bir değerlendirme biçimi sunmaktadır. İnsanların sınırlı bilgi işleme kapasitesi dikkate alınmadan tasarlanan sistemler, karar hataları ve güven kaybına yol açabilir. Bu nedenle bilgi akışı

düzenlenmeli, kritik veriler önceliklendirilmeli ve kullanıcıya uyarlanabilir arayüzler sunulmalıdır.

Ergonomi yalnızca fiziksel değil, bilişsel ve duygusal boyutlarıyla da ele alınır. Fiziksel ergonomi, kontrol paneli tasarımı ya da çalışma ortamı düzenlemesi gibi alanları kapsar. Bilişsel ergonomi, karar verme süreçlerinde bilgi sunum biçimlerini optimize eder. Duygusal ergonomi ise stres, güven ve motivasyon gibi psikososyal boyutları dikkate alır. [20]

Karar destek sistemlerinde tasarım temelli insan faktörlerinin önemi giderek artmaktadır. Kullanıcıların teknolojiye güven duyması yalnızca algoritmanın doğruluğu değil, çıktının anlaşılabilirliği ve kontrol hissiyle ilgilidir. Açıklanabilir YZ (XAI) teknikleri, şeffaf algoritmalar ve kullanıcı dostu görselleştirmeler, tasarım temelli insan faktörleri uzmanlarının katkılarını güçlendirmektedir.

2. Tasarım Disiplini: Deneyim ve Anlamın İnşası

YZ sistemleri bağlamında tasarım yalnızca arayüz estetiği, grafik unsurlar ya da kullanım kolaylığı ile sınırlı değildir. Tasarım, çok daha geniş bir zemine sahiptir ve insanın teknolojiyle kurduğu ilişkiyi deneyimsel, bilişsel ve duygusal boyutlarıyla şekillendiren bir süreçtir. Bu nedenle tasarım, insan faktörlerinin sağladığı güvenlik, ergonomi ve işlevsellik eksenini tamamlayan, fakat aynı zamanda ondan ayrışan anlam üretim işlevini üstlenmektedir. [20]

Kullanıcıların YZ sistemlerini deneyimleme biçimi, salt işlevsellikten çok daha fazlasını kapsar. Bir arayüz yalnızca bilgiyi iletmekle kalmaz; aynı zamanda kullanıcıya güven, şeffaflık ve öngörülebilirlik hislerini verir. Örneğin açıklanabilir YZ (XAI) yaklaşımları yalnızca teknik bir şeffaflık değil, aynı zamanda kullanıcıya "sistemin nasıl çalıştığını anlama" deneyimi sunar. Bu, epistemolojik düzeyde tasarımın anlam inşası işlevini ortaya koymaktadır.

Tasarımcı yalnızca bir ekran düzeni ya da estetik tercih belirlemez; aynı zamanda insanın bilişsel yükünü dengeleyen, güven duygusunu pekiştiren, mahremiyet ve özerklik gibi değerleri kullanıcı deneyimine entegre eden bir aktördür. Burada tasarım, teknik doğruluk ve ergonomi ile sınırlı kalmaz; etik ilkelerin ve toplumsal beklentilerin deneyim üzerinden hayata geçirildiği bir araç işlevi görür. [21]

Tasarım, güven kaygısını azaltan, özerklik hissini güçlendiren ve kullanıcıya kontrol duygusu sağlayan bir rol üstlenir. Buradaki kritik nokta, tasarımın işlevsel ve deneyimsel, bilişsel ve duygusal boyutları birlikte ele almasıdır. İnsan faktörleri/ergonomi "nasıl güvenli ve

verimli çalışırım?" sorusuna yanıt ararken, tasarım "bu sistem bana nasıl hissettiriyor ve benimle nasıl anlamlı bir ilişki kuruyor?" sorusuna yanıt arar.

YZ sistemleri, klasik yazılım geliştirme yaşam döngüsünden önemli ölçüde ayrılır. Geleneksel yazılım süreçleri genellikle tasarla – geliştir – test et – uygula biçiminde doğrusal bir mantıkla işlerken, YZ sistemleri çok daha karmaşık ve döngüsel bir yapıya sahiptir. Bunun temel nedeni, YZ'nin yalnızca önceden tanımlı kuralları yerine getiren deterministik bir yapı değil, veriyle birlikte sürekli öğrenen, bağlamdan etkilenen ve öngörülemeyen çıktılar üretebilen sistemler olmasıdır. Dolayısıyla YZ yaşam döngüsü, yalnızca teknik geliştirme süreçlerini değil; aynı zamanda etik riskleri, kullanıcı algısını, toplumsal etkileşimleri ve güvenlik boyutlarını da kapsamak zorundadır. Bu noktada tasarım uzmanlarının rolü kritik hale gelir; çünkü kullanıcı gereksinimleri, bilişsel sınırlar, ergonomik ilkeler ve etik beklentiler yaşam döngüsünün her aşamasına entegre edilmelidir.

3. YZ Yaşam Döngüsü ve Tasarımın Rolü: Ankara X Hastanesi Geriatri Servisi Hipotetik Örneği

Bu başlık altında insan faktörü ve tasarım uzmanlarının YZ yaşam döngüsü aşamalarında alabilecekleri görev ve sorumlulukları hipotetik bir X hastane örneğini ele alarak incelemektedir.

X Hastanesi, Ankara'da bir üniversite hastanesi olup yalnızca şehir içinden değil, Anadolu'nun pek çok farklı bölgesinden yaşlı hastalara hizmet vermektedir. Çoğu hasta yanlarında refakatçileriyle birlikte gelmekte ve serviste yoğun bir klinik akış yaşanmaktadır. Hastalara çok sayıda test uygulanmakta, basit testler kısa sürede tamamlanabilse de kapsamlı tetkikler için kimi zaman altı aya varan bekleme süreleri oluşmaktadır. Hasta sayısının fazlalığına karşın doktor sayısı yetersizdir. Teknik altyapı mevcut yoğunluğu karşılamakta zorlanmakta, bu da hem sağlık çalışanlarının iş yükünü artırmakta hem de hasta ve refakatçilerin süreç boyunca belirsizlik ve stres yaşamasına yol açmaktadır. Bu bağlamda geliştirilecek YZ destekli klinik karar destek sistemi, yalnızca teknik doğruluk üretmekle kalmamalı, aynı zamanda hasta yoğunluğu, kaynak kıtlığı ve kullanıcı deneyimi gibi gerçekçi kısıtları dikkate almalıdır.

Örnek, ISO/IEC 5338 yaşam döngüsü standardında belirtilen, bağlam analizi, gereksinim tanımlama, tasarım ve prototipleme, değerlendirme ve test adımları, eğitim ve adaptasyon, izleme ve geri bildirim, imha ve çıkış sürecini de içermektedir. Bu aşamalar

tasarımcı ve insan faktörü uzmanlarının sürece olan katkılarını daha iyi sınıflamak üzere eklenmiştir. (Tablo 2).

Tablo 2. YZ Yaşam Döngüsünde Tasarımcının İki Boyutlu Katkısı

Yaşam Döngüsü Aşaması	***(İnsan Faktörleri Boyutuyla) Katkısı***	***(Ürün/Hizmet Tasarımı Boyutu) Katkısı***	***Pratik Uygulama: X Hastanesi Geriatri Servisi***
Bağlam Analizi	Kullanıcı görev akışlarının, bilişsel kapasite ve iş yükü sınırlarının analizi; kültürel farklılıklar ve güvenlik hassasiyetlerinin belgelenmesi, "context of use" dokümantasyonu	Kullanıcı yolculuklarının görselleştirilmesi, empati haritalarının çıkarılması; hasta, doktor, hemşire ve refakatçi deneyimlerinin senaryolaştırılması	Yoğun hasta yükü, az doktor sayısı ve düşük sağlık okuryazarlığına sahip refakatçilerin sürece katılımının haritalanması
Gereksinim Tanımlama	Açık gereksinimler (doğruluk, hız, hata azaltma) + örtük gereksinimler (etik uyum, şeffaflık, güven). Ölçülebilir kriterlerin belirlenmesi	Kullanıcı deneyimi hedeflerinin tanımlanması (ör. "bilgiye 3 tıklamada ulaşma" veya "doktorun 2 dakikadan kısa sürede risk analizi yapabilmesi")	İlaç etkileşimlerini doğru tespit etme yanında refakatçiler için anlaşılır rapor çıktısı ihtiyacının gereksinimlere eklenmesi
Tasarım ve Prototipleme	Bilişsel yük yönetimi, hata toleransı, ergonomik ilkeler; şeffaflık ve açıklanabilirlik (XAI) arayüzleri	Görsel tasarım, etkileşim akışları, ikonografi, hizmet dokümantasyonu, kullanıcı dostu prototipler	Doktor ekranında uyarılarla kritik etkileşimlerin gösterilmesi; hemşire ekranında ilaç saatlerinin belirginleştirilmesi, hasta/refakatçi için Türkçe çıktı tasarımı
Değerlendirme ve Test	Kullanılabilirlik testleri, hata oranı analizi, bilişsel yük ölçümleri, etik uyum denetimi	Deneyim prototipleri, kullanıcı senaryoları, görsel ve dilsel sadelik testleri	Doktorların algoritmik önerilere güven düzeyi anketleri; hemşirelerin uyarı tepkilerinin ölçülmesi, refakatçilerin raporu anlayıp anlamadıklarının test edilmesi
Eğitim ve Adaptasyon	Sistem sınırlarının, risklerin ve güvenlik noktalarının kullanıcıya açıklanması; görev temelli sade eğitim	Deneyim odaklı eğitim materyalleri (simülasyon, video, broşür), kullanıcı yolculuğuna uygun adaptasyon stratejileri	Doktorlara kısa modüller, hemşirelere görev başında erişilebilir görsel yönergeler, refakatçilere ilaç risklerini anlatan broşürler
İzleme ve Geri Bildirim	Performans ve etik uyum metriklerinin izlenmesi; kullanıcı güveni ve iş yükü dengesi ölçümü	Kullanıcı deneyimi güncellemeleri; hizmet kalitesi ve memnuniyet ölçümleri, arayüz revizyonları	Doktorların önerileri ne ölçüde benimsediğinin, hemşirelerin hangi uyarıları gereksiz bulduğunun ve refakatçilerin çıktılardan memnuniyetinin raporlanması
İmha ve Çıkış Süreci	Veri güvenliği, etik risklerin değerlendirilmesi, şeffaf kullanıcı bilgilendirmesi	Geçiş deneyiminin tasarımı; kullanıcıya güven verici iletişim ve hizmet çıktıları	Yeni versiyona geçişte eski verilerin güvenle silinmesi; doktorlara ek eğitim; refakatçilere sürecin anlaşılır biçimde aktarılması

4.2. İnsan Merkezli Yapay Zekâ Sistemleri İçin Standartlara Dayalı Sonuçlar ve Öneriler

1. **YZ sistemlerinde sürekli görev işleme; sistemlerin durmaksızın ve gerçek zamanlı biçimde çalışarak kesintisiz görevler üstlenmesini ifade eder.** Bu tür bir işleyiş, otonom araçların trafikte sürekli veri algılaması ya da yoğun bakımda hastaların sürekli izlenmesi gibi kritik örneklerde karşımıza çıkar. Böylesi durumlarda sistemler yalnızca anlık kararlar vermekle kalmaz, aynı zamanda kullanıcıların bilişsel yükünü de doğrudan etkiler. Bunun yanında kontrol müdahalelerinin geri alınamaması, yapılan bir eylemin telafi edilememesi veya geri döndürülememesi anlamına gelir. Robot cerrahların gerçekleştirdiği cerrahi kesiler, otomatik frenleme sistemleri ya da üretim hatlarında yapılan kalıcı işlemler gibi durumlar etkileşim tasarımını daha da kritik hale getirir.

   İnsan–makine etkileşiminin niteliği bu bağlamda büyük önem taşır. Görsel uyarılar, dokunsal (haptic) geri bildirimler ve sesli ikazlar gibi bilgi iletim yöntemleri; kullanıcıya sunulacak seçenekler ve sorumluluk paylaşımının çerçevesi; ayrıca sistemin durumunu, risklerini, alternatiflerini ve sonuçlarını kullanıcıya zamanında ve güvenilir biçimde aktaran bilgi sunumu, etkileşimin temel unsurlarıdır. Ancak mevcut ergonomi standartlarının, özellikle yüksek riskli ve geri dönüşü olmayan kararların verildiği kendi kendine dinamik biçimde işleyen sistemlerde bu etkileşim ve bilgi sunumunun nasıl düzenlenmesi gerektiğini yeterince tanımlamadığı görülmektedir. Bu nedenle insan faktörleri uzmanlarının söz konusu boşlukları dolduracak yöntemler geliştirmesi ve standartların güncellenmesi, kullanıcı güvenliği ve sistemlerin güvenilirliği açısından kritik bir gerekliliktir.[15]

2. **YZ uygulamaları iş süreçlerini dönüştürmekte, görev dağılımlarını değiştirmekte ve çalışanlardan beklenen nitelikleri yeniden şekillendirmektedir.** Rutin görevlerin otomasyonu, çalışanların daha karmaşık karar alma, denetim veya yaratıcı süreçlere yönelmesini gerektirebilir. Örneğin, YZ destekli robotların standart montaj işlerini üstlenmesi durumunda işçilerden; robotları denetleme, sorun giderme ve süreç iyileştirme amacıyla veri yorumlama gibi yeni beceriler beklenmektedir. YZ sistemlerinin başarılı entegrasyonu için teknoloji ile organizasyonel yapının birlikte tasarlanması ve optimize edilmesi gereklidir. Bu bağlamda bir yandan algoritmalar,

veri akışları ve teknik süreçler kurgulanırken; diğer yandan iş tanımları, rol dağılımları, iletişim biçimleri ve çalışan deneyimi dikkate alınmalıdır. Sistem yalnızca teknik olarak değil, insan–teknoloji–organizasyon etkileşimini bütüncül biçimde destekleyecek şekilde işlemelidir.[15]

3. YZ'nin, statik teknik sistemler olmak yerine içerik ve zaman açısından dinamik bir yapıya sahip olması yani sistemin aynı girdiye her zaman aynı cevabı vermemesi; zaman, bağlam ve deneyim arttıkça davranışının değişmesi öngörülemeyen sonuçlar doğurabilmesi nedeniyle **YZ sistemlerinde güncellemeler yalnızca yeni özellik eklemek değil, aynı zamanda eski doğruluğun, güvenliğin ve kullanıcı deneyiminin de korunup korunmadığını test etmeyi gerektirir.** YZ bileşenlerinin güncellenmesi; sistemin diğer BT altyapılarının veri setleriyle bağlantısının güçlendirilmesiyle *kullanıcı arayüzlerinin* de güncellenmesini içermelidir. Bu çerçevede sürekli iyileştirme yaklaşımı, sistem otonomi seviyelerinin uygun şekilde ayarlanmasını ve insan müdahale yeteneklerinin de yeniden tanımlanmasını dikkate almalıdır.[15]

4. **Standartların doğrudan otomasyon kazanımları elde etmesine imkân verilmelidir.** Standartlar yalnızca belge formatında değil, doğrudan makineler tarafından işlenebilir ve uygulama süreçlerine entegre edilebilir biçimde tasarlanmalı, değer tabloları, parça tanımları, 3D modeller, yazılım bileşenleri, gereksinim tanımları ve test yöntemleri gibi standart içerikleri makineler tarafından doğrudan kullanılabilir hâle getirilmelidir. [15]

   Kullanıcı merkezli tasarım standartlarında (ISO 9241 serisi gibi) kişiselleştirilmiş arayüzlerin doğrulanması, kullanılabilirlik testlerinin otomatikleştirilmesi ve senaryo çeşitliliğinin artırılması için GAN (Üretici Çekişmeli Ağ) tabanlı simülasyonlardan yararlanılabilir. Aynı şekilde "SMART Standards" (akıllı standartlar) yaklaşımı kapsamında, GAN tabanlı modeller büyük ölçekli kullanıcı senaryoları üreterek standartların daha esnek, kapsayıcı ve makine tarafından işlenebilir hale getirilmesini destekleyebilir. Bu noktada GAN modelleri, gerçek kullanıcı davranışlarına benzer yapay senaryolar ve büyük ölçekli veri setleri üreterek standartlaştırma süreçlerine önemli katkılar sağlayabilir. Örneğin, bir arayüz standardı etkinliklerinde, GAN modelleri farklı yaş, kültür veya kullanım bağlamına sahip kullanıcıların

etkileşimlerini simüle edebilir. Bu da standartların daha esnek (farklı durumlara uyarlanabilir), kapsayıcı (farklı kullanıcı gruplarını gözeten) ve makine tarafından işlenebilir (XML, RDF gibi formatlarda modellenebilir) olmasına katkı sağlar.[15]

5. **Konumlandırma (Positioning), Yönlendirme (Navigation) ve Zamanlama (Timing) – kısaca KYZ verisi – hareket ve zaman bilgisini sağlayan temel altyapı unsurlarıdır** ve kritik sistemlerin güvenli çalışmasında hayati rol oynar. Örneğin, otonom araçlarda KYZ yalnızca rota takibi için değil, aracın tüm algılama, karar verme ve hareket kontrol mekanizmalarının senkronize biçimde işlemesi için gereklidir. Bu sistemde yaşanabilecek en küçük kesinti veya manipülasyon dahi güvenlik açısından ciddi riskler doğurabilir. Mevcut sektörel standartlar bu tür sistemleri parçalı biçimde düzenlemektedir. Ancak KYZ yalnızca verinin güvenilirliği bağlamında değil, aynı zamanda bu veriyi kullanan tüm sistemlerin güvenliği, dayanıklılığı, hata toleransı ve insan–sistem etkileşimi açısından da ele alınmalıdır. Bu nedenle, KYZ temelli Fiziksel-Siber Sistemler için özel standartların geliştirilmesi, veriyi kullanan teknolojilerin güvenliğini artırmak ve kritik altyapıların bütünsel dayanıklılığını güçlendirmek bakımından büyük önem taşımaktadır.[14]

6. **5338 süreçleri kapsamında işbirliğini mümkün kılacak yapının etkin şekilde işlemesi için, sistem bileşenleri arasında veri transferini düzenleyen açık, tutarlı ve güvenli kurallara ihtiyaç vardır;** böylece süreçler hem kurumsal bütünlük hem de uluslararası standartlarla uyum içinde yürütülebilir. Bu kuralların yalnızca belirli aşamalarda değil, tüm 5338 süreçlerine entegre edilmesi, veri akışının sürekliliğini, şeffaflığını ve güvenilirliğini sağlayarak işbirliği yapısının kurumsal düzeyde istikrarlı biçimde işlemesine katkı sunar.

7. **ISO 9241 serisi kullanılabilirliği "etkililik, verimlilik ve memnuniyet" ölçütleri** üzerinden, kullanım bağlamına bağlı olarak tanımlayarak diğer bütün alanlara ve YZ projelerinde de temel bir çerçeve sunar. Ancak bu standart tek başına YZ bağlamındaki gereksinimleri karşılamada yetersiz kalmaktadır. YZ sistemlerinde açıklanabilirlik, önyargı ve belirsizlik yönetimi, kontrol edilebilirlik, etik ve hukuki sorumluluklar gibi ek boyutlar öne çıkar; bunlar örneğin ISO 9241-11'in kapsamı dışında kalır. Bu nedenle, 9241-11 alanın tanımları için gerekli bir temel sağlamakla

birlikte, YZ projelerinde mutlaka ISO/IEC 25059 (YZ kalite modeli), ISO/IEC 23894 (risk yönetimi), ISO/IEC 42001 (AI yönetim sistemi) gibi tamamlayıcı standartlar ve çerçevelerle birlikte uygulanması önerilmektedir.

8. **İnsan faktörü uzmanı *bağlam analizi ve iş/görev analizi* sürecine katılarak sistemin kullanım ortamını, kullanıcı özelliklerini, veri ilişkilerini ve stratejik etkilerini değerlendirir;** ardından *paydaş ihtiyaçlarının ve gereksinimlerinin tanımlanması* sürecinde tüm paydaşların beklentilerini ölçülebilir kriterlere dönüştürür; *tasarım çözümleri üretmek* aşamasında görev analizleri, prototipler ve senaryolarla çözümleri somutlaştırır; *bilgi edinme sürecinde* ise hangi verilerin gerekli olduğu, nasıl toplanacağı ve nasıl temsil edileceği konularını kullanıcı bağlamıyla ilişkilendirir; sonrasında *tasarımın değerlendirilmesi ve geçerleme* süreçlerinde kullanılabilirlik testleri, geri bildirimler ve kullanım kalitesi ölçütleri üzerinden sistemin geçerliliğini kontrol eder; ayrıca *eğitim ve bilgilendirme* sürecinde kullanıcıların doğru kullanım, riskler ve sınırlar hakkında bilinçlenmesini sağlar; son olarak *imha sürecinde* kullanıcı alışkanlıkları, güven beklentileri ve iş süreçleri üzerindeki etkileri dikkate alır. Bu süreçlerde ISO 9241 insan -sistem etkileşimi standartlar ailesine (ISO 9241-210, 9241-11, 9241-110, 9241-112, 9241-171 gibi), ISO/TR 16982 ve ISO 6385, 26800, 10075, 11064 gibi ergonomi standartlarına uygunluk sağlar; ISO/IEC 25059, 23894, 8200, 42001 ve 38507 standartlarından geri bildirim alır. Sistem mühendisleri, yazılım/donanım geliştiriciler, kalite güvence, risk yönetimi, bilgi mühendisleri, güvenlik-gizlilik uzmanları ile etik danışmanlar ve hukukçularla işbirliği yapar.

9. **Geliştirilen sistemlerin yalnızca geliştirme aşamasında değil, tüm yaşam döngüsü boyunca kullanıcı merkezli yöntemlerle sürekli değerlendirilmesi gerekir.** Kullanıcı merkezli araştırma yöntemleri, tasarım sürecinde insan faktörü uzmanlarının en önemli araçlarını oluşturur. *Görev analizi*, kullanıcıların sistemi hangi adımlarla, hangi zorluklarla ve hangi bilişsel yük altında kullandığını ortaya koyarken; *senaryo temelli değerlendirme,* gerçek kullanım bağlamına benzer senaryolarla sistemin işlevselliğini ve kullanıcı davranışlarını test etmeyi sağlar. *Prototip testleri*, erken aşama arayüzler veya işlevsel prototipler üzerinden kullanıcı geri bildirimlerini toplarken; *kullanılabilirlik testleri*, görev tamamlama süresi, hata oranı ve memnuniyet gibi ölçütler aracılığıyla sistemin etkinliğini ve verimliliğini

değerlendirir. Bunun yanı sıra *anket ve mülakatlar*, kullanıcıların algılarını, beklentilerini ve deneyimlerini nitel ve nicel olarak sunar; *gözlem ve saha çalışmaları*, kullanıcıların doğal ortamda sistemi nasıl kullandığını inceleyerek bağlamsal veriler sağlar.

10. Bu yöntemlerin birlikte uygulanması**, tasarım sürecinin yalnızca işlevsel değil, aynı zamanda kullanılabilir, güvenli ve kullanıcı odaklı olmasını** garanti altına alır. İnsan faktörü uzmanı açısından bu yöntemler, standartların hangi süreçte gerekli olduğuna dair önemli gerekçeler sunar. Dolayısıyla, İnsan faktörü uzmanının, standartların hangi aşamada gerekli olduğuna dair tespitleri; kullanıcı ihtiyaçlarının doğru zamanda kontrol edilmesini, gereksiz tekrarların önlenmesini ve tüm süreçlerin daha hızlı, verimli ve uyumlu yürütülmesini sağlar. Bu yaklaşım hem kaliteyi yükseltir hem de zaman/emek tasarrufu kazandırır.

11. **İnsan faktörleri uzmanları için uluslararası standartların uygulanması hem fikri mülkiyet haklarının hem de uzmanlık bilgisinin (Know-How) korunmasında stratejik bir güvence işlevi görmektedir.** YZ sistemlerinde şeffaflık ve hesap verebilirlik ilkeleri, aynı zamanda gizlilik risklerini gündeme getirmektedir. Bu risklerin etkin biçimde yönetilmesi, standartlara uyumun belgelendiği sertifikasyon mekanizmaları ile mümkün olabilir. Sertifikasyon, yalnızca kurumsal güvenilirliğin bir göstergesi olmakla kalmaz; aynı zamanda gizli bilginin korunmasına, paydaş güveninin artırılmasına ve uluslararası düzeyde rekabet avantajının güçlendirilmesine hizmet eder.

12. **Standartların uygulanması, kurumsal süreçlerde düzen ve güvenilirlik sağlamakla** birlikte, belirli kurallar çerçevesinde hareket etmeyi zorunlu kıldığından **yenilikçi yaklaşımların geliştirilmesini kimi zaman sınırlayabilir**. Katı standart uygulamaları, esnek düşünme alanını daraltarak alternatif yöntemlerin veya özgün çözümlerin ortaya çıkmasını güçleştirebilir. Bu nedenle kurumlar, standart uyumunu sağlarken aynı zamanda yenilikçiliği destekleyen esnek mekanizmalar kurmalı, standartların rehberlik işlevini inovasyonu engellemeyecek şekilde yorumlamalıdır.

13. **Büyük veri analizi ile kullanıcı araştırmalarının hangi durumlarda kullanılacağına ilişkin kararın insan faktörleri uzmanları tarafından verilmesi**, elde edilen bilgilerin bağlama uygun, güvenilir ve kullanıcı odaklı olmasını sağlar.

Sistemin ihtiyaç duyduğu bilgi türünü ve kullanıcıların etkileşim biçimlerini analiz ederek en uygun yöntem seçildiğinde, hatalı sonuçların ortaya çıkma olasılığı önemli ölçüde azalır. Yalnızca büyük veriye dayalı analizler, kullanıcı davranışlarının ardındaki motivasyonları ve bağlamsal nüansları gözden kaçırabilir.

Sonsöz olarak, başlangıçta belirtildiği gibi raporun bu bölümü, insan faktörleri alanının YZ'nin (i) kullanım bağlamını anlamak ve tanımlamak, (ii) kullanıcı gereksinimlerini tanımlamak, (iii) tasarım çözümleri üretmek, (iv) tasarımı değerlendirmek gibi yaşam döngüsünde üstleneceği görevleri metin içinde değinilen standartlar çerçevesinde ele almıştır. İnsan faktörleri alanı ve insan merkezli tasarım yaklaşımı YZ'nin toplum tarafından benimsenmesi, etkin, verimli ve memnuniyet ile kullanılmasını sağlayabilir. Aynı zamanda, insan faktörleri ve insan merkezli tasarımın YZ ile olan ilişkisini inceleyen orijinal ve derleme çalışmaların son dönemlerde yıllık onlu sayılardan, yıllık on binli sayılara ulaşması konunun hem akademi hem de iş dünyası için ele alınması gereken önemli bir başlık olduğunu göstermektedir. Bu artışın ardındaki neden, YZ tarafından sunulması beklenen güven, şeffaflık ayrıca uyulması gereken etik ilkelerin doğrudan insan faktörleri alanına girmesidir.

Bu raporda yer alan standartlar, YZ ve insan faktörleri etkileşiminin önemini vurgulamak amacıyla örnek olarak seçilmiştir. Rehber niteliğindeki bu liste yalnızca yol gösterici niteliktedir; tam bir kapsam sunmaz ve bağlayıcılığı yoktur. Bu alandaki çalışmalarda, ulusal ve uluslararası güncel standartların yanı sıra, konuya ilişkin güncel rapor ve yayınların da detaylı bir şekilde incelenmesi tavsiye edilir.

4.3. Kaynakça

Kaynakça, bu raporun hazırlanması sırasında yararlanılan çalışmaları içermektedir. Rapor açık kaynaklar kullanılarak hazırlanmıştır.

International Organization For Standardization

[1] ISO 9241-210:2019 Ergonomics of human-system interaction — Part 210: Human-centred design for interactive systems. Retrieved from https://www.iso.org/standard/77520.html

[2] ISO/IEC 5338:2023 Information technology — Artificial intelligence — YZ system life cycle processes. Retrieved from https://www.iso.org/standard/81118.html

[3] ISO/IEC/IEEE 15288:2023 Systems and software engineering — System life cycle processes*. Retrieved from https://www.iso.org/standard/81702.html

[4] ISO 9241-11:2018 Ergonomics of human-system interaction — Part 11: Usability: Definitions and concepts. Retrieved from https://www.iso.org/standard/63500.html

[5] ISO 9241-110:2020 Ergonomics of human-system interaction — Part 110: Interaction principles. Retrieved from https://www.iso.org/standard/75258.html

[6] ISO 9241-171:2008 Ergonomics of human-system interaction — Part 171: Guidance on software accessibility. Retrieved from https://www.iso.org/standard/39080.html

[7] ISO 9241-112:2025 Ergonomics of human-system interaction — Part 112: Principles for the presentation of information*. Retrieved from https://www.iso.org/standard/87518.html

[8] International Organization for Standardization. Retrieved from https://www.iso.org/home.html

[9] ISO/TR 16982:2002 Ergonomics of human-system interaction — Usability methods supporting human-centred design*. Retrieved from https://www.iso.org/standard/31176.html

[10] ISO/IEC 25059:2023 Software engineering — Systems and software Quality Requirements and Evaluation (SQuaRE) — Quality model for YZ systems. Retrieved from https://www.iso.org/standard/80655.html

[11] ISO/IEC 23894:2023 Information technology — Artificial intelligence — Guidance on risk management. Retrieved from https://www.iso.org/standard/77304.html

[12] ISO/IEC 42001:2023 Information technology — Artificial intelligence — Management system. Retrieved from https://www.iso.org/standard/42001

[13] ISO/IEC 38507:2022 Information technology — Governance of IT — Governance implications of the use of artificial intelligence by organizations. Retrieved from https://www.iso.org/standard/56641.html

[14] ISO/IEC TS 8200:2024 Information technology — Artificial intelligence — Functional safety and YZ systems. Retrieved from https://www.iso.org/standard/83012.html

Raporlar ve Makaleler:

[14] National Institute of Standards and Technology. (2023). Foundational PNT Profile: Applying the Cybersecurity Framework for the Responsible Use of Positioning,

Navigation, and Timing (PNT) Services (https://nvlpubs.nist.gov/nistpubs/ir/2023/NIST.IR.8323r1.pdf)

[15] Deutsches Institut für Normung. (2022). German standardization roadmap on artificial intelligence (2nd ed.). Retrieved from https://www.dke.de/resource/blob/2017010/9da6a36165da13249ab535eff3af3b03/nr-ki-english---download-data.pdf

[16] Stanford University, Human-Centered Artificial Intelligence. (2024). Artificial intelligence index report 2024. Retrieved from https://hai.stanford.edu/assets/files/hai_ai_index_report_2025.pdf

[17] Simonetta, A., & Paoletti, M. C. (2024). SO/IEC Standards and design of an artificial intelligence system. In CEUR Workshop Proceedings (Vol. 3916, pp. 39-43). Retrieved from https://ceur-ws.org/Vol-3916/paper_07.pdf

[18] National Institute of Standards and Technology. (2023) Artificial Intelligence Risk Management Framework (AI RMF 1.0)(NIST YZ 100-1). Retrieved from https://nvlpubs.nist.gov/nistpubs/ai/NIST.AI.100-1.pdf

[19] Nahar, N., Zhou, S., Zhang, H., Kästner, C., & Lewis, G. (2023). *A Meta-Summary of Challenges in Building Products with ML Components: Collecting Experiences from 4758+ Practitioners. Project Report. Carnegie Mellon University. Retrieved from osf.io/y5edu

[20] John Chris Jones. 1970. Design Methods: Seeds of Human Futures. John Wiley & Sons, New York, NY.

[21] Herbert A. Simon. 1969. The Sciences of the Artificial. MIT Press, Cambridge, MA.

[22] Stefanova-Stoyanova, Varbinka, Petko Danov, and Kamelia Raynova. "Improving Human-Computer Interaction with Generative Adversarial Networks by Covering ISO 9241 Standards in Interaction Design." *2024 32nd National Conference with International Participation (TELECOM)*. IEEE, 2024.

[23] Oviedo, J., Rodriguez, M., Trenta, A., Cannas, D., Natale, D., & Piattini, M. (2024). ISO/IEC quality standards for YZ engineering. *Computer Science Review, 54,* 100681. https://doi.org/10.1016/j.cosrev.2024.100681


Yapay Zekâ Sistemleri ve
Teknik Süreçler

5. YAPAY ZEKÂ SİSTEMLERİ ve TEKNİK SÜREÇLER

5.1. Yapay Zekâ Geliştirme Süreçlerinde Uluslararası Standartların Entegrasyonu: Teknik Yönetici Perspektifi

Ulaş Çakır

Savunma Sanayi Başkanlığı, m.u.cakir@gmail.com

YZ teknolojilerinin hızlı evrimi, standartlaştırma alanında da karmaşık ve çok katmanlı bir yapının ortaya çıkmasına neden olmuştur. Yapılan araştırmalar, YZ standartlarının sıfırdan oluşturulmak yerine, ISO/IEC 12207 gibi mevcut yazılım ve sistem mühendisliği standartlarının üzerine inşa edildiğini ve bu sayede ISO/IEC 5338 [13] gibi yeni standartlarla YZ'ye özgü süreçlerin entegre edildiğini göstermektedir. Bu bütünleşik yaklaşım, NIST YZ RMF'nin esnek ilkeleri [18] ile ISO/IEC 42001'in yapılandırılmış yönetim sistemini [19] bir araya getirerek, YZ projeleri için hem rehberlik hem de sertifikasyon imkânı sunmaktadır. Ancak, özellikle savunma ve havacılık gibi kritik sektörlerdeki DO ve MIL-STD standartlarının YZ'ye özgü süreçlerle nasıl entegre edileceği konusunda hala önemli bir bilgi boşluğu bulunmaktadır. Bu kısım, YZ geliştirme sürecinin her aşamasında hangi standartların nasıl uygulanacağını, veri hazırlığından model geliştirmeye kadar olan süreçleri, söz konusu sektörel standartların bakış açısıyla inceleyecektir.

5.1.1. Yapay Zekâ Sistemi Yaşam Döngüsü ve ISO/IEC 5338 Çerçevesi

YZ sistemlerinin geliştirilmesi, geleneksel yazılım geliştirme yaşam döngüsüne (SDLC) benzer aşamaları içermekle birlikte, YZ'ye özgü tekrarlayan ve döngüsel süreçlerle de zenginleşmiştir. Tipik olarak bu yaşam döngüsü, gereksinim analizi, veri toplama ve hazırlama, model geliştirme, doğrulama, dağıtım ve sürekli bakım aşamalarını içerir [1]. YZ sistemlerinin doğasında bulunan sürekli öğrenme ve adapte olma yeteneği, bu döngünün doğrusal olmaktan ziyade, sürekli bir geri bildirim ve iyileştirme mekanizmasıyla çalışmasını gerektirmektedir [2].

Bu karmaşık yaşam döngüsünü yapılandırmak, yönetmek ve YZ sistemlerinin güvenilirliğini sağlamak için ISO/IEC 5338:2023 standardı kritik bir çerçeve sunmaktadır. "Bilgi teknolojisi - Yapay Zekâ Sistem Yaşam Döngüsü Süreçleri" başlıklı bu standart, YZ'ye özgü süreçlerin etkin bir şekilde tanımlanmasını, kontrol edilmesini ve yönetilmesini kolaylaştıran kapsamlı bir rehberlik sağlar [3]. ISO/IEC 5338'in temel bir prensibi, YZ sisteminin geliştirilmesi ve dağıtımı boyunca başlangıçtan itibaren güvenlik önlemlerinin alınmasını sağlamaya çalışmasıdır. Bu yaklaşım, güvenlik tehditlerinin YZ sisteminin doğal kırılganlıkları göz önüne alınarak proaktif bir şekilde ele alınmasını sağlar [4, 5]. Standart, sürekli risk yönetimi süreçlerini vurgular ve veri kalitesi ve veri kökeni gibi hususların titizlikle belgelenmesini savunur. ISO/IEC 5338, bu şekilde, YZ sistem yaşam döngüsü boyunca güvenliği, güvenilirliği ve etiği temel bir prensip olarak entegre eden ilk kapsamlı yol haritasını sunmaktadır.

Aşağıdaki tablo (Tablo 1), YZ geliştirme süreçlerinde kullanılan temel ISO/IEC standartlarını ve bunların uygulama alanlarını özetlemektedir.

Tablo 1: Yapay Zekâ Uygulama Alanları ve İlgili ISO/IEC Standartları

Uygulama Alanı	**İlgili ISO/IEC Standartları**
Analitik ve Makine Öğrenmesi (ML) için Veri Kalitesi	ISO/IEC 5259-1:2024, ISO/IEC 5259-2:2024, ISO/IEC 5259-3:2024, ISO/IEC 5259-4:2024, ISO/IEC 5259-5:2024, ISO/IEC 25012:2008
YZ Yönetim Sistemi	ISO/IEC 42001:2023
YZ Kavramları ve Terminolojisi	ISO/IEC 22989:2022
Risk Yönetimi Rehberliği	ISO/IEC 23894:2023
Otomatik Yapay Zekâ Sistemlerinin Kontrol Edilebilirliği	ISO/IEC TS 8200:2024, ISO/IEC 25059:2023
Veri Yaşam Döngüsü Çerçevesi	ISO/IEC 8183:2023
YZ Denetimi ve Sertifikasyon Kuruluşları için Gereksinimler	ISO/IEC FDIS 42005, ISO/IEC DIS 42006
Makine Öğrenmesi Kullanan YZ Sistemleri için Çerçeve	ISO/IEC 23053:2022
Etik Kaygılar	ISO/IEC TR 24368:2022, ISO/IEC 25059:2023

Aşağıdaki tablo (Tablo 2) ise, YZ yaşam döngüsü aşamalarının kritik bileşenlerini ve bu bileşenlerle doğrudan ilişkili olan başlıca uluslararası standartları özetlemektedir.

Tablo 2: YZ Yaşam Döngüsü Aşamaları ve İlgili Standartlar

Yaşam Döngüsü Aşaması	Ana Faaliyet Alanı	İlgili Standartlar ve Çerçeveler
Veri	Veri Toplama, Temizleme, Etiketleme, Dönüştürme, Depolama	ISO/IEC 5259 Serisi, ISO/IEC 25012, NIST Yapay Zekâ Risk Yönetim Çerçevesi (AI RMF), ISO/IEC 42001
Model Geliştirme	Model Tasarımı, Eğitimi, Yapılandırma, Test Etme	ISO/IEC 5338, ISO/IEC 23053, DO-178C, DO-330
Doğrulama ve Test	Performans Doğrulama, Ön Yargı Tespiti, Güvenlik Testi	ISO/IEC 5338, NIST YZ RMF, DO-178C
Dağıtım ve Operasyon	Üretime Alma, Sürekli İzleme, Bakım, Sürüm Yönetimi	ISO/IEC 5338, ISO/IEC 42001, MIL-STD
Yönetişim ve Risk	Strateji Belirleme, Risk Değerlendirmesi, Etik Uyum, Hesap Verebilirlik	ISO/IEC 42001, NIST YZ RMF, ISO/IEC FDIS 42005

5.1.2. Yaşam Döngüsü Aşamalarına Göre Standartların Uygulanması

1. Veri Yaşam Döngüsü ve Kalite Yönetimi

YZ sistemlerinin güvenilirliği ve performansı, büyük ölçüde eğitildiği verilerin kalitesine bağlıdır. Verilerin hatalı, eksik veya önyargılı olması, YZ modelinin güvenilmez ve potansiyel olarak ayrımcı sonuçlar üretmesine yol açabilir [14]. Bu nedenle, YZ geliştirme yaşam döngüsünün veri aşaması, etik ve teknik risklerin yönetilmesi için en kritik noktalardan biridir [15]. Bu aşamada, ISO/IEC 5259 serisi ve ISO/IEC 25012 gibi standartların kullanımı hayati önem taşımaktadır.

ISO/IEC 5259 serisi, analitik ve makine öğrenmesi için veri kalitesine özel olarak odaklanan bir dizi standarttır. Serinin her bir bölümü, veri kalitesi yönetiminin farklı yönlerini ele alır: ISO/IEC 5259-1, temel kavramları ve terminolojiyi tanımlayarak genel bir giriş sunar. ISO/IEC 5259-2, veri kalitesi için ölçüm modelleri ve metrikleri belirler; bu standart, veri risklerinin azaltılması için ISO/IEC 42001 (YZ Yönetim Sistemi) ile entegre bir şekilde kullanılır [6]. ISO/IEC 5259-3, veri yaşam döngüsü boyunca uygulanması gereken veri kalitesi yönetimi süreçleri ve pratikleri hakkında rehberlik sağlar. Son olarak, ISO/IEC 5259-4, veri kalitesini sürekli olarak izlemek ve iyileştirmek için yapılandırılmış bir süreç çerçevesi sunar.

ISO/IEC 25012, yazılım mühendisliği bağlamında veri kalitesi modelini tanımlayan daha genel bir standarttır. Bu standart, veri kalitesi özelliklerini iki ana kategoriye ayırır: verinin kendisiyle ilgili olan (Accuracy, Consistency, Completeness, Currentness, Credibility) ve verinin belirli bir sistem içinde kullanımıyla ilgili sisteme bağlı veri kalitesi. Bir YZ sistemi için verinin tamlığı (Completeness) ve tutarlılığı (Consistency) gibi kendisiyle ilgili olan özellikleri, modelin önyargılı sonuçlar üretme riskini doğrudan etkileyen faktörlerdir. ISO/IEC 25012, bu özelliklerin değerlendirilmesi için kullanılabilir.

Bir YZ sisteminin güvenilirliği ve tarafsızlığı, büyük ölçüde eğitim verilerinin kalitesine bağlıdır. Bu durum, etik bir sorunun (önyargı ve ayrımcılık) kökeninin, veriden kaynaklı teknik bir duruma dayandığını göstermektedir. Bu bağlantı, etik tasarım (ethics-by-design) yaklaşımının sadece soyut bir kavram olmadığını, aynı zamanda veri kalitesini yönetmek için ISO/IEC 5259 ve ISO/IEC 25012 gibi teknik standartların titizlikle uygulanmasını gerektiren bir pratik olduğunu gösterir. Bu standartlar, veri yaşam döngüsünün en başından itibaren önyargı ve riskleri azaltmak için bir yol haritası sunar. Bu nedenle, ISO/IEC 5338 çerçevesi içinde veri yaşam döngüsü aşamasında bu standartların uygulanması, YZ sistemlerinin hem teknik olarak sağlam hem de etik olarak sorumlu olmasını sağlamanın ilk ve en kritik adımıdır.

Tablo 3: Veri Kalitesi Yönetimi ve Etik Kapsamları

Standart	Kapsam	Etik ve Güvenilirlik Bağlantısı
ISO/IEC 5259 Serisi	Analitik ve ML için veri kalitesi ölçümü, yönetimi ve sürekli iyileştirilmesi.	Dengesiz veya eksik verilerden kaynaklanan önyargı ve ayrımcılık risklerini azaltır.
ISO/IEC 25012	Veri kalitesi özelliklerinin (Verinin Kendisine Bağlı, Sisteme Bağlı) belirlenmesi ve değerlendirilmesi.	Tamlık, tutarlılık ve doğruluk gibi özelliklerin eksikliği etik olmayan sonuçlara yol açabilir.
ISO/IEC 8183	YZ'de veri yaşam döngüsü için bir çerçeve sunar.	Verinin kökenini (provenance) ve yaşam döngüsünü izleyerek hesap verebilirliği artırır.

2. Model Geliştirme, Doğrulama ve Operasyonel Gözetim

Model geliştirme aşaması, YZ sistemlerinin çekirdeğini oluşturur ve ISO/IEC 23053 bu süreç için temel bir çerçeve sunar. Bu aşamada, modelin doğru bir şekilde eğitildiğinden ve doğrulandığından emin olmak kritik öneme sahiptir. Doğrulama, modelin eğitimde kullanılmayan verilerle test edilmesi yoluyla yapılır. Bu adım, modelin yeni girdi değerlerine ne kadar iyi tepki verebileceğini gösterir.

Geleneksel yazılımdan farklı olarak, YZ modellerinin performansı zaman içinde veri sapması (data drift) gibi nedenlerle düşebilir veya bozulabilir. Bu durum, geleneksel yazılım

mühendisliğindeki "dağıt ve unut" yaklaşımından köklü bir ayrılışı temsil etmektedir. YZ sistemlerinin güvenilirliğini sürdürmek için sürekli operasyonel gözetim (AIOps) hayati hale gelmiştir. ISO/IEC 5338'in sürekli doğrulama (continuous validation) gerekliliği ve ISO/IEC 42001'in sürekli izleme ve iyileştirme (continuous monitoring & improvement) döngüsü, bu yeni operasyonel ihtiyacı karşılamaktadır. Bu döngü, YZ sistemlerinin performansının sürekli olarak gözden geçirilmesi ve yönetişim stratejilerinin güncellenmesi için bir süreç sağlar.

Bu bağlamda, YZ operasyonel yönetiminde (AIOps) başarı, anomali tespiti ve veri sapması analizi gibi teknik izleme araçlarını, ISO/IEC 42001 tarafından desteklenen tanımlı yönetim süreçleriyle birleştirmeyi gerektirir. Bu yaklaşım, operasyonel sapmaları organizasyonel riskler ve iş hedefleriyle ilişkilendirerek proaktif bir şekilde yönetmeyi mümkün kılar.

3. YZ Yönetim Sistemi, Etik ve Risk Yönetimi

YZ sistemlerinin güvenilirliği, sadece teknik performanstan ibaret değildir; aynı zamanda bu sistemleri geliştiren ve kullanan kuruluşların etik ve sorumlu yönetim süreçlerini ne kadar benimsediğiyle de yakından ilişkilidir. YZ riskleri, sadece algoritmik hatalardan değil, aynı zamanda organizasyonel eksikliklerden de kaynaklanabilir. Bu nedenle, YZ yönetim sistemleri, YZ'nin kurumsal risk yönetimi ve etik bağlamına entegrasyonu için bir temel sağlamaktadır [7]. Bu aşamalarda, ISO/IEC 42001, ISO/IEC 23894 ve ISO/IEC TR 24368 standartları kritik bir rol oynamaktadır.

ISO/IEC 42001:2023, bir YZ Yönetim Sistemi (YZYS) kurmak, uygulamak, sürdürmek ve sürekli iyileştirmek için gereken kriterleri tanımlayan temel bir standarttır. Bu standardın amacı, sorumlu bir şekilde geliştirilen ve kullanılan YZ sistemlerini teşvik etmektir. ISO/IEC 42001, etik ilkeler gibi adillik, ayrımcılık yapmama ve gizliliğe saygı konularının önemini vurgulayarak, YZ'yi kurumsal düzeyde yönetişim ve liderliğe dahil etmeyi amaçlamaktadır. Ayrıca, bu standart, ISO/IEC 27001 (Bilgi Güvenliği Yönetim Sistemi) gibi diğer yönetim sistemleriyle entegre edilebilecek bir Yüksek Seviyeli Yapıya (High Level Structure) sahiptir. Bu entegrasyon, YZ'ye özgü risklerin mevcut kurumsal risk yönetimi süreçlerine dâhil edilmesini kolaylaştırır.

ISO/IEC 23894:2023, YZ kullanımından kaynaklanan risklerin yönetimi konusunda somut rehberlik sunar. Bu standart, güvenilir YZ'nin temelinin risk yönetimi olduğunu vurgulamaktadır. ISO/IEC 23894, risk yönetimini yalnızca teknik bir süreç olmaktan çıkarıp, daha geniş sosyal ve ekonomik hedeflerle ilişkilendirir [17]. Standart, risk yönetimi sürecinin

amacını, değer yaratmak ve korumak, performansı artırmak ve hedeflere ulaşmayı desteklemek olarak tanımlar. Bu standart, NIST YZ RMF gibi çerçevelerle de uyumluluk göstererek, genel amaçlı YZ modelleri için risk yönetimi uygulamalarını güçlendirir.

ISO/IEC TR 24368:2022, YZ'nin yol açtığı etik ve sosyal kaygılara kapsamlı bir genel bakış sunar. Bu standart, YZ sistemlerinin şeffaflık, hesap verebilirlik ve önyargılardan kaçınma gibi etik ilkeleri teknik gereksinimlerle birleştirerek, sorumlu YZ geliştirme için pratik bir çerçeve sağlar. Örneğin, şeffaflık kavramı, *ISO/IEC 25059'da* teknik bir özellik olarak ele alınırken, bu ISO/IEC 24368'de kullanıcıların YZ kararlarını anlayıp güvenebilmelerinin bir yolu olarak daha geniş bir çerçevede ele alınır.

ISO/IEC 42001 ve ISO/IEC 23894 gibi standartlar, algoritmik detaylardan ziyade organizasyonel yapıları ve süreçleri tanımlamaktadır. Bu durum, YZ'ye özgü risklerin sadece teknik birimlerin değil, tüm kuruluşun (üst yönetim, hukuk, etik kurulları, vb.) sorumluluğunda olduğunu göstermektedir [16]. Risk yönetiminin üst düzey yönetim ve denetim gerektiren bir kurumsal fonksiyon olarak ele alınması, etik ve sosyal hedeflere ulaşmak için teknik çözümlerin ötesinde bütüncül bir yönetim çerçevesine ihtiyaç duyulduğunu ortaya koyar. Bu standartlar, YZ'nin potansiyel risklerini sistematik olarak ele almak ve farklı departmanları ortak bir amaç etrafında birleştirmek için bir köprü görevi görmektedir. Bu yaklaşım, sadece yasal uyumluluğu sağlamakla kalmaz, aynı zamanda YZ'ye olan güveni de artırmayı hedefler.

Tablo 4: Yapay Zekâ Yönetişimi ve Risk Yönetimi Standartları

Standart	Kapsam	Anahtar Rolü
ISO/IEC 42001	Yapay Zekâ Yönetim Sistemi (YZYS) kurulması.	YZ'yi kurumsal politikalara, yönetişim ve risk yönetimi süreçlerine entegre etmek.
ISO/IEC 23894	YZ risk yönetimi için rehberlik.	YZ yaşam döngüsü boyunca risklerin tanımlanması, değerlendirilmesi ve azaltılması için somut pratikler sağlamak.
ISO/IEC TR 24368	YZ'nin etik ve sosyal kaygıları.	Teknik standartlarla etik ilkeleri birleştirerek sorumlu YZ geliştirme çerçevesi sunmak.

4. Güvenilirlik, Şeffaflık ve Güvenlik

YZ sistemlerinin karar alma süreçlerinin şeffaf olmaması (kara kutu sorunu), güvenilirlik ve hesap verebilirlik konularında ciddi endişeleri beraberinde getirmektedir [12]. Bu durum, kullanıcı güvenini sarsarak YZ teknolojilerinin benimsenmesini zorlaştırmaktadır. Bu zorlukların üstesinden gelmek için, ISO/IEC 25059 ve ISO/IEC TS 8200 gibi standartlar, YZ sistemlerinin iç işleyişini daha anlaşılır, kontrol edilebilir ve potansiyel saldırılara karşı daha dayanıklı hale getirmeyi amaçlamaktadır [10].

ISO/IEC 25059:2023, yazılım ve sistem kalitesi gereksinimleri ve değerlendirmesi (Systems and software Quality Requirements and Evaluation - SQuaRE) serisinin bir parçası olarak YZ sistemlerine özgü yeni kalite özellikleri tanımlar. Bu özellikler arasında şeffaflık, kontrol edilebilirlik ve işlevsel uyarlanabilirlik gibi etik zorluklarla doğrudan ilgili olan kavramlar yer alır. Şeffaflık, bu standarda göre, paydaşların YZ sisteminin nasıl çalıştığına dair ayrıntılı bilgilere erişebilmesi anlamına gelir. Kontrol edilebilirlik, insan müdahalesini mümkün kılarak hesap verebilirlik ilkesini destekler. Bununla birlikte, bu standardın bazı noktalarda insan gözetimi ve etik bütünlük gibi konularda boşluklar barındırdığı ve bu alanlarda diğer standartlarla tamamlanması gerektiği belirtilmektedir.

ISO/IEC TS 8200:2024, otomatik YZ sistemlerinin kontrol edilebilirliği konusunda rehberlik sunar. Bu, özellikle kritik durumlarda veya beklenmeyen sonuçlarla karşılaşıldığında insan müdahalesinin sağlanması açısından temel bir gerekliliktir. Bu standart, YZ sistemlerinin otonom operasyonları ile insan gözetimi arasındaki dengeyi kurmayı hedefler.

Güvenlik, YZ yaşam döngüsünün yalnızca sonradan düşünülen bir yönü olmaktan çıkıp, ISO/IEC 5338 standardının da vurguladığı gibi, en başından itibaren temel bir tasarım prensibi haline gelmelidir. Geleneksel siber güvenlik risklerinin yanı sıra, YZ sistemleri, algoritmaların manipüle edilmesini hedefleyen adversarial saldırılar gibi tehditlere karşı da savunmasızdır. ISO/IEC 5338, YZ'nin bu özel güvenlik kırılganlıklarını tanıyarak, sürekli risk yönetimi ve sürekli doğrulama süreçlerini zorunlu kılar. Bu, veri kümesi kayması (data drift) veya kavram kayması (concept drift) gibi sistemin bütünlüğünü tehlikeye atabilecek durumları düzenli olarak test etmeyi içerir.

Geleneksel yazılım geliştirme yaklaşımlarının aksine, YZ'nin doğasında bulunan çekişmeli (adverserial) saldırılara karşı kırılganlığı, güvenliğin veri toplama ve model geliştirme aşamasından itibaren ele alınmasını gerektirmektedir. ISO/IEC 5338 gibi standartlar, bu kolektif sorumluluğu teknik ve süreçsel gerekliliklerle somutlaştırarak, YZ sistemlerinin doğuştan güvenli (secure-by-design) olmasını sağlamayı hedeflemektedir. Bu yaklaşım, YZ güvenliğinin, sadece siber güvenlik ekibinin değil, veri bilimciler, model geliştiriciler ve hatta iş birimi liderlerinin ortak sorumluluğu haline geldiğini göstermektedir.

Tablo 5: Güvenilirlik Özellikleri ve İlgili Standartlar

Özellik	Kapsam	İlgili Standartlar
Şeffaflık	Sistemin nasıl çalıştığına dair bilgilere erişim, "kara kutu" etkisini azaltma.	ISO/IEC 25059, ISO/IEC TR 24368
Kontrol Edilebilirlik	İnsan müdahalesinin mümkün ve etkili olması.	ISO/IEC TS 8200, ISO/IEC 25059
Güvenlik	Adversarial saldırılar gibi tehditlere karşı dayanıklılık.	ISO/IEC 5338, ISO/IEC 25059

5.1.3. Sektörel Standartlar ve Entegrasyon

YZ sistemlerinin havacılık, savunma ve diğer yüksek riskli veya özel gereksinimli alanlardaki uygulamaları, uluslararası standartlara ek olarak, sektöre özel katı düzenlemelerle de yönetilmektedir. Bu standartlar, YZ sisteminin kendisinden ziyade, onun çalıştığı donanım ve yazılım platformlarının güvenlik, güvenilirlik ve çevresel dayanıklılık gereksinimlerini belirler.

1. NIST Yapay Zekâ Risk Yönetimi Çerçevesi (AI RMF)

ABD'de yaygın olarak kullanılan bir çerçeve olan *NIST YZ RMF*, YZ ile ilişkili riskleri yönetmek için kapsamlı bir rehberlik sunmaktadır. Özellikle genel amaçlı YZ (General Purpose AI) ve temel modeller [8] (foundation models) gibi karmaşık sistemler için tasarlanmış olan bu çerçeve, ISO/IEC 23894 gibi uluslararası standartlarla uyumluluk göstererek, onlara özgü risk yönetim süreçleri için somut bir profil oluşturur [21].

NIST YZ RMF, risk yönetimini dört ana fonksiyona ayırarak yapılandırır:

- **Yönetişim (Govern):** YZ risk yönetimi için kurumsal politikaları, rolleri, sorumlulukları ve şeffaf bir kültürü belirler.
- **Haritalama (Map):** YZ sisteminin amacını, potansiyel yararlarını ve risklerini, kullanım senaryoları ve etkileşimde bulunduğu paydaşlar bağlamında anlamayı ve belgelemeyi hedefler.
- **Ölçme (Measure):** Güvenilir YZ özelliklerinin değerlendirilmesi için uygun yöntem ve metrikleri uygular, bu sayede sistemin önyargıları, güvenliği ve performansını ölçer.
- **Yönetme (Manage):** Değerlendirilen risklere yanıt verme, onları önceliklendirme ve azaltma stratejilerini planlama ve uygulama işlevini içerir.

NIST YZ RMF, YZ'nin yasal düzenlemelerle karşılaşacağı bir geleceğe proaktif bir hazırlık aracı olarak değerlendirilmelidir. ISO/IEC 42001'in Avrupa Birliği YZ Yasası (AIA) gerekliliklerini yerine getirmek için bir temel oluşturması gibi [11, 20], bu çerçeveler yasal zorunlulukların pratikte nasıl uygulanacağına dair somut bir yol haritası sunar. Bir YZ sistemini standartlara uygun geliştirmek, sadece iyi uygulama olmaktan öte, olası bir yasal denetimde güçlü bir kanıt sunma yeteneği sağlar [9]. Bu durum, pek çok standardın, henüz yasal zorunluluğu olmayan alanlarda bile endüstri standardı haline gelme potansiyeli taşıdığını gösterir.

2. Havacılık (DO Serisi)

Genel standartların yanı sıra bazı sektörlere göre düzenlenmiş standartlar da mevcuttur. Bu standartlar, YZ sistemlerinin kendisinden ziyade, bu sistemleri barındıran donanım ve yazılım platformlarının güvenlik, güvenilirlik ve çevresel dayanıklılık gereksinimlerini belirlemek üzere organize edilmişlerdir. Bu durum, YZ'nin sadece yazılım ve veri bilimiyle sınırlı kalmayıp, sistem mühendisliği ve donanım tasarımı gibi alanlarla da iç içe geçtiği durumları da ele almak için yol gösterici nitelikte olduğunu göstermektedir.

Havacılık için DO-178C ve DO-254:

- ***DO-178C,* hava araçlarındaki yazılımlar için tasarım güvencesi ve ürün güvencesi tanımlayan temel bir standarttır.** Uçuş yazılımlarının güvenilirliğini ve kalitesini sağlamak için kullanılan, havacılık endüstrisi için bir standartlar ve kılavuzlar bütünüdür. RTCA (Radio Technical Commission for Aeronautics) tarafından yayınlanan bu standart, yazılımın tasarlanmasından test edilmesine kadar tüm yaşam döngüsünü kapsar ve yazılımın güvenlik kritikliği seviyesine göre (A'dan E'ye kadar) titiz bir doğrulama süreci gerektirir. Bu kılavuz, havacılık sistemlerinde kullanılan yazılımların işlevsel güvenliğini sağlamak için kapsamlı bir süreç ve kanıt zinciri oluşturmayı hedefler.

  YZ geliştiricileri açısından DO-178C, geleneksel yazılım geliştirme süreçlerinden farklılık gösteren YZ modellerinin doğrulama ve validasyonuna özel dikkat gerektirir. YZ algoritmalarının deterministik olmaması, test ve doğrulama süreçlerini karmaşıklaştırır. Bu nedenle, YZ geliştiricilerinin, modelin öğrenme verilerini, eğitim sürecini ve karar alma mantığını şeffaf bir şekilde belgelemeleri ve her bir çıktının neden bu şekilde üretildiğini gösterebilmeleri hayati önem taşır. Ek olarak, modelin beklenen ve beklenmeyen koşullar altında nasıl davrandığını kanıtlamak, potansiyel hataları ve

sapmaları belirlemek için kapsamlı bir test stratejisi oluşturmak ve bu süreçleri DO-178C'nin gerektirdiği titizlik seviyesine uygun hale getirmek zorundadırlar.

- ***DO-254* ise hava araçlarındaki elektronik donanımların tasarımı için rehberlik sağlar.** Havacılık sistemlerinde kullanılan karmaşık donanım (özellikle FPGA, ASIC ve PLD gibi programlanabilir donanımlar) için bir geliştirme, doğrulama ve sertifikasyon standardıdır. RTCA tarafından yayınlanan bu belge, donanımın güvenlik ve güvenilirlik gereksinimlerini karşılamasını sağlamak amacıyla donanım geliştirme yaşam döngüsünün her aşamasında uyulması gereken kılavuzları belirler. Tıpkı DO-178C'nin yazılım için yaptığı gibi, DO-254 de donanımın tasarımından, entegrasyonuna ve testine kadar tüm süreçlerin izlenebilir, kontrollü ve belgelenmiş olmasını şart koşar.

  DO-254, donanım katmanında yer alan YZ işlemcileri veya YZ hızlandırıcıları gibi özel donanım bileşenleri söz konusu olduğunda kritik bir öneme sahiptir. Geleneksel donanım geliştirmeden farklı olarak, YZ donanımları genellikle sinir ağları gibi esnek ve uyarlanabilir yapılara dayanır. Bu durum, donanım tasarımının "statik" ve tamamen önceden tanımlanmış olmasını zorunlu kılan DO-254 gereksinimleriyle çelişebilir. Bu nedenle, YZ geliştiricileri, donanımın donanım bileşenleri üzerindeki YZ algoritmalarının nasıl çalıştığını, potansiyel hataları ve belirsizlikleri nasıl yöneteceğini ve performans sapmalarının nasıl minimize edileceğini titizlikle belgelemelidir. Ayrıca, YZ donanımının doğruluğunu ve güvenilirliğini, uçuşa elverişli bir sistemde kullanılmak üzere kanıtlamak için kapsamlı donanım doğrulama testleri ve hata analizi süreçleri uygulamaları gerekmektedir.

Bu standartların temel gerekliliği, sistem gereksinimleri ile tasarım, kodlama, doğrulama ve test faaliyetleri arasında net, doğrulanabilir bir "izlenebilirlik" (traceability) sağlamaktır. YZ tabanlı bir sistemin bu standartlara uygunluğu, her bir nöral ağ katmanının veya algoritma parçasının, tanımlanmış bir sistem gereksinimine nasıl katkıda bulunduğunun kanıtlanmasını gerektirir. Bu izlenebilirlik, hata ayıklama süreçlerini kolaylaştırır, güvenlik risklerini yönetmeye yardımcı olur ve sertifikasyon sürecini destekler.

3. Savunma Sanayii (MIL-STD)

Savunma sanayi, YZ sistemlerinin hızla benimsendiği bir alandır; ancak bu entegrasyon, sistemlerin güvenilirlik ve dayanıklılık açısından en yüksek standartları karşılamasını gerektirir. MIL-STD-461 ve MIL-STD-810 gibi askeri standartlar, bu kritik gereksinimlerin karşılanması için bir çerçeve sunar. YZ algoritmaları ne kadar sofistike olursa olsun, eğer

çalıştıkları donanım fiziksel ve elektromanyetik olarak zorlu ortamlara karşı savunmasızsa, sistemin işlevini kaybetmesine neden olabilir. Bu nedenle, YZ geliştirme süreçlerinin, sadece yazılımsal doğrulama ile sınırlı kalmayıp, donanımın çevresel faktörlere (titreşim, sıcaklık, nem) ve elektromanyetik parazitlere karşı dayanıklılığını güvence altına alacak şekilde bu standartlarla uyumlu hale getirilmesi hayati önem taşır. Bu uyum, YZ destekli sistemlerin savaş sahası gibi en zorlu koşullarda bile kararlı, güvenilir ve öngörülebilir bir şekilde çalışmasını sağlamanın anahtarıdır.

- ***MIL-STD-810*** **(Çevresel Mühendislik), bir ekipmanın hizmet ömrü boyunca karşılaşacağı çevresel koşullara (sıcaklık, nem, titreşim, şok vb.) dayanıklılığını test etmeyi hedefler.** ABD Savunma Bakanlığı (DoD) tarafından belirlenen bir standartlar dizisidir ve askeri ekipmanların çevresel koşullara dayanıklılığını test etmek için kullanılır. Bu standart, bir ürünün askeri ortamlarda (aşırı sıcaklık, nem, titreşim, şok, tuzlu sis, kum, toz, rüzgar ve basınç gibi) nasıl performans gösterdiğini değerlendirir. MIL-STD-810'un amacı, askeri donanımların zorlu operasyonel şartlarda dahi güvenilir bir şekilde çalışmaya devam ettiğinden emin olmaktır. Bu standart, belirli bir test yöntemi veya prosedürü değil, bir dizi test metodu ve kılavuzudur; bu sayede üreticiler, ürünlerinin belirli çevresel zorluklara ne kadar dayanıklı olduğunu kanıtlayabilirler.

  MIL-STD-810, fiziksel donanım üzerinde çalışan YZ sistemleri için büyük önem taşır. YZ sistemleri genellikle yüksek performanslı işlemciler ve sensörler kullanır; bu bileşenler, sıcaklık değişimleri, titreşim ve şok gibi çevresel streslere karşı hassas olabilir. Bu nedenle, YZ sistemlerini geliştiren mühendislerin, algoritmaların sadece yazılımsal olarak değil, aynı zamanda fiziksel olarak da çevresel koşullardan en az düzeyde etkilendiğinden emin olmaları gerekir. Geliştiriciler, donanım platformunun MIL-STD-810 testlerinden geçtiğini ve YZ donanımının aşırı sıcaklıklarda, titreşimli ortamlarda veya yüksek irtifalarda dahi doğru ve stabil bir şekilde çalıştığını doğrulamalıdır. Aksi takdirde, YZ modelinin karar verme yeteneği fiziksel çevresel faktörler nedeniyle tehlikeye girebilir.

- ***MIL-STD-461*** **(Elektromanyetik Uyumluluk),** ekipmanların elektromanyetik ortamda düzgün çalışmasını sağlamak için test gereksinimlerini belirler. Askeri sistemlerde kullanılan elektrikli ve elektronik ekipmanların elektromanyetik uyumluluğunu (EMC) kontrol eden bir standartlar dizisidir. Bu standart,

ekipmanların hem istenmeyen elektromanyetik emisyonlar üretmesini hem de harici elektromanyetik girişimlere karşı bağışıklık göstermesini sağlamak için tasarlanmıştır. Bu, bir sistemin operasyon sırasında diğer sistemlerin çalışmasına engel olmamasını ve aynı zamanda çevresel elektromanyetik gürültüden etkilenmeyerek güvenilir bir şekilde çalışmasını garanti altına alır. MIL-STD-461, bu gereksinimleri test etmek için çeşitli metotlar ve prosedürler sunar ve özellikle kara, hava ve deniz platformları gibi zorlu ortamlarda kritik önem taşır.

YZ sistemlerinin donanım katmanında ne kadar sağlam olması gerektiğine dair önemli ipuçları sunar. YZ algoritmaları, genellikle yüksek hızlı sinyal işleme gerektiren işlemciler, bellekler ve sensörler gibi bileşenler üzerinde çalışır. Bu bileşenler, elektromanyetik parazitlere karşı oldukça hassas olabilir. Yüksek frekanslı işlemci faaliyetleri, doğal olarak elektromanyetik emisyonlar üretebilir ve bu da çevredeki diğer hassas sistemleri etkileyebilir. Benzer şekilde, harici kaynaklardan gelen radyo frekansı (RF) gürültüsü veya elektromanyetik darbeler (EMP), YZ donanımının performansını düşürebilir, hatalı veriye veya modelin kararsız çalışmasına neden olabilir. Bu nedenle, YZ geliştiricilerinin, donanım tasarımlarının MIL-STD-461'in gerektirdiği elektromanyetik girişim (EMI) ve elektromanyetik duyarlılık (EMS) testlerinden geçecek şekilde filtreleme ve topraklama tekniklerini uygulamaları şarttır. Bu, YZ destekli sistemlerin savaş alanında veya kritik görevlerde güvenilir bir şekilde çalışmasını garantilemek için hayati bir adımdır.

Bu standartlar, özellikle askeri uygulamalar gibi zorlu koşullarda çalışan YZ sistemlerinin, çevresel faktörler nedeniyle arızalanmamasını veya performans kaybı yaşamamasını garanti eder. En iyi YZ modeli bile, güvenilir bir donanım platformunda çalışmadığı sürece işlevsiz kalabilir. Bu nedenle, YZ'nin fiziksel dünyadaki güvenilirliği, onu barındıran fiziksel ve yazılımsal platformun da doğrulanmış ve sertifikalandırılmış olmasını gerektirir.

Havacılık ve savunma gibi yüksek riskli sektörlerde, bir YZ sisteminin standartlara uyumunun sağlanması, sadece modelin kendisinin güvenilirliğine değil, aynı zamanda onu barındıran fiziksel ve yazılımsal platformun da doğrulanmış ve sertifikalandırılmış olmasına bağlıdır. Bu durum, ISO/IEC 5338 gibi üst düzey yaşam döngüsü standartlarının, sektöre özgü sertifikasyon standartlarını kapsayacak şekilde genişlemesi gerektiğini göstermektedir. Bu

disiplinler arası yaklaşım, YZ'nin sadece yazılım mühendisliği ve veri bilimi disiplinleriyle sınırlı kalmayıp, aynı zamanda sistem mühendisliği ve donanım tasarımı gibi alanlarla da iç içe geçtiğini ortaya koymaktadır.

4. Değerlendirmeler

YZ teknolojilerinin gelişimi, standartlaştırma alanında da kapsamlı bir dönüşümü beraberinde getirmiştir. Analizler, YZ standartlarının artık parçalı ve izole bir yapıdan, birbirini tamamlayan ve entegre bir ekosisteme dönüştüğünü açıkça göstermektedir. Bu ekosistemde, ISO/IEC 5338, YZ sistemlerinin yaşam döngüsü için yol gösterici bir ana çerçeve olarak öne çıkmaktadır. Bu çerçeve, YZ'ye özgü riskleri ve gereklilikleri sistematik bir şekilde ele alarak, YZ sistemlerinin "ethics-by-design" ve "secure-by-design" yaklaşımlarıyla geliştirilmesini sağlamaktadır.

Yönetim katmanında, ISO/IEC 42001 gibi standartlar, YZ'yi kurumsal politikalara ve yönetişim süreçlerine entegre ederek, YZ'nin sadece teknik bir sorun olmaktan çıkıp, tüm kuruluşun sorumluluğu haline geldiğini vurgulamaktadır. Bu yönetim çerçevesi, ISO/IEC 23894 gibi risk yönetimi standartlarıyla desteklenerek, YZ'den kaynaklanan risklerin sistematik olarak tanımlanmasını, değerlendirilmesini ve yönetilmesini sağlamaktadır. Teknik seviyede ise, ISO/IEC 5259 serisi ve ISO/IEC 25012 gibi veri kalitesi standartları, YZ modellerinin güvenilirliğini ve etik çıktılarını doğrudan etkileyen en temel unsurları ele almaktadır. Veriden kaynaklanan önyargı ve risklerin yönetilmesi, YZ'nin güvenilirliğini sağlamanın ilk ve en kritik adımıdır.

YZ sistemlerinin askeri ve havacılık gibi kritik sektörlerde yaygınlaşması, bu sistemlerin sadece algoritmik olarak değil, aynı zamanda fiziksel ve çevresel olarak da en yüksek güvenilirlik ve emniyet standartlarını karşılamasını zorunlu kılıyor. Bu bağlamda, DO-178C (yazılım), DO-254 (donanım), MIL-STD-810 (çevresel dayanıklılık) ve MIL-STD-461 (elektromanyetik uyumluluk) gibi standartlar, YZ geliştirme süreçlerinin ayrılmaz bir parçası haline gelmektedir. YZ geliştiricileri, artık sadece algoritmaların doğruluğunu değil, aynı zamanda bu algoritmaları barındıran donanımın ve yazılımın, zorlu operasyonel koşullar altında dahi öngörülebilir ve güvenilir bir şekilde çalışabildiğini kanıtlamak zorundadır. Bu standartlar, YZ sistemlerinin karmaşık ve belirsiz yapısını yönetmek için bir yol haritası sunar, potansiyel hataları ve riskleri en aza indirir ve en nihayetinde YZ'nin kritik görevlerde güvenle kullanılmasının önünü açar. Dolayısıyla, bu standartlara uyum, YZ'nin kritik sistemlere entegrasyonu için sadece bir gereklilik değil, aynı zamanda bir zorunluluktur.

Başarılı bir YZ projesi için, bu standartların her birinin ayrı ayrı uygulanması değil, YZ yaşam döngüsü boyunca birbiriyle tutarlı ve entegre bir şekilde kullanılması esastır. Gönüllü standartlar ve çerçeveler (örneğin NIST YZ RMF), gelecekteki yasal zorunluluklara (örneğin AB YZ Yasası) proaktif bir şekilde hazırlanmak için kritik araçlar olarak işlev görmektedir. Bu nedenle, kuruluşların YZ süreçlerini şimdiden bu standartlara göre yapılandırması, sadece yasal uyum risklerini azaltmakla kalmayacak, aynı zamanda YZ sistemlerine olan güveni artırarak bu teknolojinin daha geniş bir alanda başarılı bir şekilde benimsenmesini sağlayacaktır.

Belirtilen standartlar ekosisteminden yola çıkarak, başarılı bir YZ projesi geliştirmek isteyen kuruluşlar için entegre bir yaklaşım benimsemek esastır. Başlangıç noktası olarak, ISO/IEC 5338 standardını bir ana çerçeve olarak kullanarak YZ sisteminin tüm yaşam döngüsünü yapılandırmak hayati önem taşır. Bu ana çerçeve altında, yönetim ve organizasyonel seviyede ISO/IEC 42001 ile YZ yönetişimini kurumsal politikalara dahil etmek, teknik seviyede ise ISO/IEC 5259 gibi standartlarla veri kalitesini ve önyargı yönetimini güvence altına almak gerekir. Son olarak, özellikle askeri ve havacılık benzeri kritik sektörlerde kullanılacak sistemler için, donanım ve yazılımın çevresel ve elektromanyetik dayanıklılığını güvence altına almak amacıyla MIL-STD-810, MIL-STD-461, DO-178C ve DO-254 gibi sektörel standartlarla uyumluluğun sağlanması zorunludur. Bu entegre ve çok katmanlı yaklaşım, sadece yasal uyum risklerini azaltmakla kalmayacak, aynı zamanda YZ sistemlerinin güvenilirliğini ve güvenilirliğini artırarak bu teknolojinin daha geniş bir alanda başarılı bir şekilde benimsenmesini sağlayacaktır.

5.1.4. Kaynakça

[1] Simonetta, A., & Vetrò, A. (2024). ISO/IEC Standards and Design of an Artificial Intelligence System. In CEUR Workshop Proceedings (Vol. 3916). CEUR-WS. https://ceur-ws.org/Vol-3916/paper_07.pdf

[2] Simonetta, A., Vetrò, A., Paoletti, M. C., & Torchiano, M. (2021). Integrating SQuaRE data quality model with ISO 31000 risk management to measure and mitigate software bias. In Proceedings of the 3rd International Workshop on Experience with SQuaRE Series and Its Future Direction (IWESQ 2021) (Vol. 3114, pp. 17-22). CEUR-WS. http://ceur-ws.org/Vol-3114/paper-04.pdf

[3] Simonetta, A., & Paoletti, M. C. (2022). The use of maximum completeness to estimate bias in AI-based recommendation systems. In R. Fazzolari & A. A. Jaber (Eds.), Proceedings of the Scholars' Yearly Symposium of Technology, Engineering, and Mathematics (SYSTEM) (Vol. 3360, pp. 76-84). CEUR-WS. http://ceur-ws.org/Vol-3360/

[4] Simonetta, A., Trenta, A., Paoletti, M. C., & Vetrò, A. (2021). Metrics for identifying bias in datasets. In ICYRIME 2021 International Conference of Yearly Reports on Informatics Mathematics, and Engineering 2021 (Vol. 3118, pp. 10-17). CEUR-WS. https://ceur-ws.org/Vol-3118/p02.pdf

[5] Simonetta, A., Nakajima, T., Paoletti, M. C., & Venticinque, A. (2022). Fairness metrics and maximum completeness for the prediction of discrimination. In Woodstock'21: Symposium on the irreproducible science (Vol. 3356, pp. 13-20). CEUR-WS. http://ceur-ws.org/Vol-3356/

[6] McIntosh, T. R., Susnjak, T., Liu, T., Watters, P., Nowrozy, R., & Halgamuge, M. N. (2024). From COBIT to ISO 42001: Evaluating Cybersecurity Frameworks for Opportunities, Risks, and Regulatory Compliance in Commercializing Large Language Models. arXiv. https://doi.org/10.48550/arXiv.2402.15770

[7] Gualo, F., Rodriguez, M., Verdugo, J., Caballero, I., & Piattini, M. (2021). Data quality certification using ISO/IEC 25012: Industrial experiences. Journal of Systems and Software, 176, Article 110938. https://doi.org/10.1016/j.jss.2021.110938.

[8] Barrett, A. M., Newman, J., Madkour, N., Nonnecke, B., Hendrycks, D., Murphy, E. R., Jackson, K., & Raman, D. (2025). YZ Risk-Management Standards Profile for General-Purpose YZ Systems (GPAIS) and Foundation Models (Version 1.1). Center for Long-Term Cybersecurity, University of California, Berkeley. https://cltc.berkeley.edu/publication/ai-risk-management-standards-profile-v1-1/

[9] Del Castillo, F., & Al-Naddaf, M. (2023). A new YZ hazard management framework. arXiv. https://doi.org/10.48550/arXiv.2310.16727

[10] Sun, K., Zhu, Z., & Lin, Z. (2025). Integrating ISO/IEC standards for robust YZ security and resilience against adversarial attacks. International Journal of Advanced Smart Convergence, 14(1), 85–100. https://doi.org/10.7236/IJASC.2025.14.1.85

[11] Hohl, M., & Simonetta, A. (2024). Interplay of ISMS and AIMS in context of the EU YZ Tüzüğü. https://doi.org/10.48550/arXiv.2412.18670

[12] Othman, A. S., Al-Otaibi, M., & Al-Amri, S. (2024). The Role of Explainable Artificial Intelligence (XAI) in Achieving Trustworthiness in YZ Systems. arXiv:2412.11384.

[13] International Organization for Standardization. (2023). Information technology—Artificial intelligence—AI system life cycle processes (ISO/IEC 5338:2023). ISO.

[14] Gualo, F., Rodriguez, M., Verdugo, J., Caballero, I., & Piattini, M. (2021). Data quality certification using ISO/IEC 25012: Industrial experiences. Journal of Systems and Software, 176, Article 110938. https://doi.org/10.1016/j.jss.2021.110938

[15] Al-Bayati, A. M., & Al-Bayati, S. A. (2024). Trustworthy YZ from principles to practice: A framework for ensuring ethical, reliable, and secure artificial intelligence systems. arXiv. https://doi.org/10.48550/arXiv.2403.16808

[16] The European Commission (2019). Ethics Guidelines for Trustworthy AI.

[17] The OECD (2019). OECD YZ Principles.

[18] NIST (2023). NIST YZ Risk Management Framework (AI RMF).

[19] International Organization for Standardization. (2023). Information technology—Artificial intelligence—Management system (ISO/IEC 42001:2023). ISO/IEC

[20] Pötsch, J. (2024). Interplay of ISMS and AIMS in context of the EU AI Act. arXiv. https://doi.org/10.48550/arXiv.2412.18670

[21] Schnitzer, R., Hapfelmeier, A., Gaube, S., & Zillner, S. (2024). YZ hazard management: A framework for the systematic management of root causes for YZ risks. arXiv. https://doi.org/10.48550/arXiv.2310.16727

5.2. Yazılım Mühendisliği Açısından Yapay Zekâ Standartları ve Düzenlemelerinin Değerlendirilmesi

Dr. Cemal Gemci

TBD Yönetim Kurulu Üyesi (Sayman), OSTİM Teknik Üniversitesi, cgemci@gmail.com

---

YZ, onlarca yıldır dünya çapında oldukça aktif ve yoğun bir araştırma alanı olmuştur. Son zamanlarda, YZ alanı, akıl yürütme, karar verme, filtreleme, yazılım kişiselleştirme ve uyarlanabilirlik, doğal dil anlama, derin öğrenme, makine öğrenimi ve veri madenciliği alanlarında kritik işlevsellik ve önemli iyileştirmeler sağlamak için yeniden ele alınmaktadır. Bu geniş kapsamlı faaliyet, büyük veri, akıllı telefon ve mobil yazılım uygulamaları, Nesnelerin İnterneti ve her türlü insan faaliyeti ve mesleğinde geniş bir uygulama yelpazesindeki önemli ve gelişen yeni teknolojik gelişmelerin araştırma ve geliştirme baskısı altındadır. Önceki gelişmeler ışığında, YZ kullanımı, Uzman Sistemler gibi ilgili teknolojilerin ve Bilgi Mühendisliği, Bilgi Edinimi veya Bilgi Temsili gibi daha geleneksel prosedürlerin ötesine geçmiş ve diğer yazılım biçimlerine de entegre edilmesi gerekmektedir. Bu nedenle, mevcut ilgili araştırmalar, YZ iyileştirmelerinin yazılıma verimli bir şekilde dahil edilmesine ve yazılımların YZ ile güçlendirilmesine yönelik çabaları içermektedir. Amaç iki yönlüdür: (i) yazılımın kendi başına veya bir denetleyici/eğitmenin yardımıyla öğrenmesini ve gelişmesini (yani, belirli görevleri yerine getirmede daha iyi, daha kullanıcı dostu ve daha verimli hale gelmesini) sağlayan algoritmalar, mekanizmalar, metodolojiler ve prosedürler geliştirmek ve (ii) YZ'nin (en azından kısmen) yazılım geliştirmeyi otomatikleştirmek için kullanımı da dahil olmak üzere tüm Yazılım Mühendisliği sürecini iyileştirmek ve bilgi mühendisliği ve bilgi ediniminin, prototiplemenin ve akıllı yazılım modüllerinin hızlı uygulama geliştirmesinin dahil edilmesini yansıtmak (Maria Virvou et all).

2022 yılında YZ üreticisi OpenAI'ın ChatGPT ürününü duyurmasıyla birlikte YZ'nin hayatımızdaki rolü hızlanmaya başladı ve günümüzde her alana bu etkinin yayıldığını görebiliyoruz. Birçok YZ ürünü bulunuyor ve bu ürünlerin kullanım amacı birbirinden farklı olacak şekilde sunulmaya devam etmektedir. İlgili YZ araçlarını sunan firmalar ise YZ mühendisleri sayesinde ürünü doğru standartlarda sunarak, hedef kitleye sorunsuz ürünü oluşturmak adına çalışmalar yapmaktadır. (Amazon)

YZ mühendislerinin temel amacı, makineleri ve bilgisayar programlarını akıllı hale getirmektir. Onlar bu programlara insan gibi düşünmeyi, karar vermeyi öğretirler. İnsan Zekâsını çoklu yönlerde taklit edebilecek çıktıları sunmayı hedefleyen YZ mühendisleri, insanlar gibi düşünüp karar verebilen bilgisayar programları üreten kişilerdir. YZ mühendisleri bilgisayarlara konuşmayı, resimleri anlamayı, problem çözmeyi öğreterek bir alt küme çerçevesinde doğru çıktıyı sunmayı amaçlarlar.

Mesela bir YZ mühendisi, ev işlerinde yardımcı olacak bir robot programı yapabilir. Bu robota bulaşıkları yıkamayı, ütü yapmayı, yerleri süpürmeyi öğretebilir. Böylece robot akıllı bir şekilde ev işlerini yapabilir. Ya da YZ mühendisi, telefondaki sesli asistanı programlayarak biz soru sorduğumuzda cevap verebilmesi için sesli asistana konuşmayı öğretir.

Bir başka örnek verecek olursak, bir robot köpek aldınız ve ona 'otur' komutu veriyorsunuz. Ama robot köpek sadece havlıyor, oturmuyor. Çünkü ona oturmayı kimse öğretmemiş. İşte YZ mühendisi, robot köpeğe kamera ve sensörler yardımıyla otur komutunu algılamayı ve buna uygun hareket etmeyi programlayabilir. Böylece robot artık aldığı komutu öğrendiği için uygulayabilir.

Teknoloji alanından bir örnek verirsek, cep telefonlarımızın kameraları her geçen gün gelişmeye devam ediyor. Bu kameralarda gece modu var ve bu sayede karanlıkta fotoğraf çekebiliyoruz. Peki telefon kamerası gece modunu nasıl algılamakta? İşte burada YZ devreye giriyor. YZ mühendisleri, telefon kamerasına gece fotoğraflarını iyileştirebilecek özel bir programlama ile sağlıyorlar. Bu programa milyonlarca gece ve gündüz fotoğrafını örnek gösteriyorlar. Böylece telefon, karanlık bir ortamda fotoğraf çekildiğini anlayınca gece programını çalıştırıyor ve fotoğrafı sanki gündüz çekilmiş gibi aydınlatıyor. Bu sayede telefonumuz, gözümüz gibi değilse bile beyin gibi düşünüp fotoğrafı otomatik olarak iyileştirebiliyor.

Özetle YZ mühendisleri bilgisayarda kodlama yaparak ve dataları kullanarak, bu programları geliştirerek günlük hayatımızı kolaylaştırıyorlar. Yani YZ mühendisliği, makineleri insan gibi düşünebilir hale getirme işidir diyebiliriz.

Her YZ anahtar kelimesi aşağıdaki YZ alanlarından biriyle manuel olarak ilişkilendirilir [S. Russell]:

- Makine öğrenimi ve veri madenciliği (MLDM), gizli ilişkileri bulmak ve tahminlerde bulunmak için geniş veri koleksiyonlarını analiz eden teknikleri kapsar.

- Bilgi temsili ve akıl yürütme (KRR), gerçek dünya kavramlarını somutlaştırma biçimlerini ve yeni bilginin nasıl çıkarılacağını veya nedensel olguların nasıl açıklanacağını ifade eder.
- Belirli bir amaca göre en uygun çözümleri bulan algoritmaları derleyen problem çözme için arama ve optimizasyon (SO).
- İletişim ve algı (CP), yani dil işleme ve görme gibi insan yeteneklerini anlama ve simüle etme.

5.2.1. Proje Planlama Süreci

Proje planlama sürecinin amacı, etkili ve uygulanabilir planlar oluşturmak ve bunları koordine etmektir. Bu süreç, projenin tamamı boyunca devam eden ve planların düzenli olarak gözden geçirilmesini içeren sürekli bir süreçtir.

- Proje yönetimi ve teknik faaliyetlerin kapsamını belirler,
- Süreç çıktıları, görevler ve teslimatlar tanımlar,
- Görevlerin yürütülmesine yönelik takvimleri, başarı ölçütlerini içerecek şekilde oluşturur,
- Görevlerin yerine getirilmesi için gereken kaynakları tahmin eder.

YZ destekli yazılım geliştirmede proje planlama süreci, klasik yazılım projelerinden farklı olarak hem yazılım mühendisliği disiplinini hem de veri odaklı YZ süreçlerini kapsamalıdır. Çünkü YZ tabanlı sistemlerde sadece kod geliştirmek değil; veri toplama, model seçimi, eğitim, doğrulama ve sürekli güncelleme de projenin kritik parçalarıdır.

Aşağıda YZ destekli yazılım geliştirmede genel bir proje planlama süreci adımları verilmiştir:

1. ***Gereksinim Analizi***
    - Klasik gereksinimler: Kullanıcı ihtiyaçları, sistem hedefleri, fonksiyonel ve fonksiyonel olmayan gereksinimler.
    - YZ'ye özgü gereksinimler:
    - Çözülecek problemin gerçekten YZ ile çözülebilir olup olmadığı.
    - Kullanılacak veri kaynaklarının tanımlanması (etik, güvenilir, güncel).
    - Modelin doğruluk, hız, bellek kullanımı gibi kalite kriterleri.
    - YZ'nin kararlarının açıklanabilirlik ve güvenilirlik düzeyi.

2. ***Proje Kapsamı ve Yol Haritası***
    - YZ modülü ile klasik yazılım bileşenlerinin sınırlarını belirleme.
    - Örnek: Sistem mimarisinde hangi kısımlar geleneksel yazılım, hangi kısımlar YZ modeliyle desteklenecek?
    - İteratif (çevik) planlama: YZ projelerinde belirsizlik yüksek olduğu için Agile/Scrum yaklaşımı daha uygundur.
3. ***Veri Planlaması***
    - Veri toplama: Gerekli veri miktarı, çeşitliliği ve kaynağı (kamu verisi, sensörler, kullanıcı girdileri).
    - Veri işleme: Temizleme, etiketleme, normalizasyon, artırma (augmentation).
    - Veri güvenliği ve gizlilik: KVKK, GDPR gibi regülasyonlara uygunluk.
4. ***Teknoloji ve Model Seçimi***
    - Kullanılacak YZ türü: Makine öğrenmesi, derin öğrenme, doğal dil işleme, bilgisayarla görü vb.
    - Geliştirme ortamı: TensorFlow, PyTorch, Scikit-learn, Hugging Face gibi çerçeve (framework) sistemleri.
    - Donanım ihtiyaçları: GPU/TPU, bulut servisleri.
5. ***Zaman ve Kaynak Planlaması***
    - Sprint bazlı takvim: Önce prototip, sonra model optimizasyonu, ardından entegrasyon.
    - Veri toplama → Model eğitimi → Test → Yazılım entegrasyonu → Kullanıcı testleri şeklinde fazlara bölme.
    - İnsan kaynağı: Yazılım geliştiriciler, veri mühendisleri, YZ mühendisleri, test uzmanları (Maria Virvou).

5.2.2. Risk yönetimi

ISO/IEC 23894 standardına uygun tanımlanan risk yönetimi hedefleri şunlardır: Adil olma (fairness), Mahremiyet (privacy), Güvenilirlik (reliability), Şeffaflık (transparency), Açıklanabilirlik (explainability), Hesap verebilirlik (accountability), Erişilebilirlik (availability), Bütünlük (integrity) ve Sürdürülebilirlik / Bakım yapılabilirlik (maintainability).

YZ Risk Yönetiminin İlkeleri Risk yönetimi, bir organizasyonun ihtiyaçlarına entegre, yapılandırılmış ve kapsamlı bir yaklaşım ile yanıt vermelidir. Temel ilkeler, bir organizasyonun

öncelikleri belirlemesine ve belirsizliklerin hedefleri üzerindeki etkilerini nasıl ele alacağına karar vermesine olanak tanır. Bu ilkeler, stratejik veya operasyonel olsun, tüm organizasyon seviyeleri ve hedefleri için geçerlidir. Sistemler ve süreçler genellikle farklı ortamlar için belirli kullanım durumlarında çeşitli teknolojiler ve işlevlerin bir kombinasyonunu kullanır. Risk yönetimi, tüm teknolojileri ve işlevsellikleri ile birlikte genel sistemi ve çevre ile paydaşlar üzerindeki etkileri dikkate almalıdır. YZ sistemleri, bir organizasyonun hedefleri üzerinde olumlu veya olumsuz etkileri olan yeni riskler ortaya çıkarabilir veya mevcut risklerin meydana gelme olasılığını değiştirebilir. Ayrıca, organizasyon tarafından özel bir dikkate alınmayı gerektirebilirler. Bu belgede, bir organizasyonun uygulayabileceği risk yönetimi ilkeleri, çerçeveleri ve süreçleri hakkında ek yönergeler yer almaktadır.

Risk yönetimi çerçevesinin amacı, organizasyonun risk yönetimini önemli faaliyetlerine ve işlevlerine entegre etmesine yardımcı olmaktır. Rehberlik ISO 31000:2018, 5.1'de belirtilmiştir. Risk yönetimi, bir organizasyon için ilgili bilgilerin toplanmasını kapsar, böylece kararlar alınabilir ve risklerle başa çıkılabilir. Yönetim kurulu, toplam risk iştahını ve organizasyonel hedefleri belirler, ancak risklerin tanımlanması, değerlendirilmesi ve yönetimi için karar verme sürecini organizasyon içindeki idare işlevlerine devreder. ISO/IEC 38507 [1] YZ sistemlerinin geliştirilmesi, satın alınması veya kullanılması hakkındaki organizasyon yönetiminin ek dikkate alımlarını tanımlar. Bu dikkate alımlar arasında yeni fırsatlar, potansiyel risk iştahı değişiklikleri ve organizasyon tarafından YZ'nin sorumlu kullanımını sağlamak için yeni yönetim ilkeleri bulunmaktadır. Bu, ISO 31000:2018, 5.2'de tanımlanan dinamik ve aşamalı olarak tekrar eden organizasyonel entegrasyonu desteklemek için bu belgede tanımlanan risk yönetimi süreçleri ile birlikte kullanılabilir.

ISO/IEC 23894'e göre Risk değerlendirme faaliyetleri, YZ sistemleriyle ilişkili tüm riskleri içermeli ve bu riskleri ele almak için uygun önlemlerin yer aldığı bir risk azaltma planı ve ilgili risk yönetimi kayıtları oluşturulmalıdır. ISO/IEC 23894, YZ kullanan ürün, sistem ve hizmetleri geliştiren, üreten, devreye alan ve kullanan kuruluşlara yönelik risk yönetimi rehberliği sağlar. Ancak bu standart, Emniyet (Safety) ve Güvenlik (Security) amaçlı olarak YZ kullanan ürün ve hizmetlere özel risk yönetimine yönelik değildir. Bu nedenle, emniyet ve güvenlik hedefleriyle YZ kullanan kuruluşlar, YZ'ye özgü süreçlerin yanı sıra, uygun uluslararası risk yönetimi standartlarını da dikkate almalıdır.

Bireylerin teknoloji ile nasıl etkileşimde bulunduğunu ve buna nasıl yanıt verdiğini etkileyebileceği beklentisini göz önünde bulundurulduğunda, YZ sistemlerinin

geliştirilmesinde yer alan organizasyonların, geliştirdikleri YZ sistemlerinin daha fazla kullanımına yönelik ilgili mevcut bilgileri takip etmeleri önerilir; ayrıca YZ sistemlerinin kullanıcıları, bu sistemlerin ömrü boyunca kullanımları hakkında kayıt tutabilirler. YZ'nin sürekli gelişen yeni bir teknoloji olması nedeniyle, tarihsel bilgiler sadece sınırlı ölçüde mevcut olabilir ve geleceğe yönelik beklentiler hızla değişebilir. Organizasyonlar bunu dikkate almalıdır. Eğer ele alınacaksa, YZ sistemlerinin dahili kullanımının düşünülmesi gerekir. Müşteriler ve dış kullanıcılar tarafından YZ sistemlerinin kullanımının izlenmesi, fikri mülkiyet hakları ve sözleşme veya piyasa spesifik kısıtlamalarla sınırlı olabilir. Bu tür kısıtlamalar, YZ risk yönetim süreci aracılığıyla belgelenmeli ve iş koşulları yeniden değerlendirmeyi gerektirdiğinde güncellenmelidir.

En ileri düzeydeki sürekli risk yönetimi türü, bir YZ sisteminin, bir dünya modeli ve kurallar temelinde muhakeme yoluyla aktif olarak karar risk analizi gerçekleştirmesidir. Otonom risk yönetimine ek olarak, zarar riskleri şu yollarla da azaltılabilir: —Zararlı kararların riskli durumlarda alınmadığını doğrulamak için yeterli test senaryosu kapsamı —Makine öğrenimi durumunda, bu riskli durumların ve doğru karar örneklerinin eğitim verisine dâhil edilmesi, —Otomatik risk yönetimi süreçlerinin insan tarafından sonradan değerlendirilmesi.

5.2.3. Sonuç ve Değerlendirmeler

YZ'nin, yazılım mühendislerinin yazılım süreci yaşam döngüsündeki görevlerle başa çıkma biçimini değiştirmesi bekleniyor. Son zamanlarda, YZ'nin farklı dallarına dayalı önerilerde bir artış yaşandı ve bu da yeni araştırma alanlarının ortaya çıkmasına ve endüstriyel uygulamalara yeni olanaklar açılmasına yol açtı. Örneğin, arama tabanlı yazılım mühendisliği, yazılım geliştirme sürecindeki karmaşık görevlere daha iyi veya optimize edilmiş çözümler aramak için algoritmalar sağlar. Ayrıca, yazılım analitiği, kod depolarından, geliştirici forumlarından veya teknik dokümanlardan bilgi edinmek için makine öğrenimi ve madencilik tekniklerini kullanır. Doğal dil işleme ve otomatik akıl yürütme, yazılım yapıtlarından ve dokümanlardan bilgi modelleme ve çıkarım yapmaya yarayan tekniklerin kullanıldığı diğer YZ alanlarıdır. Kısacası, profesyonellerin artık daha fazla kaynağı var, ancak yine de YZ yöntemlerinin etkili kullanımı için kapsamını ve sorunlarına nasıl uyduğunu anlamaları gerekiyor. (Maria Virvou et all).

Yeni YZ yöntemleri sürekli olarak ortaya çıksa ve aralarındaki engeller belirsizleşse de, yöntemlerin çoğunun kökleri hâlâ Russell ve Norvig'in "Artificial Intelligence: A Modern Approach, 4th edn." kitabında sunulan köklü alanlara dayanmaktadır: makine öğrenimi ve veri

madenciliği, bilgi temsili ve akıl yürütme, arama ve optimizasyon, iletişim ve algı. Ancak, yazılım mühendisleri bu tekniklere aşina olmayabilir ve bu da öngörücü modellerin veya meta-sezgilerin nasıl etkili bir şekilde uygulanacağını bilmeyi zorlaştırabilir [9]. Ayrıca, uygulayıcıların YZ teknolojilerinin yeteneklerini ve sınırlamalarını anlamaları gerekir [34]. Bu nedenle, doğru seçimi yapmak, net hedeflere sahip olmayı ve her tekniğin dayattığı temel varsayımları tam olarak anlamayı gerektirir. Olası etki ve riskler göz önüne alındığında, YZ'nin uygulanması gereken nokta ve süreç ve sonuçlar içindeki gerçek sorumluluğu da dikkate alınmalıdır (R. Feldt, et all).

5.2.4. Kaynaklar:

Maria Virvou et all, Handbook on Artificial Intelligence-Empowered Applied Software Engineering, VOL.1: Novel Methodologies to Engineering Smart Software Systems, ISBN 978-3-031-08201-6, Springer 2022,

C. Ebert, J. Heidrich, S. Martinez-Fernandez, A. Trendowicz, Data science: technologies for better software. IEEE Softw. 36(06), 66–72 (2019)

R. Feldt, F. Gomes de Oliveira Neto, R. Torkar, Ways of applying artificial intelligence in software engineering, in Proceedings of 6th International Workshop on Realizing Artificial Intelligence Synergies in Software Engineering (RAISE) (2018), pages 35–41

Amazon, https://aws.amazon.com/tr/what-is/machine-learning/

S. Russell, P. Norvig, Artificial Intelligence: A Modern Approach, 4th edn.(Pearson, 2020)

Maria Virvou, Artificial Intelligence-Enhanced Software and Systems Engineering, Vol 2. (Springer 2022)

5.3. YZ Güvenliği ve Ulusal Standartlar: Uluslararası Çerçeveler Işığında Güvenilir Bir YZ Platformu İçin Değerlendirme

Sinan Silim

Şans Girişim Ortak Girişimi & CEO, Thalvera Cyber Security, silim.sinan@gmail.com

---

Yapay Zekâ teknolojilerinin savunma, sağlık, finans ve ulaştırma gibi kritik sektörlerde artan entegrasyonu, ulusal altyapılar ve kamu güvenliği açısından stratejik bir güvenlik meselesi haline gelmiştir. Bu çalışma, bir ülkenin YZ standartlarını oluşturması amacıyla, güvenilir bir YZ platformunun temel güvenlik gerekliliklerini ve uluslararası standartlar bağlamında benimsemesi gereken stratejileri detaylandırmaktadır. Analizler, YZ güvenliğinin, yalnızca geleneksel siber güvenlik önlemlerinden ibaret olmadığını, aksine sistemin tüm yaşam döngüsüne yayılan çok boyutlu ve proaktif bir yaklaşım gerektirdiğini göstermektedir (Farnham, 2025).

Uluslararası arenada, YZ güvenliğine yönelik üç ana yaklaşım öne çıkmaktadır: ABD merkezli **NIST Yapay Zekâ Risk Yönetimi Çerçevesi (AI RMF)**, Avrupa Birliği'nin **Yapay Zekâ Yasası (EU AI Act.)** ve küresel bir yönetim sistemi standardı olan **ISO/IEC 42001** (Şahin Ardıyok, 2023). Bu çerçeveler, sırasıyla gönüllü bir kılavuz, bağlayıcı bir regülasyon ve pratik bir yönetim sistemi sunarak farklı felsefeleri yansıtmaktadır. Rapor, bu yaklaşımların güçlü yönlerini bir araya getiren hibrit bir ulusal strateji önermektedir. Bu strateji, yüksek riskli YZ sistemleri için bağlayıcı regülasyonlar getirmeyi, diğer sistemler için ise inovasyonu destekleyen gönüllü standartları teşvik etmeyi hedeflemektedir (*NIST YZ Risk Management Framework: A Simple Guide to Smarter YZ Governance*, 2025b).

Güvenilir bir YZ platformu, veri zehirlenmesi, düşmanca saldırılar, model hırsızlığı ve istem enjeksiyonu gibi YZ'ye özgü tehditlere karşı dayanıklı olmalıdır. Bu tehditlerle mücadele için, platformun tüm yaşam döngüsü boyunca (veri toplama, eğitim, dağıtım ve operasyon) "tasarımla güvenlik" (secure-by-design) prensibi benimsenmeli ve kapsamlı teknik ile yönetsel tedbirler uygulanmalıdır. Bu tedbirler arasında hassas verilerin şifrelenmesi, güvenilir veri kaynaklarının kullanılması, düşmanca eğitim (adversarial training) tekniklerinin uygulanması ve sürekli izleme mekanizmalarının entegrasyonu yer almaktadır (Microsoft, 2025) (NCSC, 2023).

Sonuç olarak, ulusal bir YZ güvenlik standardının amacı, inovasyonu kısıtlamak değil, aksine kamu ve özel sektörde YZ'ye olan güveni artırarak bu teknolojinin sorumlu ve sürdürülebilir gelişimini sağlamaktır. Çalışma, YZ yönetişiminin, teknolojinin kendisinden ziyade, onu nasıl yöneteceğimize ve kullanacağımıza dair stratejik bir mesele olduğuna dikkat çekmektedir.

5.3.1. Ulusal Bir Güvenlik Stratejisinin Gerekliliği

YZ teknolojileri, modern toplumun ve ekonominin temelini oluşturan kritik altyapılardan (enerji ve su sistemleri gibi) finansal hizmetlere, sağlık uygulamalarından ulaştırmaya kadar birçok alanda giderek daha merkezi bir rol oynamaktadır. YZ sistemlerinin bu denli yaygınlaşması, ulusal düzeyde bir güvenlik stratejisinin oluşturulmasını zaruri kılmaktadır. Zira YZ sistemlerinin güvenliği, yalnızca siber güvenlik uzmanlarının teknik bir sorunu olmaktan çıkmış, ulusal güvenlik, ekonomik istikrar ve toplumsal güven açısından hayati bir öneme sahip olmuştur (Şahin Ardıyok, 2023).

YZ'nin siber güvenlik alanındaki rolü, ikili bir paradoks sunmaktadır: YZ, bir yandan siber saldırıları makine hızında tespit edip önleyebilen güçlü bir savunma aracı olarak öne çıkarken, diğer yandan kötü niyetli aktörler tarafından siber saldırıların otomasyonu ve manipülasyonu için de kullanılabilmektedir. Geleneksel siber güvenlik yöntemleri (virüsler, Truva atları, DDoS saldırıları, IP sahtekarlığı ve yemleme gibi) YZ ile otomasyon kazandıkça daha gelişmiş ve karmaşık hale gelmektedir. Bu durum, savunma mekanizmalarının da insan müdahalesini aşan bir hızda çalışması gerekliliğini ortaya koymaktadır. *Darktrace* ve *IBM Watson* gibi platformlar, anormal davranışları tespit ederek ve tehditlere gerçek zamanlı müdahale ederek YZ'nin proaktif bir güvenlik aracı olarak nasıl kullanılabileceğini göstermektedir. Dolayısıyla, bir ülkenin ulusal YZ stratejisi, sadece bu teknolojinin güvenlik açıklarına karşı korunmasını değil, aynı zamanda YZ'nin kendisinin ulusal siber savunma kapasitesini nasıl artırabileceğini de düşünmesini gerektirmektedir (Jamil, 2024) (Arsenault, 2024).

Bu çalışma, bir YZ platformunun uluslararası standartlar ışığında güvenilir hale getirilmesi için sahip olması gereken güvenlik standartları ve önlemlerine derinlemesine bir bakış sunmaktadır. Bu rapor, ulusal politika yapıcıların, YZ'nin sunduğu potansiyel faydaları en üst düzeye çıkarırken, beraberinde getirdiği riskleri de etkin bir şekilde yönetmelerine yardımcı olacak kapsamlı bir yol haritası sağlamayı amaçlamaktadır.

5.3.2. Yapay Zekâ Güvenliğine Yönelik Uluslararası Standartlar ve Düzenleyici Çerçeveler

YZ güvenliği konusunda küresel düzeyde ortak bir standart bulunmamakla birlikte, farklı coğrafyalar ve kurumlar tarafından benimsenen çeşitli çerçeveler ve regülasyonlar mevcuttur. Bu yaklaşımlar, gönüllü kılavuzlardan bağlayıcı yasalara kadar geniş bir yelpazede yer almaktadır ve her biri farklı bir felsefeye dayanmaktadır. Ulusal bir standart oluştururken, bu mevcut çerçevelerin incelenmesi ve güçlü yönlerinin harmanlanması stratejik bir gerekliliktir.

1. NIST Yapay Zekâ Risk Yönetimi Çerçevesi (AI RMF)

Ulusal Standartlar ve Teknoloji Enstitüsü (NIST) tarafından geliştirilen YZ Risk Yönetimi Çerçevesi (AI RMF), ABD'de ve uluslararası alanda geniş kabul gören, gönüllü bir rehberdir. Temel amacı, inovasyonu engellemeden kuruluşların YZ sistemlerine özgü riskleri, sistemin tüm yaşam döngüsü boyunca yönetmelerine yardımcı olmaktır. Çerçeve, YZ risklerini belirleme, değerlendirme ve azaltma konusunda yapılandırılmış bir yaklaşım sunar ve "Güvenilir YZ" (Reliable AI) kavramı üzerine inşa edilmiştir. Bu kavram, yalnızca güvenliği değil, aynı zamanda şeffaflık, adalet, hesap verebilirlik ve gizlilik gibi ilkeleri de kapsamaktadır (Farnham, 2025).

AI RMF'nin kalbinde, YZ risk yönetimini organize eden dört ana işlev bulunmaktadır:

- **Yönetişim (Govern):** YZ risklerini yönetmek için yönetim yapılarını tanımlamayı, rolleri ve sorumlulukları atamayı içerir.
- **Haritalama (Map):** YZ yaşam döngüsü boyunca riskleri proaktif olarak tanımlama ve değerlendirme sürecidir.
- **Ölçme (Measure):** YZ sistemlerinin performansını, etkinliğini ve risklerini nicel olarak değerlendirmeyi sağlar.
- **Yönetme (Manage):** Tespit edilen riskleri önceliklendirme ve azaltma adımlarını içerir.

NIST, kuruluşların YZ risk yönetimi yetkinliklerini değerlendirmeleri için dört olgunluk seviyesi tanımlar: Kısmi (Partial), Risk Odaklı (Risk-Informed), Tekrarlanabilir (Repeatable) ve Adaptif (Adaptive). Bu seviyeler, kuruluşlara mevcut durumlarını kıyaslamak (benchmark) ve gelecekteki iyileştirmeleri önceliklendirme olanağı sunar (Farnham, 2025).

2. Avrupa Birliği Yapay Zekâ Yasası (EU AI Act.)

Avrupa Birliği YZ Yasası, yasal olarak bağlayıcı olan ve dünya genelinde bu alandaki ilk kapsamlı düzenleme olma niteliği taşıyan bir mevzuattır. Yasanın temel amacı, AB iç pazarında insan merkezli, güvenli ve güvenilir YZ sistemlerinin geliştirilmesini ve benimsenmesini teşvik etmektir. AB Yasa'sı, YZ sistemlerini potansiyel risklerine göre sınıflandıran bir risk temelli yaklaşım benimser. Düzenlemelerin katılığı, sistemin temel haklar üzerindeki potansiyel etkisine göre artmaktadır (European Union, 2024).

Yasa, dört ana risk kategorisi belirlemiştir:

- **Kabul Edilemez Risk**: Bireylerin temel haklarına açık tehdit oluşturan sistemler yasaklanmıştır. Örnek olarak, sosyal puanlama sistemleri veya serbest iradeyi manipüle eden YZ sistemleri verilebilir (Şahin Ardıyok, 2023).
- **Yüksek Risk**: Kritik altyapı, tıp, hukuk, istihdam ve kolluk kuvvetleri gibi alanlardaki YZ sistemleri bu kategoriye girer ve katı yasal yükümlülüklere tabidir. Bu yükümlülükler; yüksek kaliteli veri setleri, risk yönetim sistemlerinin kurulması, teknik dokümantasyon, insan gözetimi ve yüksek düzeyde siber güvenlik gerekliliklerini içermektedir (Şahin Ardıyok, 2023).
- **Sınırlı Risk**: Chatbot'lar gibi belirli sistemler için kullanıcıların bir makineyle etkileşimde bulunduğunun açıkça belirtilmesi gibi şeffaflık kuralları zorunludur. Ayrıca, derin sahteler (deepfakes) gibi YZ tarafından oluşturulan içeriklerin etiketlenmesi gerekmektedir (Şahin Ardıyok, 2023).
- **Minimal Risk**: Uygulamaların büyük çoğunluğunu oluşturan bu kategoriye (örn. spam filtreleri) giren sistemler için ek yükümlülükler getirilmemiştir. Ancak şirketler, bu sistemler için gönüllü olarak ek davranış kuralları benimseyebilir (Şahin Ardıyok, 2023).

3. ISO/IEC 42001:2023 Yapay Zekâ Yönetim Sistemi

ISO/IEC 42001:2023, kuruluşların YZ sistemlerini yönetmesi için uluslararası bir yönetim sistemi standardıdır. Bu standart, YZ ile ilgili riskleri ve fırsatları sistematik bir şekilde ele alarak, YZ teknolojilerinin güvenli ve etik bir şekilde geliştirilmesi, konuşlandırılması ve kullanılması için bir çerçeve sunar. Standart, ISO'nun bilinen "planla-uygula-kontrol et-önlem al" (P-U-K-Ö) döngüsüne benzer bir yapıyı takip eder ve şu ana bölümleri içerir: Liderlik, Planlama, Destek, İşletme ve Değerlendirme. ISO/IEC 42001, YZ sistemlerinin kurumsal politikalara ve değerlere uygunluğunu sağlamaya odaklanarak, kuruluşların disiplinli bir

şekilde YZ operasyonlarını yürütmelerini desteklemektedir. Bu yönetim sistemi, YZ'yi operasyonlarına entegre eden herhangi bir kuruluş için bir yol haritası görevi görmektedir (Ekaglobal, 2023).

4. ISO/IEC 5338:2023 Yapay Zekâ Yaşam Döngüsü Süreçleri

ISO/IEC 5338:2023, YZ sistemlerinin yaşam döngüsünü tanımlayan kapsamlı bir çerçeve sunar. Bu standart, makine öğrenimi ve sezgisel sistemlere dayanan YZ sistemlerinin yaşam döngüsü süreçlerini ve ilgili kavramları tanımlar. ISO/IEC 5338, mevcut YZ süreç standartlarıyla entegre olarak, YZ sistemlerinin etkin bir şekilde tanımlanmasını, kontrol edilmesini, uygulanmasını ve iyileştirilmesini kolaylaştırır. Bu standart, YZ geliştirme süreçlerini düzene sokarak çeşitli alanlarda inovasyonu, güvenilirliği ve verimliliği teşvik etmeyi amaçlamaktadır. Bu teknik süreçler, ISO/IEC 42001 gibi yönetim sistemleriyle birlikte kullanıldığında, sorumlu YZ (Responsible AI) uygulamalarının organizasyonel, yönetimsel ve mühendislik düzeyinde nasıl hayata geçirileceği sorusuna yanıt verir.

5. ISO/IEC TR 24368:2022 Etik ve Toplumsal Kaygılar

ISO/IEC TR 24368:2022 standardı, YZ'nin etik ve toplumsal kaygılarına ilişkin üst düzey bir genel bakış sunar. Bu standart, geliştiriciler, düzenleyiciler ve ilgili tüm paydaşlar için YZ sistemlerinde şeffaflık, hesap verebilirlik ve adalet gibi ilkelerin benimsenmesine yardımcı olur. TR 24368, kuruluşların etik kaygıları sistematik olarak ele alabilmeleri için yapılandırılmış bir yaklaşım sağlar ve YZ'yi farklı toplumsal değerlerle uyumlu hale getirmeyi teşvik eder. Standart, etik risk değerlendirmelerinin YZ sisteminin tüm yaşam döngüsü boyunca, ilk tasarımdan veri toplamaya, model eğitimine, dağıtıma ve operasyonlara kadar kilit aşamalarda gerçekleştirilmesi gerektiğini belirtir. Ayrıca, hesap verebilirliğin teknik ekiplerden öte, kurumsal liderlik ve yönetişim yapılarını da kapsaması gerektiğini vurgular.

6. Küresel Politika Trendlerinin Karşılaştırmalı Analizi ve Stratejik Sonuçlar

NIST'in gönüllü, esnek ve inovasyon odaklı yaklaşımı ile AB'nin bağlayıcı ve hak merkezli yasal çerçevesi, YZ yönetişimi konusundaki iki farklı felsefenin ürünleridir. NIST, özel sektörün kendi risklerini yönetme kapasitesine güvenirken, AB, YZ'nin temel haklara potansiyel zararlarını önlemek için proaktif ve yasal bir denetim mekanizması kurmayı amaçlamaktadır. Bu farklılıklar, bir ülkenin ulusal YZ standardını oluştururken tek bir modeli kopyalamasının yetersiz kalacağını göstermektedir. Bunun yerine, bu iki yaklaşımın güçlü yönlerini harmanlayan hibrit bir model benimsenmesi daha uygun bir strateji olacaktır (Farnham, 2025).

NIST gibi bir çerçeve, YZ'nin geniş kullanım alanlarında inovasyonu teşvik etmek için özel sektöre rehberlik sunabilirken, AB Yasası'ndan ilham alınarak yüksek riskli veya kritik alanlardaki YZ sistemleri için bağlayıcı regülasyonlar getirilebilir. ABD Başkanı Joe Biden'ın 2023 tarihli bir idari emriyle, ulusal güvenlik açısından riskli YZ sistemleri geliştirenlerin güvenlik testi sonuçlarını hükümetle paylaşmasını zorunlu kılması, NIST RMF gibi gönüllü bir çerçevenin dahi zamanla stratejik sektörlerde bağlayıcı unsurlarla desteklenebileceğinin bir kanıtıdır. Bu, ulusal bir standardın, başlangıçta gönüllü bir çerçeve olarak tasarlanmasının, ileride yasal gerekliliklere dönüşebilecek esnekliğe sahip olması gerektiğini ortaya koymaktadır (Farnham, 2025) (Şahin Ardıyok, 2023) .

Çin'in devlet merkezli ve kontrollü yaklaşımı ise, veri egemenliğinin ve ulusal güvenliğin birincil öncelik olarak ele alındığı farklı bir modeli yansıtmaktadır. Bu model, YZ sistemlerinin içeriğinin "sosyalist değerlerle uyumlu" olmasını dahi zorunlu kılmaktadır. Bir ulusal YZ standardı, bu farklı yaklaşımları değerlendirerek kendi özgün ihtiyaçlarına ve değerlerine en uygun dengeyi bulmalıdır (Ye, 2023).

Aşağıdaki tablo (Tablo 1), uluslararası çerçevelerin temel özelliklerini karşılaştırmalı olarak sunmaktadır:

Tablo1: (Farnham, 2025), (Wiz Experts Team, 2025), (European Union, 2024).

Standart/ Yasa Adı	Kapsam ve Yapı	Hukuki Statüsü	Temel Felsefesi	Öne Çıkan Özellikler
NIST YZ RMF	Tüm YZ yaşam döngüsünü kapsayan dört temel işlev (Govern, Map, Measure, Manage) üzerine kurulu bir kılavuz.	Gönüllü	İnovasyonu engellemeden risk yönetimi.	"Güvenilir YZ" ilkeleri (adalet, şeffaflık), olgunluk seviyeleri.
AB Yapay Zekâ Yasası	Risk temelli bir sınıflandırma sistemiyle YZ uygulamalarını düzenler.	Bağlayıcı	İnsan merkezli, güvenli ve güvenilir YZ.	Yüksek riskli sistemlere katı gereklilikler, kabul edilemez riskler için yasaklar.
ISO/IEC 42001:2023	Kuruluşların YZ sistemlerini yönetmeleri için bir yönetim sistemi standardı.	Gönüllü	Disiplinli ve sistematik bir YZ yönetişimi.	ISO'nun P-U-K-Ö döngüsüne benzer yapı, risk ve fırsat yönetimi.
ISO/IEC 5338:2023	Yapay Zekâ sistemlerinin yaşam döngüsü süreçlerini tanımlar.	Gönüllü	Yaşam döngüsü süreçlerini standartlaştırma.	YZ sistemlerinin geliştirilmesini, kontrolünü ve iyileştirilmesini kolaylaştırır.
ISO/IEC TR 24368:2022	YZ'nin etik ve toplumsal kaygılarına genel bakış.	Gönüllü (Teknik Rapor)	Etik ve toplumsal kaygıların ele alınması.	Şeffaflık, hesap verebilirlik, adalet ilkelerini vurgular ve etik risk değerlendirmeleri için rehberlik sunar.

5.3.3. Yapay Zekâ Sistemlerine Yönelik Tehditler ve Güvenlik Açıkları

YZ sistemlerinin benzersiz yapısı, geleneksel siber güvenlik risklerinin ötesinde, bu teknolojiye özgü yeni tehditleri de beraberinde getirmektedir. Bu tehditlerin anlaşılması, etkili bir güvenlik stratejisi oluşturmanın ilk adımıdır. Bu tehditler, veri, model ve altyapı olmak üzere üç ana kategoride incelenebilir (NCSC, 2023).

1. Veri Kaynaklı Tehditler

YZ modelleri, eğitim verileri kadar iyi performans gösterirler. Bu durum, veri zehirlenmesi ve gizlilik ihlalleri gibi saldırıları YZ sistemleri için ciddi birer risk haline getirmektedir (Chen, 2023).

- **Veri Zehirlenmesi (Data Poisoning):** Bu saldırı türü, saldırganların bir YZ sisteminin eğitim setine kasıtlı olarak yanıltıcı veya kötü niyetli veriler enjekte etmesiyle meydana gelir. Yanlış verilerle eğitilen model, yanlış veya zararlı çıktılar üretebilir. Bu durum, özellikle sürekli öğrenen ve geri beslemeye dayalı sistemler için ciddi bir tehlike oluşturur (Amos, 2022).
- **Gizlilik İhlalleri:** YZ eğitim veri setleri, şirketler için ticari sır niteliği taşıyan veya bireyler için hassas kişisel veriler içerebilir. Modelin işlevselliğinin tersine mühendislikle çözülmesi (model inversiyonu) gibi saldırılar, kamuya açık olması amaçlanmayan gizli verilerin ifşa olmasına yol açabilir. Bu, veri koruma ve gizlilik mevzuatlarına uyum açısından ciddi riskler taşımaktadır (Kaminsky, 2024).

2. Model ve Algoritma Kaynaklı Tehditler

Bu tehditler, YZ modelinin kendisine veya onunla etkileşime geçme şekline odaklanır.

- **Düşmanca Saldırılar (Adversarial Attacks):** En bilinen YZ tehditlerinden biridir. Saldırganlar, bir YZ sistemine giriş verisi üzerinde, insan gözüyle fark edilemeyecek kadar küçük değişiklikler yaparak modelin yanlış sınıflandırma veya karar vermesini sağlayabilir. Örneğin, bir dur tabelasına eklenen küçük bir çıkartma, otonom bir aracın onu bir hız limiti işareti olarak algılamasına neden olabilir. Bu saldırılar, özellikle yüz tanıma sistemleri ve otonom araçlar gibi yüksek riskli uygulamalar için kritik bir zafiyet oluşturmaktadır (Microsoft, 2025).
- **Model Hırsızlığı (Model Theft):** Bir saldırganın, bir YZ modelinin işlevselliğini ve mimarisini tersine mühendislik yoluyla kopyalamasıdır. Bu saldırı, modelin fikri

mülkiyetini ve ticari sırlarını tehlikeye atar ve bir şirketin rekabet avantajını doğrudan hedef alır (Microsoft, 2025).

- **İstem Enjeksiyonu (Prompt Injection):** Özellikle büyük dil modellerinde (LLM) yaygın olan bu saldırı, kullanıcı girdisi içine gizlenmiş zararlı komutlarla modelin güvenlik önlemlerini atlamasını veya istenmeyen çıktılar üretmesini hedefler. OWASP'ın LLM'lere yönelik en kritik güvenlik açıkları listesinde yer almaktadır (Kaminsky, 2024).

3. Altyapı ve Sistem Kaynaklı Tehditler

YZ sistemleri, karmaşık bir altyapı üzerinde çalışır. Bu altyapılardaki geleneksel güvenlik açıkları, YZ'ye özgü riskleri tetikleyebilir.

- **Tedarik Zinciri Zafiyetleri**: YZ sistemleri genellikle üçüncü taraf kütüphaneler, önceden eğitilmiş modeller ve açık kaynak kodlar gibi bileşenlerden oluşan karmaşık bir tedarik zinciri kullanır. Bu zincirdeki herhangi bir zafiyet, tüm YZ platformunu istismara açık hale getirebilir (Trendmicro, 2024).
- **Hizmet Reddi (Denial of Service - DoS):** YZ hizmetini aşırı yükleyerek veya modelin yanıt vermesini engelleyerek sistemi kullanılamaz hale getirme saldırılarıdır (Trendmicro, 2024).
- **Yetki İhlalleri:** YZ kaynakları için uygun kimlik ve erişim yönetiminin uygulanmaması, saldırganların sistemlere sızmasına ve kötü amaçlı eylemlerde bulunmasına olanak tanıyabilir (Microsoft, 2025).

5.3.4. Güvenilir Bir YZ Platformu İçin Güvenlik Standartları ve Tedbirler

Güvenilir bir YZ platformu oluşturmak, yalnızca saldırılara karşı koruma sağlamakla sınırlı değildir; aynı zamanda sistemin adil, şeffaf ve hesap verebilir olmasını da gerektirir. Bu nedenle, güvenlik önlemleri, YZ sisteminin tüm yaşam döngüsü boyunca (veri toplama, model eğitimi, dağıtım ve operasyon) entegre edilmelidir (Farnham, 2025).

1. YZ Yaşam Döngüsü Boyunca Güvenlik Yaklaşımı

Geleneksel siber güvenlik, genellikle bir sistemin operasyonel aşamasına odaklanırken, güvenilir bir YZ için anahtar, tasarım aşamasından itibaren güvenliğin entegre edilmesidir (secure-by-design). Veri toplama aşamasında yapılan bir hata veya veri zehirlenmesi, modelin eğitiminde yanlılığa (bias) veya yanlış sonuçlara yol açabilir. Bu durum, sistemin operasyonel

kullanımda yanlış veya zararlı kararlar almasına neden olur. Bu nedenle, YZ yaşam döngüsünün her aşamasında alınan güvenlik önlemleri, bir sonraki aşamada ortaya çıkabilecek risklerin maliyetini büyük ölçüde azaltmaktadır (NCSC, 2023) (Trendmicro, 2024).

2. ***Veri Yönetişimi ve Hazırlığı (Yaşam Döngüsü: 1. Aşama)***

YZ sistemlerinin temeli verilerdir. Bu nedenle, veri güvenliği ve yönetişimi, YZ güvenliğinin en kritik başlangıç noktasıdır.

- **Gizlilik ve Koruma:** Hassas veriler, yetkisiz erişimi engellemek için şifreleme teknikleri kullanılarak korunmalıdır. Ayrıca, kişisel verilerin anonimleştirilmesi veya anonim verilerin tercih edilmesi, veri ihlali riskini azaltmaktadır (Microsoft, 2025). .
- **Bütünlük ve Kalite:** Eğitim için kullanılan verilerin güvenilir ve doğrulanabilir kaynaklardan geldiğinden emin olunmalıdır. Veri zehirlenmesi riskini azaltmak için verilerin düzenli olarak temizlenmesi, doğrulama setleri ve çapraz doğrulama yöntemleri gibi tekniklerin kullanılması gereklidir. Kalitesiz veya dengesiz veriler, YZ modelinde önyargılara ve yanlışlıklara yol açabilir (Microsoft, 2025).

3. ***Model Geliştirme ve Eğitim (Yaşam Döngüsü: 2. Aşama)***

Modelin kendisinin güvenilir ve dayanıklı olması, YZ güvenliğinin temelini oluşturur.

- **Güvenli Tasarım:** Güvenlik, adalet ve şeffaflık ilkeleri, modelin tasarım sürecine baştan itibaren entegre edilmelidir (SAP, 2024).
- **Düşmanca Eğitim:** YZ modelleri, düşmanca saldırılara karşı direncini artırmak için özel olarak tasarlanmış düşmanca eğitim (adversarial training) teknikleriyle eğitilmelidir. Bu, modelin yanıltıcı girdilere karşı daha sağlam olmasını sağlar (NIST, 2025) (Palo Alto Networks, 2025).
- **Kapsamlı Test ve Denetim:** Modellerin olası saldırılara karşı güvenlik açıklarını belirlemek için düzenli ve kapsamlı testler yapılmalıdır. Sistemin zafiyetlerini proaktif olarak tespit etmek için etik siber saldırganlardan (hacker) oluşan "kırmızı takım" (red team) uygulamaları kullanılabilir (Gürkan Coşkun, 2024).

4. ***Dağıtım ve Operasyonel Süreçler (Yaşam Döngüsü: 3. Aşama)***

YZ sistemi operasyonel hale geldiğinde, sürekli izleme ve yönetişim mekanizmaları devreye girmelidir.

- **Sürekli İzleme ve Güncelleme:** YZ sistemleri, anormal davranışları ve güvenlik açıklarını gerçek zamanlı olarak tespit etmek için otomatik izleme araçlarıyla denetlenmelidir. Güvenlik açıklarını kapatmak ve ortaya çıkan tehditlere karşı dayanıklılığı artırmak için modeller düzenli olarak güncellenmelidir (Microsoft, 2025).
- **Erişim Kontrolü:** Yalnızca yetkili kişilerin YZ sistemlerine ve verilerine erişimini sağlamak için rol tabanlı erişim kontrolü (RBAC) ve çok faktörlü kimlik doğrulama (MFA) gibi güçlü mekanizmalar uygulanmalıdır (Microsoft, 2025).
- **Altyapı Güvenliği:** YZ platformunun üzerinde çalıştığı altyapı, geleneksel siber güvenlik önlemleri (güvenlik duvarları, ağ saldırı önleme sistemleri) ve Sıfır Güven (Zero Trust) mimarisi ile korunmalıdır (Microsoft, 2025).
- **Denetlenebilirlik ve Hesap Verebilirlik:** Özellikle hayati tehlike arz eden veya temel hakları etkileyen uygulamalarda, YZ sistemleri bağımsız olarak denetlenebilir olmalı ve olumsuz etkileri minimize etmek için bir yedek planı bulunmalıdır (İstanbul Barosu, 2021).

Aşağıdaki tablo (Tablo 2), YZ'ye özgü başlıca tehditleri ve bunlara karşı önerilen koruma yöntemlerini özetlemektedir:

Tablo 2: (İstanbul Barosu, 2021), (Microsoft, 2025), (NIST, 2025), (Trendmicro, 2024).

Tehdit Adı	Tanımı	Etki Alanı	Önerilen Koruma Önlemleri (Teknik ve İdari)
Veri Zehirlenmesi	Eğitim verilerine kötü niyetli veri enjekte ederek modelin hatalı çıktılar vermesini sağlamak.	Model bütünlüğü, güvenilirliği.	Veri kaynaklarının doğrulanması, düzenli veri temizliği, veri setlerinin doğrulama için bölünmesi.
Düşmanca Saldırılar	Giriş verilerinde insan gözüyle algılanamayan değişiklikler yaparak modeli yanıltmak.	Model doğruluğu, sistem güvenliği.	Düşmanca eğitim (adversarial training), kapsamlı testler ve denetimler, kırmızı takım uygulamaları.
Model Hırsızlığı	Bir modelin işlevselliğini veya mimarisini tersine mühendislikle kopyalamak.	Fikri mülkiyet, ticari sırlar.	Fikri mülkiyet haklarının korunması, model parametrelerinin gizlenmesi ve şifrelenmesi, erişim kontrolü.
İstem Enjeksiyonu	Zararlı komutları kullanıcı girdisi içine gizleyerek modelin güvenlik önlemlerini atlamasını sağlamak.	Sistem bütünlüğü, çıktının güvenilirliği.	Giriş ve çıkış denetimleri, istem enjeksiyonu tespit algoritmaları, model çıktılarının sürekli izlenmesi.
Gizlilik İhlalleri	Eğitim verilerindeki hassas bilgilerin yetkisiz şekilde ifşa edilmesi.	Gizlilik, veri koruma, yasal uyumluluk.	Hassas verilerin şifrelenmesi, veri anonimleştirme teknikleri, rol tabanlı erişim kontrolü.

5.3.5. Ulusal Bir Yapay Zekâ Güvenlik Standardı İçin Stratejik Öneriler

Ulusal bir YZ güvenlik standardı oluşturmak, karmaşık ve çok yönlü bir süreçtir. Bu süreç, uluslararası en iyi uygulamaların dikkatli bir şekilde analiz edilmesini ve ülkenin kendi benzersiz koşullarına uyarlanmasını gerektirir.

1. Uluslararası Standartların Adaptasyonu İçin Hibrit Bir Model

Tek bir uluslararası standardın tamamen kopyalanması, bir ülkenin kendine özgü ihtiyaçları, mevcut yasal çerçevesi ve teknoloji ekosistemi göz önüne alındığında yetersiz kalacaktır. Bu nedenle, en pragmatik yaklaşım, farklı çerçevelerin güçlü yönlerini birleştiren hibrit bir model benimsemektir. ISO/IEC 5338 gibi yaşam döngüsü standartları, YZ'nin gelişim ve operasyonel süreçlerini düzenleyerek teknik bir temel oluşturur. ISO/IEC TR 24368 gibi etik kılavuzlar, adaletten şeffaflığa kadar geniş bir yelpazedeki toplumsal kaygıları ele alarak bu teknik temelin üzerine bir değerler çerçevesi inşa edilmesini sağlar. NIST'in esnek ve gönüllü yaklaşımı, özel sektörün inovasyonunu destekleyerek geniş ölçekte YZ'nin benimsenmesini teşvik edebilir. Aynı zamanda, AB'nin risk sınıflandırması, finans, sağlık, hukuk ve savunma gibi kritik alanlardaki YZ sistemleri için zorunlu minimum güvenlik ve şeffaflık gerekliliklerini belirleyebilir. Bu strateji, temel hakları korurken, YZ ekosisteminin genelinde inovasyon için alan bırakacaktır (Farnham, 2025).

ISO/IEC 42001 ise, bu standartların kurumsal düzeyde nasıl uygulanacağına dair pratik bir yönetim sistemi sunarak süreci kolaylaştırabilir.

2. Uygulama ve Yönetişim Mekanizmaları

Etkin bir standardın oluşturulması, sağlam bir yönetişim yapısı ve uygulama mekanizmalarına bağlıdır.

- **Sorumlu YZ İçin Yönetişim Yapısı:** YZ'nin çok disiplinli doğası gereği (veri bilimi, hukuk, etik, siber güvenlik), tek bir birimin yeterli olması beklenemez. Bu nedenle, farklı bakanlıkları, düzenleyici kurumları, akademiyi ve özel sektörü bir araya getiren bir "YZ Görev Gücü" veya "YZ Yönetişim Kurulu" oluşturulması önerilmektedir. Bu yapı, YZ stratejisinin sürekli olarak güncellenmesini ve ilgili tüm paydaşların katılımıyla yürütülmesini sağlayacaktır (SAP, 2025) (Para Dergi, 2024).

- **Belgelendirme ve Denetim:** Özellikle yüksek riskli olarak sınıflandırılan YZ sistemleri için bağımsız, üçüncü taraf denetimlerini veya ulusal bir belgelendirme mekanizmasını zorunlu kılmak, bu sistemlerin güvenilirliğini ve denetlenebilirliğini artıracaktır. Bu, aynı zamanda kamuoyunun YZ'ye olan güvenini pekiştirecektir (İstanbul Barosu, 2021).

3. Kapasite Geliştirme ve İş birliği

Küresel siber güvenlik sektöründeki yaklaşık 4 milyon kişilik uzman açığı göz önüne alındığında, YZ güvenliği konusunda yetkinlik geliştirmek kritik bir öneme sahiptir.

- **Ulusal Yetkinlik Geliştirme:** YZ güvenliği alanında uzman yetiştirmek için kamu, akademi ve özel sektör işbirliğine yatırım yapılmalıdır. Üniversite programları, sertifika eğitimleri ve araştırmalar teşvik edilmelidir (Microsoft, 2025).

- **Farkındalık ve Eğitim:** YZ sistemlerini geliştiren, kullanan ve yöneten tüm paydaşlar için YZ riskleri ve güvenilir kullanım yöntemleri hakkında sürekli eğitim programları düzenlenmelidir. Çalışanların farkındalığının artırılması, insan hatası kaynaklı güvenlik açıklarını minimize etmenin en etkili yollarından biridir (SAP, 2025) (Trendmicro, 2024).

5.3.6. ISO/IEC 5338 YZ Yaşam Döngüsünde Kritik Siber Güvenlik Aşamaları

YZ sistemleri, kritik altyapılardan tüketici hizmetlerine kadar geniş bir yelpazede kullanılmakta ve giderek artan ölçüde güvenlik tehditlerine maruz kalmaktadır. ISO/IEC 5338 standardı, YZ mühendisliği için yaşam döngüsü odaklı bir yaklaşım sunmaktadır. Bu çalışmada, yaşam döngüsünün dört aşaması — Gereksinimler & Tasarım, Veri Yaşam Döngüsü, Entegrasyon & Dağıtım, İşletim & Bakım — siber güvenlik açısından "çok kritik" kategorisinde incelenmiştir. Her aşama için güvenlik riskleri, odak alanları ve alınabilecek önlemler detaylandırılmıştır.

YZ sistemlerinin güvenliği yalnızca modelin doğruluğu veya performansı ile değil, aynı zamanda yaşam döngüsü boyunca maruz kaldığı tehditlerin yönetimiyle de doğrudan ilişkilidir. ISO/IEC 5338, YZ mühendisliği yaşam döngüsünü tanımlarken güvenlik ve risk yönetimini her aşamaya entegre etmektedir. Bununla birlikte, bazı aşamalar siber güvenlik açısından daha kritik bir konumda yer almaktadır.

Bu bölüm, ISO/IEC 5338 kapsamında öne çıkan dört kritik aşamayı (2, 3, 6 ve 7) ele almakta, neden kritik olduklarını, hangi güvenlik odaklarına sahip olmaları gerektiğini ve hangi önlemlerin uygulanabileceğini tartışmaktadır.

1. Gereksinimler ve Tasarım (Requirements and Design) Aşaması

1.1 Kritikliğin Nedeni

- Yanlış veya eksik güvenlik gereksinimleri, daha sonraki aşamalarda kapatılamayan mimari zafiyetlere yol açar.
- Erişim kontrolü, loglama ve kimlik doğrulama eksiklikleri, sistem bütünlüğünü tehlikeye atar.

1.2 Güvenlik Odakları

- Secure-by-Design prensipleri
- Kimlik ve erişim yönetimi (IAM)
- API güvenliği ve şifreleme
- İnsan faktörleri (ISO 9241-210)

1.3 Önlemler ve Yol Gösterici Standartlar

- Tehdit modelleme metodolojilerinin (örn. STRIDE, MITRE ATLAS) erken aşamada uygulanması
- Zero Trust mimarisi benimsenmesi
- Güvenlik gereksinimlerinin ISO/IEC 27001 ve ISO/IEC 27005 çerçeveleriyle uyumlu tanımlanması

2. Veri Yaşam Döngüsü (Data Life Cycle) Aşaması

2.1 Kritikliğin Nedeni

- Veri manipülasyonu (data poisoning, adversarial örnekler) modelin güvenilirliğini doğrudan zedeler.
- Gizlilik ihlalleri, yasal ve etik uyumsuzluklara yol açar.

2.2 Güvenlik Odakları

- Veri kaynağı doğrulama ve bütünlük kontrolleri
- Hassas verilerin maskeleme ve anonimleştirme yöntemleri
- Adversarial veri temizliği ve kalite kontrolü

2.3 Önlemler ve Yol Gösterici Standartlar

- *Hash* tabanlı bütünlük doğrulama mekanizmaları
- *Differential privacy* yöntemleri ile hassas veri koruması
- Eğitim verilerinde *adversarial* örnek tespiti
- ISO/IEC 27701 ve ISO/IEC 27018 entegrasyonu

3. Entegrasyon ve Dağıtım (Integration and Deployment) Aşaması

3.1 Kritikliğin Nedeni

- Üretime alınan modeller saldırganlar için en görünür yüzeydir.
- API güvenlik açıkları, *container* yanlış yapılandırmaları ve CI/CD zafiyetleri *supply chain* saldırılarına kapı aralar.

3.2 Güvenlik Odakları

- Container güvenliği ve imaj doğrulama
- API rate *limiting*, erişim kontrolü ve WAF
- *DevSecOps* süreç entegrasyonu
- Sürüm onayı ve *sandboxing*

3.3 Önlemler ve Yol Gösterici Standartlar

- CI/CD *pipeline'nına* güvenlik testleri (SAST, DAST) entegre etmek
- Kubernetes ve container imajlarını SBOM doğrulamalı ve imzalı kullanmak
- Dağıtım öncesi güvenlik denetimleri ve kırmızı takım (Red Teaming) testleri yapmak

4. İşletim ve Bakım (Operation and Maintenance) Aşaması

4.1 Kritikliğin Nedeni

- Canlı ortamda saldırılar çoğunlukla bu aşamada gerçekleşir.
- Yetersiz izleme, saldırıların uzun süre fark edilmemesine yol açar.
- Loglama eksiklikleri adli analizde (forensics) büyük zafiyet yaratır.

4.2 Güvenlik Odakları

- *Anomaly detection* sistemleri
- Güvenlik bilgi ve olay yönetimi (SIEM)
- Sürekli loglama ve olay müdahalesi (IR)
- Model güncellemeleri ve yama yönetimi

4.3 Önlemler ve Yol Gösterici Standartlar

- SOC ile *pipeline* loglarının entegre takibi
- SOAR tabanlı otomatik müdahale mekanizmaları
- API kullanımında kötüye kullanım tespit kuralları
- ISO/IEC 27035 uyumlu olay müdahale planları

5. Siber Güvenlik Bakış Açısı ile Değerlendirme

YZ sistemlerinin güvenliği, yaşam döngüsünün her aşamasında ele alınmalıdır. Bununla birlikte, ISO/IEC 5338 kapsamında Gereksinimler & Tasarım, Veri Yaşam Döngüsü, Entegrasyon & Dağıtım, İşletim & Bakım aşamaları siber güvenlik açısından en kritik noktaları oluşturmaktadır.

Bu aşamalarda secure-by-design yaklaşımı, adversarial veri koruması, güvenli dağıtım ve sürekli izleme ön planda tutulmalıdır. Böylece hem teknik hem de regülasyonel düzeyde güvenli, güvenilir ve sürdürülebilir YZ çözümleri hayata geçirilebilir.

5.3.7. Sonuç ve Geleceğe Yönelik Bakış

YZ güvenliği, tek seferlik bir proje veya durağan bir hedef değildir; sürekli evrilen tehditler ve teknolojiler karşısında sürekli adaptasyon ve iyileştirme gerektiren dinamik bir süreçtir. Ulusal bir YZ standardı, teknolojinin gelişimini kısıtlamaktan ziyade, onu sorumlu ve güvenli bir şekilde yönlendirerek toplumsal ve ekonomik faydaları maksimize etme potansiyeline sahiptir.

Bu çalışmanın ortaya koyduğu üzere, güvenilir bir YZ platformunun temelini, kapsamlı bir veri yönetişimi, modelin bütünlüğünü koruyan teknik önlemler ve tüm yaşam döngüsüne yayılan güçlü bir altyapı güvenliği oluşturmaktadır. Ancak en gelişmiş teknik önlemler bile, etkin bir yönetişim ve insan gözetimi olmaksızın yetersiz kalacaktır. YZ'nin en büyük tehditlerinin, kontrol edilemeyen bir teknolojiden değil, insan gözetiminin ve yönetişiminin eksikliğinden kaynaklandığına dair yaygın bir kanaat bulunmaktadır. Geleceğin YZ'sini şekillendirirken, teknolojinin kendisinden çok, onu nasıl yöneteceğimiz, etik değerlerle nasıl uyumlu hale getireceğimiz ve kamuoyunda nasıl güven tesis edeceğimiz belirleyici olacaktır. Bu nedenle, bir ülkenin YZ stratejisi, inovasyonu teşvik ederken, güvenlik ve güvenilirlik ilkelerinden asla taviz vermemelidir (İstanbul Barosu, 2021) (Wiz Experts Team, 2025).

5.3.8. Kaynakça

Amos, Z. (2022, October 10). Veri Zehirlenmesi: Bir Çözümü Var mı? Retrieved September 1, 2025, from Unite.AI website: https://www.unite.ai/tr/data-poisoning-is-there-a-solution/

Arsenault, B. (2024). YZ çağında siber tehditlerin üstesinden gelmek ve savunmayı güçlendirmek. Retrieved September 1, 2025, from Microsoft.com website: https://www.microsoft.com/tr-tr/security/security-insider/intelligence-reports/cyber-signals-issue-6-navigating-cyberthreats-and-strengthening-defenses

Chen, M. (2023, December 20). 6 Common YZ Model Training Challenges. Retrieved September 1, 2025, from Oracle.com website: https://www.oracle.com/tr/artificial-intelligence/ai-model-training-challenges/

Ekaglobal. (2023). ISO/IEC 42001:2023 YZ Yönetim Sistemi. Retrieved September 1, 2025, from Ekaglobal.com.tr website: https://ekaglobal.com.tr/hizmet/iso-iec-42001-2023-yapay-zeka-yonetim-sistemi.html

European Union. (2024). Regulation - EU - 2024/1689 - EN - EUR-Lex. Retrieved from Europa.eu website: https://eur-lex.europa.eu/legal-content/EN/TXT/?uri=OJ%3AL_202401689

Farnham, K. (2025, July 24). NIST YZ Risk Management Framework: A simple guide to smarter YZ governance. Retrieved from Diligent.com website: https://www.diligent.com/resources/blog/nist-ai-risk-management-framework

Gürkan Coşkun. (2024, November 8). Güvenilir YZ Kavramı. Retrieved September 1, 2025, from YZ Kanunu website: https://www.yapayzekakanunu.com/post/g%C3%BCveni-li-r-yapay-Zekâ-kavrami

İstanbul Barosu. (2021). Güvenilir YZ İlkeleri. Retrieved September 1, 2025, from istanbulbarosu.org.tr website: https://www.istanbulbarosu.org.tr/files/komisyonlar/yzcg/guveniliryapayZekâicinetikkilavuz.pdf

Jamil, A. (2024, September 3). Umetech. Retrieved from Umetech.net website: https://www.umetech.net/blog-posts/successful-implementations-of-ai-in-cyber-defense

Kaminsky, S. (2024, February 21). ChatGPT, Gemini ve diğer Yapay Zekâlar nasıl güvenle kullanılır. Retrieved September 1, 2025, from Kaspersky.com.tr website: https://www.kaspersky.com.tr/blog/how-to-use-chatgpt-ai-assistants-securely-2024/12053/

Microsoft. (2025). YZ Güvenliği nedir? YZ Sistemlerini Koruyun | Microsoft Güvenlik. Retrieved September 1, 2025, from Microsoft.com website: https://www.microsoft.com/tr-tr/security/business/security-101/what-is-ai-security

NCSC. (2023, November 14). Case study: The cyber security of artificial intelligence. Retrieved from www.ncsc.gov.uk website: https://www.ncsc.gov.uk/collection/annual-review-2023/technology/case-study-cyber-security-ai

NIST. (2025). Artificial Intelligence: Adversarial Machine Learning | NCCoE. Retrieved September 1, 2025, from www.nccoe.nist.gov website: https://www.nccoe.nist.gov/ai/adversarial-machine-learning

Palo Alto Networks. (2025). What Is Adversarial YZ in Machine Learning? Retrieved September 1, 2025, from Palo Alto Networks website: https://www.paloaltonetworks.com/cyberpedia/what-are-adversarial-attacks-on-AI-Machine-Learning

Para Dergi. (2024, April 19). Ülkelerin YZ ile ilgili düzenlemeleri. Retrieved September 1, 2025, from paradergi website: https://www.paradergi.com.tr/dunya/2024/04/19/ulkelerin-yapay-Zekâ-ile-ilgili-duzenlemeleri

Şahin Ardıyok. (2023, December 22). Yapay Zekâya Dünyada İlk Regülasyon Müdahalesi Avrupa Birliği'nden Geldi: AB YZ Yasası. Retrieved August 31, 2025, from Rekabet ve Regülasyon website: https://www.rekabetregulasyon.com/yapay-Zekâya-dunyada-ilk-regulasyon-mudahalesi-avrupa-birliginden-geldi-ab-yapay-Zekâ-yasasi/

SAP. (2025). İş dünyasında etkili YZ uygulaması: Başarı için adımlar. Retrieved September 1, 2025, from Sap.com website: https://www.sap.com/turkey/resources/effective-ai-implementation-in-business

SAP . (2024, August 9). YZ etiği nedir? Yapay Zekâda etik rolü | SAP. Retrieved September 1, 2025, from Sap.com website: https://www.sap.com/turkey/resources/what-is-ai-ethics

Trendmicro. (2024). YZ ile İlgili Güvenlik Riskleri Nelerdir? Retrieved September 1, 2025, from Trend Micro website: https://www.trendmicro.com/tr_tr/what-is/ai/security-risks.html

Wiz Experts Team. (2025, January 31). NIST AI Risk Management Framework: A tl;dr. Retrieved from wiz.io website: https://www.wiz.io/academy/nist-ai-risk-management-framework

Ye, J. (2023, July 14). China says generative AI rules to apply only to products for the public. *Reuters*. Retrieved from https://www.reuters.com/technology/china-issues-temporary-rules-generative-ai-services-2023-07-13/

5.4. Veri Bilimci ve Veri Mühendisi Açısından ISO/IEC 5338 Standardının Değerlendirilmesi

Doç. Dr. Meltem Eryılmaz [a], Prof. Dr. Gökçe Nur Yılmaz [b]

[a] *TBD YK Üyesi, OSTİM Teknik Üniversitesi Yazılım Müh. Böl., meltem.eryilmaz@ostimteknik.edu.tr*
[b] *TBD Yönetim Kurulu Üyesi, TED Üniversitesi Bilgisayar Müh. Bölüm Bşk., nur.gkc@gmail.com*

Yapay Zekâ (YZ) sistemlerinin güvenli, şeffaf, hesap verebilir ve etik ilkelere uygun olarak tasarlanabilmesi, geliştirilmesi ve sürdürülebilir biçimde yönetilebilmesi hem uluslararası hem de ulusal düzeyde kabul görmüş standartlara uyumla mümkündür. Uluslararası ISO/IEC standart ailesi, IEEE 7000 serisi ve Avrupa Birliğinin (AB) YZ Tüzüğü gibi düzenlemeleri veri bilimcilerden yazılım mühendislerine, yöneticilerden politika yapıcılara kadar tüm paydaşlar için ortak bir referans çerçevesi sunarken; ISO/IEC 42001, kurumların YZ yönetim sistemlerinin planlanması, uygulanması, izlenmesi ve sürekli iyileştirilmesi için üst düzey bir yönetişim yapısı oluşturur.

Bu çerçeveyi tamamlayan ISO/IEC 22989, YZ kavramları ve terminolojisini standardize ederek farklı disiplinler arasında ortak bir dil kurulmasını sağlarken; ISO/IEC 23053, özellikle makine öğrenmesi tabanlı YZ sistemlerinin bütün yaşam döngüsünü yapılandıran bir süreç modeli sunar. Öte yandan ISO/IEC 23894, teknik, etik, güvenlik ve mahremiyet risklerinin tüm yaşam döngüsü boyunca yönetilmesini zorunlu kılarak sorumlu YZ geliştirme uygulamalarını güçlendirir. Şeffaflık, açıklanabilirlik ve hesap verebilirliği teknik bir gereklilik hâline getiren ISO/IEC 25059 ile IEEE 7001, kullanıcıların ve paydaşların sistemin işleyişini anlamasını sağlayan kritik ilkeleri belirler. ISO 5338'in 6.3 ve 6.4 maddeleri ile ISO/IEC standart ailesi, veri bilimci ve mühendislerin çalışmalarında rehber olarak kullanabileceği detaylı, yapılandırılmış ve bütünleşik çerçeveler sunmakta; aynı zamanda YZ sistemlerinin geliştirilmesi, uygulanması ve denetlenmesi için bütüncül bir yol haritası oluşturmaktadır. Bütün bu standartlar ile birlikte ISO/IEC 5259 serisinin veri kalitesinin ölçülmesi, izlenmesi ve iyileştirilmesine yönelik çerçeveleri, YZ sistemlerinin tutarlı bir kalite, güvenlik ve etik düzeyini korumasında tamamlayıcı bir rol oynar.

5.4.1. Uluslararası Standartların YZ Sistemlerindeki Rolü

YZ sistemlerinin topluma sağlayacağı güven, şeffaflık ve sorumluluk ancak uluslararası kabul görmüş standartların etkin şekilde uygulanmasıyla mümkün olur. ISO/IEC 42001 standardı, farklı ölçek ve sektörlerde faaliyet gösteren kurumlar için YZ yönetim sistemlerinin tüm yaşam döngüsünde güvenli, etik ve sürdürülebilir çözümler geliştirilmesini hedefler. Bu standart, YZ sistemlerinin planlanmasından dağıtımına, performans değerlendirmelerinden sürekli iyileştirme süreçlerine kadar tüm aşamalarda etik prensiplerin (adalet, ayrımcılığın önlenmesi, veri gizliliği ve insan haklarına saygı) aktif olarak yer almasını zorunlu kılar. Böylece, sistemler arasında ülkeye ve uygulama alanına bağlı çeşitlilik olsa da; ortak bir kalite, şeffaflık ve etik seviyesinin korunması hedeflenir.

Ayrıca, ISO/IEC 22989 standardı, YZ terminolojisi ve temel kavramlar konusunda hem disiplinler arası iletişimde ortak zemin oluşturmakta hem de uygulamada bütünlük yaratmaktadır. Bu sayede veri bilimi ekipleri ile yönetişim birimleri, geliştirme, eğitim, sertifikasyon gibi farklı alanlarda daha uyumlu ve koordineli çalışabilmektedir. ISO/IEC 23053 ise özellikle makine öğrenmesi temelli YZ sistemlerinde, sorumlu ve etkin kullanım için gereken temel ilkeleri ve yapısal çerçeveleri ortaya koyar. Kurumsal risk yönetimi açısından ISO 31000 ve ISO/IEC 23894 standartları ile entegrasyon, hem öngörülen hem de beklenmedik tehlikelerin belirlenmesi ve yönetilmesi açısından önemli bir avantaj sunar.

5.4.2. YZ Yaşam Döngüsüne Özgü Standartlar ve Uzmanlık Alanları

YZ sistemlerinin yaşam döngüsünde, ulusal ve uluslararası standartlara göre başarılı ve etik uygulamalar geliştirebilmek için farklı uzmanlık alanlarıyla iş birliği yapmak gereklidir. Süreç bazında bu ihtiyaçlar aşağıdaki tabloda detaylı olarak belirtilmektedir.

Süreç / Aşama	İlgili Standartlar	İşbirliği Gereken Uzmanlık Alanları
1. Problem Tanımı ve Gereksinim analizi	ISO/IEC 22989, ISO/IEC 23053	• Alan uzmanları (domain experts): Uyugulamanın yapılacağı sektöre özgü bilgi (sağlık, eğitim, enerji) • İnsan-bilgisayar etkileşimi (HCI) uzmanları: Kullanıcı ihtiyaçlarının doğru modellenmesi • Etik uzmanları: Temel etik sorunların erken aşamada ele alınması
2. Veri Toplama, Hazırlama ve Yönetimi	ISO/IEC 27001, ISO/IEC 29100, ISO/IEC 23053	• Veri bilimciler ve istatistikçiler: Veri kalitesi, önyargıların giderilmesi, örnekleme yöntemleri • Hukukçular / KVKK ve GDPR uzmanları: Kişisel verilerin korunması ve yasal uygunluk • Siber güvenlik uzmanları: Veri güvenliği önlemleri

Süreç / Aşama	İlgili Standartlar	İşbirliği Gereken Uzmanlık Alanları
3. Model Geliştirme ve Doğrulama	ISO/IEC 23053, ISO/IEC 23894	• Makine öğrenmesi / derin öğrenme uzmanları • Yazılım mühendisleri: Entegrasyon, sürüm kontrolü, algoritma optimizasyonu • Etik ve sosyal bilimciler: Model önyargısı, adalet, toplumsal etkiler
4. Uygulama / Yaygınlaştırma / Eğitim	IEEE 7000 serisi, ISO/IEC 24029 (eğitimde YZ)	• Eğitim teknolojisi uzmanları: Sürekli izleme, performans doğrulama • Kullanıcı deneyimi (UX) uzmanları: Kullanılabilirlik ve erişilebilirlik • Proje yöneticileri / İş geliştirme uzmanları: Sürdürülebilirlik ve adaptasyon
5. İzleme, Denetim ve Etki Değerlendirmesi	ISO/IEC 23894, ISO 31000 (Risk Yönetimi), IEEE 7001	• Kalite güvence uzmanları: Sürekli izleme, performans doğrulama • Bağımsız denetçiler / etik kurullar: Şeffaflık ve hesap verebilirlik • Politika yapıcılar / regülasyon uzmanları: Yasal uyumluluk ve kurumsal etki analizi

5.4.3. Uygunluk Sağlanması Gereken Standartlar

Bu yaşam döngüsü aşamalarının her birinde hangi uluslararası standartlara uyum sağlanması gerektiği ise aşağıdaki tabloda bütüncül olarak sunulmaktadır.

Alan	Standart	Amaç
Yönetim & Yönetişim	ISO/IEC 42001	YZ Yönetim Sistemi kurulumu ve iyileştirilmesi
Terminoloji	ISO/IEC 22989	Ortak kavramsal çerçeve
Veri Kalitesi	ISO/IEC 25012, ISO/IEC 5259 serisi	Veri doğruluğu, bütünlük, denge
Güvenlik & Mahremiyet	ISO/IEC 27001, ISO/IEC 29100	Güvenlik ve mahremiyet çerçevesi
ML Çerçevesi	ISO/IEC 23053	ML tabanlı YZ sistemlerinin düzenli yaşam döngüsü
Risk Yönetimi	ISO/IEC 23894, ISO 31000	Teknik–etik risk yönetimi
Etik & Şeffaflık	ISO/IEC 25059, IEEE 7001	Açıklanabilirlik ve adalet
Etki Analizi	ISO/IEC 42005, ISO/IEC 42006	Sosyal, hukuki ve etik etki değerlendirmesi
Yeni Teknikler	ISO/IEC TS 8200	Yeni YZ yöntemlerine uyum

5.4.4. Veri Kalitesi, Şeffaflık ve Risk Yönetimi

YZ projelerinin teknik başarısı ve etik uygunluğunun sağlanmasında, veri yönetimi ve kalite standartları büyük önem taşır. ISO/IEC 5259 serisi; analitik ve makine öğrenmesi süreçlerinde kaliteli, bütünlüklü ve güncel verinin temini, sürdürülmesi ve gerektiğinde

iyileştirilmesine yönelik adımları açıkça ortaya koymaktadır. Eğitim sürecinde kullanılan verinin dengeli, kapsayıcı, hatasız ve eksiksiz olması, hem algoritmik başarım hem de toplumsal adalet açısından hayati önemdedir. Eksik veya dengesiz veriler, YZ sistemlerinde önyargı ve ayrımcılık risklerini artırmakta, teknik hata oranlarını ise yukarı çekmektedir.

Bunun yanı sıra, karar alma süreçlerinde insan denetiminin mümkün kılınması, sistemlerin şeffaf ve hesap verebilir olmasını sağlar. Standartlaştırılmış istatistiksel analizler, otomasyonlu kalite ölçüm ve denetim protokolleri, proje çıktılarında güven inşasına ve yasal uyumluluğun sağlanmasına ayrıca katkıda bulunur. Özellikle modelin çalışması, kullanılan veri setleri ve karar mekanizmalarına dair şeffaf dokümantasyon, gerek denetimlerde gerekse son kullanıcı güveninde kritik rol oynar.

5.4.5. Etik Risklerin Yönetimi

YZ sistemlerinin gerçek dünyada başarılı, adil ve toplum yararına kullanılabilmesi için etik risklerin yönetimi bütüncül bir yaklaşımla ele alınmalıdır. “Ethics-by-design” ilkesi, etik değerlerin yalnızca nihai ürüne sonradan eklenmesi değil; aksine, tasarım ve geliştirme süreçlerinin tamamında, baştan itibaren gözetilmesini ifade eder. ISO/IEC 42001 ve ISO/IEC 25059 standartlarında teknik ve yönetsel gereklilik olarak yer alan bu yaklaşım, model şeffaflığı, insan kontrolü, hesap verebilirlik ve işlevsel adalet gibi başlıkları öne çıkarır. Özellikle ISO/IEC 25059, şeffaflığı teknik bir parametreye dönüştürerek, sistem işleyişinin kamuya açık ve anlaşılabilir olmasını talep etmektedir. Bu, kullanıcıların veya paydaşların modelin nasıl çalıştığını, aldığı kararları ve potansiyel riskleri sorgulayabilmesini sağlar.

Etik riskler yalnızca eğitim verisinden değil, model mimarisinden, tasarım tercihlerinden ve kullanım senaryolarından da kaynaklanabilmektedir; bu tür uyumsuzluklar, model çıktılarının toplumda ayrımcılığa, güven erozyonuna veya doğrudan hukuki sorunlara yol açmasına neden olabilir. Bu nedenle ISO/IEC 5259-2 ve ISO/IEC 42001, YZ sistemlerinde sürekli kalite denetimi, etik gözetim ve risk izleme mekanizmalarının kurumsallaştırılmasını zorunlu kılar. Çoklu ve dinamik olarak güncellenen metriklerin, örneğin önyargı analizleri, dengesizlik ölçümleri ve performans değişkenlikleri, düzenli biçimde izlenmesi, gerektiğinde sistematik müdahalelerin yapılmasını mümkün kılarak etik uyumu teknik düzeyde güvence altına alır. Dolayısı ile yapay zekânın etik kullanımı bireysel haklar, veri mahremiyeti, toplumsal adalet ve kurumsal hesap verebilirlik gibi temel değerlerle doğrudan ilişkili olduğundan, şeffaflık, açıklanabilirlik ve etik sorumluluk ilkelerinin standartlaştırılması hem

kullanıcı güveninin tesis edilmesi hem de toplumsal meşruiyetin sürdürülebilirliği açısından kritik öneme sahiptir.

5.4.6. Ulusal Perspektif

YZ yaşam döngüsünün her aşamasında standart uyumu kritik önemdedir. Veri kalitesi, model geliştirme süreçleri, açıklanabilirlik, etik tasarım ilkeleri, risk yönetimi ve denetim mekanizmaları, raporun temel bileşenlerini oluşturmaktadır. Türkiye'nin YZ yönetişiminin uluslararası standartlarla uyumlu ve aynı zamanda yerel ihtiyaçlara uygun bir yapıda geliştirilmesi, sürdürülebilir bir ulusal YZ stratejisinin temelini oluşturmaktadır.

Türkiye Bilişim Derneği'nin şekillendirdiği ulusal yönetişim perspektifi, YZ ekosisteminin gelişiminde hem uluslararası standartlarla uyumu esas almakta hem de Türkiye'nin kendine özgü gereksinimlerine karşılık verecek esneklik ve güncelliğe sahip bir yapı önermektedir. Bu yaklaşım, ülkenin sayısal (dijital) dönüşüm sürecinde güçlü ve sürdürülebilir bir YZ ekosistemi oluşturulmasının temel anahtarıdır.

YZ teknolojilerinin çeşitli sektörlerde güvenilir biçimde yaygınlaştırılması için, hem kamu hem de özel sektör kurumlarında şeffaflık, etik, hesap verebilirlik ve sürdürülebilirlik ilkeleri yalnızca stratejik planların değil, tüm operasyonların merkezine yerleştirilmelidir. Özellikle mevzuat geliştirme ve günlük uygulamalarda; AB YZ Tüzüğü, Genel Veri Koruma Tüzüğü (GDPR) ve Genel Veri Koruma Tüzüğü (OECD) YZ ilkeleri gibi uluslararası bağlayıcı çerçeveler ile tam uyumlu, fakat Türkiye'nin sektörel ve kültürel ihtiyaçlarına göre şekillenmiş kendi denetim ve rehberlik mekanizmaları kurulmalıdır. Bu sayede, hem uluslararası işbirlikleri ve teknoloji transferi kolaylaşacak, hem de yerel paydaşların çözüm süreçlerine katılımı artacaktır.

Bununla birlikte, disiplinler arası eğitim programlarının yaygınlaştırılması, YZ farkındalığını artırıcı projelerin desteklenmesi ve kamu-özel sektör ile akademi arasında güçlü işbirliklerinin kurulması, ekosistemin dayanıklılığı ve inovasyon kapasitesi açısından kritik önemdedir. Farklı uzmanlık alanlarından gelen ekiplerin ortak hedefler için uyumlu çalışması, hem teknik mükemmellik hem de etik ve toplumsal sorumluluk boyutunda yüksek standartların yakalanmasını sağlar.

Veri bilimi projeleri özelinde, nitelikli ve yetkin profesyonellerden oluşan ekiplerin kurulması, kurumlar arası bilgi paylaşımının teşvik edilmesi ve şeffaf denetim

mekanizmalarının etkinleştirilmesi gereklidir. Bu kapsamda, kamu ve özel sektör uygulamalarında YZ odaklı politika, standart ve iyi uygulama rehberlerinin ortak bir çatı altında bütünleştirilmesi, bütüncül ve tutarlı bir sayısal (dijital) dönüşüm süreci oluşturur.

Türkiye'nin sahip olduğu KVKK başta olmak üzere, ilgili ulusal veri koruma ve güvenlik regülasyonlarının, Avrupa'daki ve dünyadaki standart ve düzenlemelerle entegre çalışması sağlanmalıdır. Ulusal regülasyonların, AB YZ tüzüğü ve uluslararası risk-etik standartlarıyla uyumlu şekilde güncellenmesi, Türkiye'nin küresel dijital rekabet gücünü önemli ölçüde artıracak ve toplumun YZ'ya duyduğu güveni derinleştirecektir.

5.4.7. Teknoloji Şirketleri Tarafından Geliştirilen Yapay Zekâ Usul ve Esasları (Protokoller)

Kurumsal ölçekli YZ uygulamalarının güvenilir, sürdürülebilir ve düzenleyici çerçevelerle uyumlu biçimde işletilebilmesi, teknoloji şirketlerinin uluslararası standartlar doğrultusunda yapılandırılmış protokoller benimsemesini gerektirmektedir. ISO/IEC ve IEEE standart ailesine dayanan bu protokoller, YZ yaşam döngüsünün her aşamasında etik tasarım ilkelerinin kurumsallaştırılmasını, veri kalitesinin metodolojik olarak güvence altına alınmasını, risklerin nicel ve nitel yöntemlerle sürekli değerlendirilmesini, sistemlerin sosyal, hukuki ve etik etkilerinin bütüncül biçimde analiz edilmesini ve bağımsız denetim mekanizmalarıyla kurumsal hesap verebilirliğin sağlanmasını hedeflemektedir. Bununla birlikte, açıklanabilirlik odaklı teknik gereksinimler, model davranışının şeffaflaştırılmasını ve karar süreçlerinin izlenebilirliğini zorunlu kılarak güvenilir YZ yönetişim modellerinin temel bileşenlerinden birini oluşturur. Aşağıdaki tabloda, bu protokollerin kavramsal ve teknik bileşenleri ayrıntılı biçimde sunulmaktadır.

Teknoloji Şirketleri İçin YZ Protokolleri

Protokol	Standart Dayanağı	Açıklama
Etik Tasarım Protokolü	ISO 25059, IEEE 7001	Şeffaflık, adalet, hesap verebilirlik
Veri Kalitesi Protokolü	ISO/IEC 5259	Kaynak doğrulama, denge kontrolü
Risk Yönetimi Protokolü	ISO/IEC 23894, ISO 31000	Sürekli risk analizi
Etki Değerlendirme Protokolü	ISO/IEC 42005, 42006	Sosyal, etik, hukuki etki analizi
Denetim & Sertifikasyon Protokolü	ISO/IEC 42006	Bağımsız dış denetim

Şeffaflık Protokolü	IEEE 7001	Açıklanabilir model tasarımı

5.4.8. Sonuç

Tüm bu değerlendirmeler, YZ ekosisteminin güvenilir, etik ve sürdürülebilir bir şekilde gelişebilmesi için uluslararası standartlarla uyumlu, disiplinler arası ve çok katmanlı bir yönetişim yaklaşımına ihtiyaç duyulduğunu açıkça ortaya koymaktadır. ISO/IEC standart ailesinin sağladığı kavramsal, teknik ve yönetsel çerçeveler; veri kalitesi, açıklanabilirlik, etik tasarım, risk yönetimi ve denetim gibi YZ uygulamalarının temel yapı taşlarını sistematik biçimde düzenlerken, ulusal düzeyde geliştirilecek strateji ve politikaların da bu bütüncül modele entegre edilmesini gerekli kılmaktadır. Türkiye'nin YZ alanındaki regülasyonlarını ve kurumsal yapılanmalarını bu standartlarla uyumlu hâle getirmesi, yalnızca uluslararası rekabet gücünü artırmakla kalmayacak, aynı zamanda toplumsal güveni pekiştirerek YZ teknolojilerinin sosyal faydasını en üst düzeye çıkaracaktır. Bu bağlamda, güçlü bir ulusal yönetişim vizyonu, nitelikli insan kaynağı, açık ve hesap verebilir kurum içi süreçler ile sürekli denetime dayalı sürdürülebilir bir YZ ekosistemi inşa etmek, hem kamu hem özel sektör hem de akademi için stratejik bir zorunluluk hâline gelmiştir.

5.4.9. Referanslar

1. A. Simonetta and M. C. Paoletti, ISO/IEC Standards and Design of an Artificial Intelligence System, IWESQ 2024: 6th International Workshop on Experience with SQuaRE Family and its Future Direction, 2024.
2. ISO/IEC 42001:2023 – Yapay Zekâ Yönetim Sistemi Standardı
3. ISO/IEC 23053:2022 – Makine Öğrenmesi YZ Sistemleri Çerçevesi
4. A. Simonetta, M. C. Paoletti, A. Venticinque, The use of maximum completeness to estimate bias in AI-based recommendation systems, CEUR Workshop Proceedings 3360 (2022) pp. 76–84.
5. ISO/IEC 22989 Information technology — Artificial intelligence — Artificial intelligence concepts and terminology, 2022. https://www.iso.org/standard/74296.html
6. ISO/IEC 8183:2023 Information technology — Artificial intelligence — Data life cycle framework, 2023. https://www.iso.org/standard/83002.html

7. ISO/IEC 25012:2008 Software engineering — Software product Quality Requirements and Evaluation (SQuaRE) — Data quality model, 2008. URL: https://www.iso.org/standard/35736.html
8. ISO/IEC 23894 Information technology — Artificial intelligence — Guidance on risk management, 2023. https://www.iso.org/standard/77304.html
9. ISO/IEC DIS 42006 Information technology — Artifcial intelligence — Requirements for bodies providing audit and certification of artificial intelligence management systems, Under development. https://www.iso.org/standard/44546.html
10. ISO/IEC 5259-3 Artificial intelligence — Data quality for analytics and machine learning (ML), 2024. https://www.iso.org/standard/81092.html
11. ISO/IEC 25059:2023 Software engineering Systems and software Quality Requirements and Evaluation (SQuaRE) — Quality model for AI systems, 2023. URL: https://www.iso.org/standard/80655.html

5.5. Yapay Zekâ Yaşam Döngüsü, Standartlar ve Sanatsal Projelerde Veri Bilimcinin Rolü

Mert Çobanov

Refik Anadol Stüdyo, mertcobanov@gmail.com

Bu bölümün amacı, *Refik Anadol Studio* gibi büyük ölçekli veri ve YZ odaklı sanat projelerinde, YZ yaşam döngüsünü uzman perspektifinden değerlendirmektir. Geniş hacimli metin, görsel ve ses arşivlerinden YZ modelleri eğiterek dijital sergiler, veri heykelleri ve interaktif işler üreten projeler, veri toplama ve temizlemeden model eğitimi ve servis edilmesine kadar uçtan uca bir veri boru hattı (data pipeline) gerektirir. Özellikle devlet arşivleri veya kurumsal sanat vakıflarının özel arşivleri gibi hassas kaynaklardan gelen verinin gizliliği ve güvenliği son derece kritik olup, yüksek bilgi güvenliği önlemleri gerektirmektedir. Ayrıca stüdyo, son teknoloji büyük ölçekli YZ modellerini ve en yeni akademik modelleri paralel olarak eğitebilecek yüksek hesaplama gücüne sahiptir. Bu bağlamda, çalışmada YZ yaşam döngüsünün hangi süreçlerinde uzmanın görev aldığı, bu süreçlerdeki sorumlulukları, mevcut çalışma yönteminin hangi açılardan güncellenmeye ihtiyaç duyduğu ve bunlar için standartlara gereksinim olup olmadığı incelenecektir. Ardından yaşam döngüsü kapsamında uyulması gereken ulusal/uluslararası standartlar (ISO/IEC 5338 teknik raporunda notlar bölümünde belirtilenler dâhil) ele alınacak, belirlenen standartlara göre hangi aşamada hangi uzmanlık alanlarıyla işbirliği yapılması gerektiği tartışılacaktır. Son olarak standartların önerdiği yöntemler (araçlar, protokoller vb.) ve yakın gelecekte YZ sanat projelerinde bu standartların ve disiplinler arası işbirliğinin rolü değerlendirilecektir.

5.5.1. YZ Yaşam Döngüsünde Veri Bilimcinin Rolü ve Sorumlulukları

YZ Yaşam Döngüsünde Veri Bilimcinin Rolü ve Sorumlulukları: Refik Anadol Studio gibi veri odaklı YZ sanat projelerinde uzman, YZ sistem yaşam döngüsünün başlangıçtan uygulamaya tüm süreçlerinde aktif rol alır. ISO/IEC 5338:2023 standardı, YZ sistemleri için fikir aşamasından nihai dağıtıma dek uzanan yaşam döngüsü süreçlerini tanımlayarak ortak bir çerçeve sunmaktadır. Bu çerçeveye paralel olarak, uzman; veri hazırlığı, model geliştirme, doğrulama, dağıtım ve izleme gibi aşamaların tümüne katkı sağlar. Söz konusu uluslararası

standartlar, YZ'nin insan yararına hizmet etmesi gerektiğini vurgulayan insan merkezli yaklaşımlarla uyumludur. Bu nedenle uzman, teknik geliştirmelerin yanı sıra insan faktörlerini ve etik ilkeleri de gözeterek çalışır.

1. Bağlam ve Veri Toplama Aşaması

İlk adımda, projenin amacına uygun verilerin toplanması ve kullanım bağlamının tanımlanması gelir. Uzman, sanat eserinin hangi ortamda ve kimler tarafından deneyimleneceğini göz önünde bulundurarak gereksinimleri belirler. Bu kapsamda farklı paydaşlarla (arşiv sahipleri, küratörler, sanatçılar) görüşerek proje hedeflerini netleştirir ve verinin kaynağını, niteliğini ve hassasiyet düzeyini analiz eder. Örneğin, biyoloji temalı bir YZ sanat projesinde üniversite biyologları ve doğa bilimciler ile çalışarak verilerin bilimsel doğrulukla kullanılmasını sağlamak; tarihî bir arşiv projesinde antropologlar, tarihçiler veya sosyologlarla işbirliği yaparak içeriğin bağlamını doğru anlamak bu aşamadaki kritik görevlere girer. Uzman, veri toplarken gizlilik ve telif kısıtlarını da değerlendirir; hassas arşiv verilerinin izinler dâhilinde kullanılması ve kişisel verilerin korunması için gerekli hukuki onayların alınması noktasında hukuk danışmanları ile birlikte hareket eder. Toplanan ham verinin temizlenmesi, hataların ayıklanması ve standardize formatlara dönüştürülmesi de uzmanın sorumluluğundadır. ISO/IEC 5338 standardı, YZ sistemlerinde verinin yaşam döngüsünün güvenli yönetimini vurgulayarak veri kalitesi, kaynağı ve kökeni konusunda titiz bir dokümantasyon yapılmasını önermektedir. Uzman, bu doğrultuda her veri için köken/proveni bilgilerini kayıt altına alır; böylece model geliştirme sürecinde verinin evrimi izlenebilir ve projenin sonunda veri kullanımının şeffaflığı sağlanır. Aynı zamanda bu uygulama, gizlilik düzenlemelerine uyum ve hassas verilerin korunması açısından da kritiktir. Sonuç olarak, bağlam analizi ve veri hazırlığı aşamasında uzmanın görevi, proje hedeflerine uygun, güvenilir ve gizlilik ilkelerine uygun bir veri gövdesi oluşturmaktır. Bu süreçte uluslararası standartlar doğrudan bir kural koymasa da, örneğin kullanıcı merkezli tasarım standardı ISO 9241-210'un "kullanım bağlamını anlama" ilkesiyle uyumlu şekilde, uzmanın dikkat etmesi gereken noktalar için yol gösterir.

2. Model Geliştirme ve Eğitim Aşaması

İkinci aşamada, toplanan veriden değer üretecek YZ modelinin geliştirilmesi yer alır. Uzman, proje için en uygun makine öğrenimi mimarisini ve algoritmalarını seçer; Refik Anadol Studio'da yaygın olarak en son yayınlanan (state-of-the-art) modelleri ve teknikleri takip ederek bunları sanatsal ihtiyaçlara göre uyarlar. Model mimarisi belirlendikten sonra, yüksek hesaplama gücüne sahip dağıtık altyapıları kullanarak modelin eğitimini gerçekleştirir. Bu

süreçte görevler, model üst parametre optimizasyonu, eğitim sürecinin izlenmesi, ara sonuçların değerlendirilmesi ve gerektiğinde veri artırma (augmentation) veya dengeleme gibi tekniklerle model performansının iyileştirilmesini içerir. Uzman, eğitilen modelin sadece istatistiksel performansını değil, aynı zamanda sanatsal estetik çıktılar üretme kapasitesini de değerlendirir; örneğin arzu edilen görsel tarz veya etkiyi yakalayıp yakalamadığını kontrol eder. Model geliştirme aşamasında etik ve risk boyutlarını da yönetmek gerekir: Veri kaynaklı önyargılar (bias) varsa bunları tespit edip gidermek, modelin istenmeyen veya zararlı çıktılar üretmemesi için gerekli filtreleme ve düzenlemeleri uygulamak bu kapsamda yer alır. Uzman, ISO/IEC 23894 gibi YZ risk yönetimi rehberlerine başvurarak olası riskleri belirler ve azaltıcı önlemler planlar. Süreç boyunca güvenlik önlemleri gözetilir; örneğin modelin eğitiminde kullanılan verilerin ve çıktının izinsiz erişimlere karşı korunması amacıyla ISO/IEC 27001 standardındaki bilgi güvenliği kontrolleri uygulanır. Model geliştirme ve test sürecinde insan faktörü de tamamen dışlanmaz. Standartlar, kritik karar noktalarında insan gözetiminin önemini vurgular. Nitekim ISO/IEC 5338, YZ sistemlerinin geliştirilmesinde otomasyon ile insan yargısı arasında denge kurulmasını; özellikle güvenlik veya etik sonuçlar doğurabilecek kararların insanlar tarafından gözden geçirilebilir olmasını şart koşar. Bu doğrultuda, stüdyodaki uzman gerek modelin ara çıktıları gerekse nihai performansı konusunda diğer ekip üyelerinin ve gerektiğinde dış uzmanların (ör. etik danışmanlar, domain uzmanları) görüşlerini alarak model üzerinde iyileştirmeler yapar.

3. Doğrulama, Dağıtım ve Operasyon Aşaması

Uzman, model eğitiminin ardından sistemin doğrulama (validation) ve devreye alma süreçlerini yönetir. Doğrulama adımında, modelin hem teknik gereksinimleri karşıladığını hem de sanat projesinin konseptine uygun ve paydaş beklentilerini tatmin edici çıktılar ürettiğini test eder. Bu amaçla, mümkünse gerçek dünyadaki kullanım senaryolarını simüle eden veri setleriyle modeli sınar; örneğin interaktif bir sergi için modelin farklı etkileşim senaryolarına verdiği tepkileri değerlendirir. Python tarafında doğrulama süreci, veri hatlarında şema ve kalite kontrolleri için Great Expectations veya benzeri araçlarla otomatikleştirilir; üretimde dağılım kaymalarını (data/concept drift) ve kalite metriklerini düzenli izlemek için Evidently ile raporlar oluşturulur; açıklanabilirlik ve güvenilirlik için SHAP/Captum gibi kütüphaneler kullanılabilir; saldırı dayanıklılığı ise ISO/IEC TR 24029-1'deki sağlamlık ilkeleriyle uyumlu olarak Adversarial Robustness Toolbox (ART) veya Foolbox ile adversary testleri yapılarak ölçülür. Ardından, modelin sunum (servis) ortamına entegrasyonu gelir. Bu aşamada görevler, modelin optimize edilmesi (gerekiyorsa sıkıştırma veya hızlı çalışması için dönüştürme), uygun

donanım veya bulut altyapısına yerleştirilmesi, ve istemci arayüzleriyle (örneğin görsel-işitsel enstalasyonun yazılımıyla) entegrasyonunu içerir. Optimizasyon tarafında, PyTorch/ TensorFlow modellerinin TorchScript veya ONNX'e dışa aktarımı, ONNX Runtime ile dinamik/statik INT8 quantization, gerekirse bilgi damıtma (distillation) uygulanır; NVIDIA GPU'larda çıkarımda TensorRT'den yararlanılır, Apple/Intel hedefleri için coremltools/OpenVINO tercih edilebilir. Servisleme mimarisi olarak düşük gecikmeli REST/gRPC için FastAPI + Uvicorn/Gunicorn yaygınca kullanılır, yoğun GPU konsolidasyonu ve model havuzu yönetimi için NVIDIA Triton Inference Server; Kubernetes üzerinde standartlaştırılmış dağıtımlar için KServe; tek paketle üretime çıkma ve gözlemlenebilirlik için BentoML; büyük dil modellerinde yüksek throughput için vLLM veya Hugging Face Text Generation Inference (TGI) tercih edilebilir. Dağıtım sırasında güvenlik ve gizlilik önlemleri devam eder; model servisinin yetkisiz erişime karşı korunması, mTLS ile servis-servis iletişimin şifrelenmesi (örn. Istio/Service Mesh), sır yönetiminin HashiCorp Vault ile yapılması, log kayıtlarının güvenli yönetimi ve eğer kullanıcı verisiyle etkileşim varsa (örneğin gerçek zamanlı izleyici verisi alan bir interaktif iş) bu verilerin otomatik PII maskeleme/anonimleştirme araçlarıyla (örn. Microsoft Presidio) işlenmesi sağlanır. Kişisel veri senaryolarında diferansiyel mahremiyet gerektiren eğitim/çıkarım akışlarında Opacus veya TensorFlow Privacy gibi kütüphaneler devreye alınabilir. ISO/IEC 27701 gibi standartlar, servis aşamasında kişisel verilerin gizliliğinin korunması için uyulması gereken ilke ve kontrolleri tanımlar; uzman, bu gerekliliklere uyarak hem kurumun hem de işbirliği yapılan arşivlerin veri kullanım politikalarına uygun hareket eder. Sistemin operasyonu süresince (örneğin sergi açıkken) uzman performansı izler, herhangi bir hata, yavaşlık veya beklenmedik *output* oluşursa müdahale eder. Gözlemlenebilirlik için Prometheus ile metriklerin toplanması, Grafana ile panellerin oluşturulması, OpenTelemetry ile uçtan uca iz (trace)/metrik/log telemetrisi üretilmesi; deney ve model sürümlerinin MLflow Model Registry veya Weights & Biases Registry ile takip edilmesi; veri ve çıktı kalitesinin Evidently panelleriyle sahada raporlanması önerilir. Trafik yönetiminde canary/blue-green uygulamalarla yüzdeli yönlendirme ve gölge trafik (shadow) desenleri kullanılarak güvenli geçişler sağlanır; KServe ve servis mesh (Istio) ile yüzdesel trafik bölme yerel olarak desteklenir. YZ sistemlerinin dinamik yapısı gereği, dağıtım sonrası da modelin periyodik izlenmesi ve gerektiğinde güncellenmesi önemlidir. Nitekim yeni yayınlanan ISO/IEC 5338 standardı, YZ sistemlerinde "sürekli doğrulama" kavramını tanıtarak modelin hizmete alındıktan sonraki performansının zaman içinde düzenli aralıklarla test edilmesini önermektedir; pratikte bu, Evidently ile drift alarmları ve SLA/SLO eşikleri, MLflow/W&B ile model versiyon atlamalarında otomatik geri

alma (rollback) koşulları, KServe canary ile kademeli geçiş adımları ve ONNX Runtime/TensorRT tabanlı performans regresyon testleriyle gerçekleştirilir. Bu yaklaşım sayesinde veri veya konsept kaymalarının (drift) tespiti ve giderilmesi, güvenlik açıklarının kapanması mümkün olacaktır. Sonuç olarak uzmanın yaşam döngüsündeki rolü, modelin ilk fikirden itibaren güvenli, etik ve amaca uygun biçimde geliştirilip kullanıma sunulması ve kullanım süresince güvenilirliğinin korunması için gerekli tüm teknik ve yönetsel adımları üstlenmektir.

5.5.2. Yöntemsel İyileştirmeler ve Standartlara Uyumluluk İhtiyacı

YZ projelerinde kullanılan mevcut çalışma yönteminin sürekli güncellenmesi gerekir. Teknolojinin hızla ilerlemesiyle birlikte, Refik Anadol Studio'da yenilikçi ortamlar en son çıkan modelleri ve araçları projelerine entegre etmektedir. Ancak yalnızca güncel modeli kullanmak yeterli değildir; verinin yönetimi, modelin geliştirilmesi ve sistemin işletimine dair yöntemlerin de gelişen en iyi uygulamalar ışığında iyileştirilmesi şarttır. Mevcut çalışma yönteminde güncellenmesi gereken başlıkların başında belgelendirme ve izlenebilirlik gelir. Büyük ölçekli veriyle çalışırken, verinin ne şekilde toplandığı, işlemlerden nasıl geçirildiği, modelin hangi versiyonlar ile eğitildiği gibi bilgiler çoğu zaman proje baskısı altında ihmal edilebilmektedir. Oysaki standartlar, bu konulara özel önem atfeder. Örneğin yeni YZ yaşam döngüsü standardı, her adımda sürekli risk yönetimi uygulanmasını ve tüm süreçlerin kayıt altına alınmasını şart koşar. Bu, YZ projelerinin doğasındaki belirsizlikler ve değişken riskler (ör. veri dağılımındaki değişimler, saldırı girişimleri, performans tutarsızlıkları) karşısında proaktif olabilmek için gereklidir. Uzmanın çalışma yöntemine entegre etmesi gereken bir diğer iyileştirme kalite güvence metriklerinin genişletilmesidir. Geleneksel yazılım projelerinde başarı çoğunlukla teknik performans (hız, doğruluk vb.) veya yatırım getirisi ile ölçülürken, YZ tabanlı sanat projelerinde güvenilirlik, açıklanabilirlik, tarafsızlık, kullanıcı güveni, şeffaflık, gizlilik ve toplumsal etki gibi daha soyut ancak kritik kalite kriterleri de önem kazanır. Uluslararası çerçeveler, YZ sistemlerinin bu çok boyutlu kalite unsurlarına göre değerlendirilmesini önermektedir. Dolayısıyla, stüdyo projelerinde yalnızca "model çalışıyor mu?" sorusuna değil, "model çıktıları güvenilir mi, önyargı içeriyor mu, izleyicide istenen etkiyi yaratıyor mu?" sorularına da yanıt arayan yöntemler benimsenmelidir. Bu amaçla uzmanın, proje başlangıcında bu kriterleri başarı ölçütlerine dönüştürüp sürece dâhil etmesi, gerekirse her kriter için ayrı test ve doğrulama adımları planlaması gerekecektir.

Mevcut yöntemlerin güncellenmesi gereken bir diğer alan, güvenlik ve gizlilik uygulamalarıdır. Veri gizliliğinin kritik olduğu projelerde dahi, pratikte bazen hız ve esneklik

uğruna güvenlik protokollerinin gevşetilebildiği görülmektedir. Özellikle işbirliği yapılan kurumlar veya şirketler, YZ standartlarına ve sertifikasyonlarına pek aşina olmadığında, proje takvimine uyabilmek adına süreçlerin kısaltılmasını ya da bazı resmi prosedürlerin esnetilmesini talep edebilirler. Bu durumda uzman, bir yandan paydaşları uluslararası standartlara uygunluğu korumanın uzun vadede güven ve kalite açısından değerine ikna etmeli, bir yandan da fiiliyatta gerekli esnekliği sağlayacak yaratıcı çözümler geliştirmelidir. Örneğin, bir şirket ayrıntılı veri kalite ölçümleri veya risk analizleri için zaman ayırmak istemiyorsa, uzman en kritik riskleri hızla değerlendirebileceği atölye çalışmaları düzenleyerek veya küçük prototip testlerle olası sorunları erkenden ortaya koyarak standartların ruhunu korumaya çalışabilir. Nihayetinde, yöntemsel iyileştirmelerin kurumsal bir çerçeveye oturtulması için standartlara ihtiyaç duyulmaktadır. Standartlar, deneyime dayalı birer "kolektif en iyi uygulama rehberi" sundukları için, projelerin sadece teknik açıdan değil, etik ve yönetişim boyutlarında da tutarlı ve güvenilir olmasını sağlarlar. Ayrıca standartlara uyum, farklı ekipler ve disiplinler arasında ortak bir dil ve beklenti seti oluşturur. Bu durum, özellikle disiplinler arası ve çok paydaşlı sanat projelerinde iletişim hatalarını azaltır, verimsizlikleri giderir ve işbirliğini kolaylaştırır. Örneğin ISO/IEC 5338 gibi yaşam döngüsü standartları sayesinde farklı kurum ve ülkelerden paydaşlar, projeye aynı referans noktası üzerinden yaklaşabilir, böylece engeller azalırken yenilikçi teknolojilerin farkındalığı ve kabulü artar. Kısacası, belirli bir standarda bağlı kalarak çalışmak, uzmanın yöntemini sistematikleştirip iyileştirirken; projenin güvenlik, etik ve kalite hedeflerini de somut kontrol noktalarına bağlamaya imkân tanır.

5.5.3. Standartlar, İşbirlikleri ve Uygulama Yöntemleri

YZ yaşam döngüsü kapsamında, özellikle ISO rehberlikleri ışığında, projede uyulması veya rehber alınması gereken çok sayıda standart bulunmaktadır. Bunların bir kısmı doğrudan YZ sistemlerinin kendisine özgü iken, bir kısmı da daha genel bilgi teknolojisi, güvenlik ve kalite yönetimi standartlarıdır. Uzmanın kendi alanına özgü veya genel geçer uygunluk sağlaması gereken başlıca standartlar ve bunların kullanım amaçları aşağıda özetlenmiştir:

1. ***YZ Yaşam Döngüsü ve Yönetişim Standartları***

ISO/IEC 5338:2023 (YZ sistem yaşam döngüsü süreçleri), YZ proje süreçlerinin uçtan uca ele alınmasını sağlayan temel standarttır. Fikir, geliştirme, doğrulama, bakım gibi adımların tanımlarını yapar ve her aşamaya güvenlik, etik ve insan faktörü boyutlarının entegre edilmesini önerir. ISO/IEC 42001:2023 (YZ için yönetim sistemi) ise organizasyonların YZ'yi kurumsal düzeyde yönetmesi için bir çerçeve sunar; bu standart, risk yönetimi, kaynak yönetimi ve

sürekli iyileştirme prensiplerini YZ projelerine uyarlayarak şirketlerin YZ alanında kalite güvencesi sağlamasına yardımcı olur. Benzer şekilde ISO/IEC 38507:2022 (YZ kullanımının yönetişim ilkeleri), üst yönetimin ve yöneticilerin YZ projelerinde hesap verebilirlik ve stratejik hizalama sağlaması için kılavuzluk eder.

2. ***Veri Kalitesi ve Veri Yönetimi Standartları:***

YZ modelinin başarısı büyük ölçüde verinin kalitesine bağlıdır. Bu nedenle ISO/IEC 5259 serisi (YZ için veri kalitesi) ve ISO/IEC 25024:2015 (veri kalitesi ölçümü) gibi standartlar, eğitilecek veri setlerinin doğruluk, bütünlük, tutarlılık gibi metriklerle değerlendirilmesine yönelik yöntemler sunar. Uzman, bu standartlar doğrultusunda arşiv verilerinin proje amaçlarına uygunluğunu ölçer ve iyileştirmeler yapar. ISO/IEC 25059:2023 (YZ sistemleri için kalite modeli) ise YZ ürün ve hizmetlerinin kalite özelliklerini tanımlar; güvenilirlik, kullanılabilirlik, bakıma açıklık gibi klasik kalite ölçütlerinin yanı sıra açıklanabilirlik, önyargısızlık, güvenlik gibi YZ'ye özgü alt özellikleri de modele dahil eder. Uzman, bu kalite modelini referans alarak projenin başarı kriterlerini geniş bir perspektifte belirler.

3. ***Güvenlik ve Gizlilik Standartları:***

YZ projelerinde bilgi güvenliği ve gizlilik, sadece kurum içi değil paydaşlarla ilişkiler açısından da kritik bir uyum alanıdır. ISO/IEC 27001:2022 (Bilgi güvenliği yönetim sistemi), proje boyunca uygulanacak güvenlik kontrollerinin çerçevesini çizer; veri erişim yetkilerinin kısıtlanması, şifreleme kullanımı, güvenlik risk analizleri gibi konularda gereksinimler ortaya koyar. Uzman, özellikle hassas arşiv verilerinin işlendiği durumlarda bu kontrollerin uygulandığından emin olur. ISO/IEC 27701:2019 (Gizlilik bilgi yönetimi), 27001'in özelinde kişisel verilerin gizliliğini korumaya odaklanmış bir uzantısıdır ve projede eğer herhangi bir kişisel veri veya gizlilik içeren içerik söz konusuysa devreye girer. Ayrıca ISO/IEC 29100:2011 (Gizlilik çerçevesi), bireysel verilerin işlenmesinde uyulması gereken temel prensipleri (rıza, veri minimizasyonu, amaç sınırlaması vb.) tanımlayarak uzmana hukuki ve etik açıdan yol gösterir. Bu standartlar sayesinde, stüdyo projelerinde verinin gizliliği konusunda uluslararası kabul görmüş bir güvence seviyesi sağlanmış olur.

4. ***Risk, Etik ve Toplumsal Etki Standartları:***

ISO/IEC 23894:2023 (YZ için risk yönetimi rehberi), YZ projelerinde ortaya çıkabilecek teknik, etik veya hukukî risklerin sistematik biçimde yönetilmesine yardımcı olur. Uzman, bu rehbere uygun şekilde riskleri tanımlar, etkilerini değerlendirir ve önleyici/azaltıcı

tedbirler planlar. Özellikle sanat projelerinde toplumsal etki öngörüsü zor olabilir; bu nedenle standardın önerdiği yapısal risk değerlendirmeleri sürpriz olumsuz etkileri önceden fark etmeyi kolaylaştırır. ISO/IEC TR 24368:2022 (YZ ile ilgili etik ve toplumsal kaygılar raporu), YZ'nin toplum üzerindeki etkileri, etik ilkeler (adalet, şeffaflık, hesap verebilirlik vb.) konusunda kapsamlı bir kılavuz sunar. Uzman, proje özelinde bu ilkelerin hangilerinin öne çıktığını belirleyip tasarım kararlarında bunları gözetir. Örneğin, veri heykeli projesi toplumda yanlış anlaşılabilecek bir tarihî konuyu ele alıyorsa, önyargı ve ayrımcılık risklerine karşı özellikle duyarlı olunmalıdır. Son olarak, ISO/IEC TS 4213:2022 (ML sınıflandırma performansı değerlendirmesi) ve benzeri teknik standartlar, modelin performansının nasıl ölçülüp raporlanacağına dair araçlar sunar. Bu, uzmanın modelin sonuçlarını nesnel ve standart bir biçimde değerlendirebilmesini ve gerektiğinde üçüncü taraflara iletişiminde ortak bir dil kullanabilmesini sağlar.

Yukarıda belirtilen standartlar ışığında, projede disiplinler arası işbirliği ihtiyacı belirgin hale gelmektedir. Standartlar, YZ yaşam döngüsünün her adımında farklı uzmanlıkların katkısını vurgular. Örneğin, ISO 9241-210 kullanıcı merkezli tasarım yaklaşımı gereği kullanım bağlamı analizinde son kullanıcıların ve alan uzmanlarının dahil edilmesini önerir. Benzer şekilde ISO/IEC 5338 ve ISO/IEC 23894, gerek sistem gereksinimlerinin belirlenmesi gerek risk değerlendirmesi safhalarında paydaş katılımının önemine dikkat çeker. Bu doğrultuda, YZ sanat projesinin bağlam ve veri toplama sürecinde tarih, biyoloji, coğrafya gibi alan uzmanlarıyla işbirliği yapılması zorunlu hale gelir. Uzman, verinin doğru anlaşılması ve kullanılabilmesi için gerektiğinde akademisyenler, araştırmacılar veya ilgili alanın profesyonellerinden danışmanlık alır (örneğin tıbbi bir veriye dayalı projede bir doktorun görüşlerini almak, müzikal bir veri setiyle çalışırken bir müzikologla çalışmak gibi). Model geliştirme ve doğrulama aşamalarında, özellikle etik ve güvenlik konularında farklı uzmanlıklarla birlikte hareket edilir: Modelin önyargı içermemesi veya telif hakkı ihlali yapmaması için hukukçular ve etik uzmanları sürece dahil olur; veri güvenliği için siber güvenlik uzmanları alt yapıyı inceler; kullanıcı deneyimi odaklı interaktif projelerde UX tasarımcıları ve insan-bilgisayar etkileşimi uzmanları, sistemin izleyiciyle etkileşimini iyileştirmek üzere geri bildirim verir. Dağıtım ve işletme aşamasında ise işin sürekliliğini sağlamak ve olası kriz senaryolarına hazırlıklı olmak adına BT operasyon ekipleriyle ve gerektiğinde acil durum planlama uzmanlarıyla (özellikle kamusal alanda sergilenen bir işin beklenmedik durumlarını yönetmek için) iletişim kurulması gerekebilir. Görüldüğü üzere, standartların çizdiği ideal süreç haritasında tek bir disiplinin tüm gereksinimleri karşılaması

beklenmez; tam tersine, her süreç adımı, kendi ihtisas alanına sahip profesyonellerin katkısıyla zenginleşir. Uzmanın rolü de bu noktada bir orkestra yöneticisi gibi, farklı uzmanlıklar arasında köprü kurmak, ortak bir dil oluşturmak ve standartların öngördüğü kriterlerin bütüncül bir biçimde sağlandığından emin olmaktır.

Son olarak, standartların önerdiği yöntem ve araçlardan bahsetmek gerekirse; bu standartlar genellikle süreç odaklı ve araç-agnostik tavsiyeler sunarlar. Örneğin ISO/IEC 5338 ve ISO/IEC 42001, YZ projelerinde sürekli iyileştirme döngüsünü temel alan bir yönetim sistemi kurulmasını önerir. Bu, Planla-Uygula-Kontrol Et-Önlem Al (PUKÖ) döngüsünün YZ yaşam döngüsüne entegre edilmesi demektir. Bu kapsamda standartlar, risk değerlendirmesinin belirli aralıklarla tekrarlanması, model performansının düzenli izlenmesi ve sistem davranışındaki değişimlerin kayıt altına alınarak gerekli ayarlamaların yapılması için protokoller tanımlar. Örneğin, sürekli doğrulama protokolü gereği uzman, her üç ayda bir modeli güncel veriyle teste tabi tutup sonuçları kıyaslayabilir ve bir rapor halinde belgeleyebilir. Yine standartlar, araç seçiminde bazı prensipler verir: Veri yönetimi için şeffaflık ve izlenebilirlik sağlayan araçların (versiyon kontrol sistemleri, veri soy ağacı izleme yazılımları gibi) kullanılması; model geliştirme için güvenli programlama çerçeveleri ve kütüphanelerin tercih edilmesi; çıktıların değerlendirilmesi için standart metriklerin ve tarafsız test setlerinin kullanılması gibi yaklaşımlar tavsiye edilir. Örneğin, veri gizliliği için eşyapılı (*homomorfik)* şifreleme veya anonimleştirme araçlarının kullanımı, güvenlik için SAST/DAST gibi yazılım güvenliği test araçlarının YZ kodlarına uygulanması öneriler arasında sayılabilir. Bazı standartlar sektörel en iyi uygulamalara da referans verir: NIST'in YZ Risk Yönetim Çerçevesi veya ISO 31000 risk yönetimi standartları, YZ proje risklerini yönetmek için uygulanabilecek metodolojiler olarak değerlendirilebilir. Ayrıca, şeffaflık ve hesap verebilirlik için proje boyunca bir "model bilgi formu" (model card) tutulması, veri setleri için "veri belgesi" (datasheet) hazırlanması gibi pratikler de bugün birçok YZ standardı ve rehberinde dolaylı olarak teşvik edilmektedir. Bu tür araçlar, YZ sisteminin ne yaptığı ne kadar güvenilir olduğu ve hangi koşullarda çalıştırıldığı konusunda paydaşlara ve son kullanıcılara açıklama sunarak projenin güvenilirliğini artırır.

5.6. Yapay Zekâ Standartları: Teknik Perspektiften İnceleme

Dr. Merve Ayyüce Kızrak

Kurucu & CEO, HUX AI, ayyucekizrak@gmail.com

5.6.1. ISO/IEC TS 4213:2022 Teknik Raporu; Makine Öğrenmesi Sınıflandırma Performansının Değerlendirilmesi ve Önyargının Azaltılması

ISO/IEC TS 4213:2022 standardı, YZ alanında makine öğrenmesi kullanan sınıflandırma ve regresyon sistemlerindeki istenmeyen önyargıyı ele almayı amaçlayan önemli bir teknik dokümandır. Bu standart, YZ sistemlerinin yaşam döngüsü boyunca ortaya çıkabilecek önyargıları tespit etmek ve azaltmak için izlenmesi gereken yöntemleri tanımlar. Geliştirilen standart, *Birleşmiş Milletler* benzeri bağımsız kuruluşlar ve endüstri tarafından desteklenen, adil ve sorumlu YZ uygulamalarını teşvik etmeyi hedefleyen bir çerçeve sunar. Bu raporda, söz konusu standardın kapsamı, getirdiği kavramlar ve önerdiği en iyi uygulamalar anlaşılır bir dille açıklanmaktadır.

Standardın Amacı ve Kapsamı

ISO/IEC TS 4213:2022, makine öğrenmesi tabanlı sınıflandırma ve regresyon görevlerinde istenmeyen önyargıların azaltılması için yol gösterici ilkeleri ortaya koyar. Standart, farklı büyüklük ve türdeki tüm kuruluşlar için geçerli olacak şekilde genel bir çerçeve sunmaktadır. Temel hedefi, YZ sistemlerinin geliştirme veya kullanım aşamalarında ortaya çıkabilecek önyargıları "tedavi etmek" yani etkilerini azaltmak veya ortadan kaldırmak için uygulanabilir adımlar tanımlamaktır. Bu kapsamda standardın ana amaçları şöyle özetlenebilir:

- **Yaşam Döngüsü Boyunca Önyargı Azaltma:** YZ sistem yaşam döngüsünün her aşamasında (planlama, veri toplama, model geliştirme, değerlendirme, devreye alma ve izleme gibi) önyargı risklerini azaltmaya yönelik teknikleri tanımlamak. Standart, önyargıyla mücadelenin sistem geliştirme sürecinin en başından sonuna dek sürmesi gereken sürekli bir çaba olduğunu vurgular.
- **Uluslararası Boşluğu Doldurma**: Bu standarttan önce, YZ sistemlerindeki önyargıların ele alınmasına dair uluslararası düzeyde detaylı bir rehber

bulunmamaktaydı. ISO/IEC TS 4213, bu boşluğu doldurarak geliştirme ve kullanım sırasında alınabilecek uygun önlemleri tanımlar ve sektöre yol gösterir.

- **Genel Geçer Teknikler:** Standartta sunulan önyargı azaltma teknikleri, belirli bir uygulama alanına (örneğin yalnızca *etiket adaleti* veya *ayrımcılık* konusuna) özgü değildir; aksine, bağlamdan bağımsız genel iyi uygulamalar niteliğindedir. Böylece farklı sektörlerde (sağlık, finans, eğitim vb.) ve farklı ölçeklerdeki YZ projelerinde uygulanabilecek genel ilkeler sunulur.
- **Düzenlemelere Uyum:** Standart, kuruluşların çeşitli bölgesel ve ulusal yasal düzenlemelere uyum sağlamasına da yardımcı olmayı hedefler. Özellikle YZ'de adillik ve önyargı konusunda çıkmakta olan mevzuat ve yönetmeliklere (örneğin, kişiselleştirilmiş hizmetlerde ayrımcılığı önleme yasaları vb.) uyumluluk için iyi bir rehber niteliğindedir.

İstenmeyen Önyargı Nedir?

YZ sistemlerinde önyargı (bias), modelin sistematik olarak belirli bir yönde hatalı veya yanlı sonuçlar üretmesine yol açan hatalı varsayım, veri veya yöntemsel tutarsızlıkları ifade eder.

Standart, *istenmeyen önyargı* kavramını vurgulayarak, modelin amaçlanmayan ve istenmeyen şekilde belirli grupları veya sonuçları kayırması durumunu tanımlar. Burada önemli bir ayrım vardır: Her önyargı tamamen kötü değildir; bir YZ sisteminin hedeflerine ulaşması için kasıtlı olarak belirli bir eğilim kazandırılması gerekebilir (istenen önyargı). Ancak, sistemin amacı dışında ortaya çıkan ve geliştiricilerin arzu etmediği önyargılar istenmeyen önyargı olarak tanımlanır. Bu tür önyargılar, genellikle toplumdaki adaletsizlikleri yansıtır veya pekiştirir ve sistemin çıktılarını haksız hale getirebilir.

Önyargının Kaynakları: ISO/IEC TS 4213 standardı, ISO/IEC TR 24027 teknik raporundaki sınıflandırmaya dayanarak önyargı kaynaklarını üç ana başlık altında ele alır:

- **İnsani (Bilişsel) Önyargılar:** Geliştiricilerin, veri etiketleyicilerin veya karar vericilerin farkında olmadan kendi ön kabulleri veya kalıp yargıları nedeniyle sisteme yansıyan önyargılardır. Örneğin, eğitim verilerini derleyen kişinin dünya görüşü veya tercihlerinin veri setine etki etmesi bu kapsamdadır. İnsani önyargılar, veri seçimi ve sınıflandırma etiketlemesinde hatalı yönlendirmelere yol açabilir.

- **Veri Kaynaklı Önyargılar:** Veri önyargıları, modelin eğitildiği veya test edildiği veri kümesindeki sistematik hatalardan kaynaklanır.

Örneğin, eğitim verilerinin toplumun belli kesimlerini yeterince temsil etmemesi (örnekleme önyargısı) modelin bu kesimler üzerinde kötü performans sergilemesine yol açabilir. Tarihsel verilerde var olan tarihsel önyargılar da (örneğin geçmişte insan kaynakları seçimlerinde kadın adaylara karşı ayrımcılık yapıldıysa, bu veriden öğrenen modelin de benzer ayrımcı sonuçlar üretmesi) bu kategoriye girer. Ayrıca ölçüm ve toplama önyargıları (verinin toplandığı cihaz veya yöntemden kaynaklanan sapmalar) ve etiketleme önyargıları (veri etiketleme sürecindeki tutarsızlıklar veya hatalar) da veri kaynaklı önyargı türleridir.

- **Mühendislik Kararlarından Kaynaklanan Önyargılar:** Bu önyargılar, sistemin tasarımında ve geliştirilmesinde alınan teknik kararlar nedeniyle ortaya çıkar.

Örneğin, modelin hangi özellikleri (feature) kullanacağına dair kararlar veya üst parametre ayarları sistem performansını belli bir gruba avantaj/dezavantaj sağlayacak şekilde etkileyebilir. Benzer şekilde, veriyi ön işleme biçiminiz bazı algoritmaları diğerlerine kıyasla avantajlı kılabilir ve bu da sonuçların adil olmayan şekilde kıyaslanmasına yol açabilir. ISO/IEC TR 24027 raporu bu tür önyargılara da dikkat çekerek, mühendislik aşamasındaki tercihlerin bile önyargı doğurabileceğini vurgulamıştır.

İstenmeyen önyargılar, farkına varılmadığı takdirde YZ destekli karar verme sistemlerinde ciddi adaletsizliklere yol açabilir.

Örneğin, bir makine öğrenmesi modeli eğitim verisindeki dengesizlik yüzünden siyahi hastalarda bir hastalığı teşhis etmede daha başarısız olursa, sağlık hizmetlerinde eşitsizlik doğabilir. Bu nedenle, önyargı kaynaklarının iyi anlaşılması ve her birine yönelik ayrı önlemler geliştirilmesi kritik önem taşır.

Yapay Zekâ Yaşam Döngüsü Boyunca Önyargı

Standart, istenmeyen önyargıyla mücadelenin tek seferlik bir düzeltme değil, yaşam döngüsü boyunca devam eden bir süreç olduğunu vurgulamaktadır. Bir YZ sisteminin yaşam döngüsü; planlama ve tasarımdan başlayıp veri toplama, model geliştirme, değerlendirme, sahaya sürme (deployment) ve izleme/audit aşamalarına kadar uzanan bir dizi aşamayı içerir. ISO/IEC TS 4213, her bir aşamada alınabilecek önyargı azaltıcı önlemleri tanımlar ve bu

uygulamaları sistem geliştirme sürecine entegre etmeyi önerir. Aşağıda, YZ sistem yaşam döngüsünün kritik aşamalarında önyargıyla mücadele için dikkat edilmesi gereken noktalar sıralanmıştır:

- **Planlama ve Tasarım:** Proje başlangıcında, sistemden beklenen adillik kriterlerinin tanımlanması gerekir. Bu aşamada ekip, başarım metriği olarak yalnızca doğruluk değil, aynı zamanda farklı gruplar arası adalet ölçütlerini de hedef olarak belirlemelidir.

Örneğin, işe alım yapan bir YZ sisteminde "erkek ve kadın adaylar için seçilme oranlarının benzer olması" gibi bir kriter baştan konulabilir. Standart, önyargı risklerinin erken değerlendirilmesini ve gereksinimlerin buna göre şekillendirilmesini tavsiye eder.

- **Veri Toplama ve Hazırlama:** Verinin toplanmasında ve ön işlenmesinde özen gösterilmesi en kritik konulardan biridir. İlk olarak, kullanılan veri kümesi hedeflenen popülasyonu temsil edecek şekilde çeşitli ve dengeli olmalıdır.

Örneğin, bir yüz tanıma sistemi geliştiriliyorsa, eğitim verisi farklı etnik köken, yaş ve cinsiyet gruplarından yeterli örnek içermelidir. İkinci olarak, veri üzerinde yapılacak ön işlemler (preprocessing) dikkatle ele alınmalıdır. Veriyi temizleme veya dönüştürme adımları, bazı hassas bilgileri çıkarıp çıkarmama kararları ileride model çıktılarında önyargıya yol açabilir.

Standart, veri ön işlemede alınacak kararların önyargı açısından etkisini değerlendirmeyi önerir.

Örneğin, veri kümelerindeki dengesizlikleri gidermek için alt örnekleme/üst örnekleme teknikleri uygulamak veya sentetik veri üretimi (data augmentation) ile az temsil edilen grupların verisini artırmak faydalı olabilir. Ayrıca, verideki aykırı değerlerin veya hatalı etiketlerin düzeltilmesi hem model performansını artırır hem de önyargı ihtimalini azaltır.

- **Model Geliştirme (Eğitim) Aşaması:** Modeli eğitirken kullanılan algoritmalar ve ayarlar da önyargıya etki edebilir. Standart, geliştiricilerin önyargı farkındalığıyla algoritmik seçimler yapmasını önerir.

Örneğin, fairness (adalet) kriterlerini göz önüne alan veya belirli demografik gruplardaki hataları cezalandıran özel kayıp fonksiyonları kullanılabilir. Algoritmik ayarlamalar olarak adlandırılan bu tür teknikler, modelin eğitim sürecine doğrudan entegre edilir.

Literatürde “fairness-aware” (adalet farkındalığına sahip) makine öğrenmesi yöntemleri olarak geçen bu yaklaşımlar, modelin belirli bir grup aleyhine daha yüksek hata yapmasını engellemeye çalışır.

Örneğin, bir kredi skorlaması modelinde, farklı etnik gruplar arasında benzer onay oranları elde etmeyi hedefleyen düzenleyiciler (regularizer) eklemek bu kapsamdadır. Bunun yanı sıra, üst parametre optimizasyonu sürecinde de önyargı metriklerini göz önünde bulundurmak gerekebilir; zira sadece genel doğruluğu maksimize eden bir üst parametre seçimi, azınlık gruplardaki performansı düşürebilir (bu durum bir çeşit üst parametre önyargısı yaratabilir).

- Model Değerlendirme ve Test: Bir modelin adilliğini anlamak için değerlendirme aşamasında sadece genel başarı ölçütlerine bakmak yeterli değildir. Standart, model performansının farklı alt gruplar için ayrı ayrı incelenmesini ve adalet metrikleri kullanılarak değerlendirilmesini önerir.

Örneğin, “hassas özellik” (cinsiyet, ırk gibi) bazında ayrı hata oranları hesaplamak, modelin herhangi bir grup için sistematik olarak daha kötü performans gösterip göstermediğini ortaya koyar.

Yaygın adalet ölçütlerinden bazıları,

- Eşit Hata Oranı (iki grup için pozitif tahminlerin yanlışlık oranlarının eşit olması),
- Demografik Parite (modelin pozitif karar verme oranının gruplar arasında benzer olması) ve
- Eşit Fırsat (gerçek olumlu örnekler için modelin pozitif tahmin oranlarının eşit olması) gibi metriklerdir.

ISO/IEC TS 4213, bu tür metriklerin kullanımını teşvik eder ve önyargının ölçülmesi için istatistiksel tekniklerin uygulanmasını önerir.

Örneğin, bir sağlık teşhis modelinde hem kadın hem erkek hastalar için duyarlılık (recall) değerlerini ayrı hesaplayıp kıyaslamak, cinsiyete dayalı bir performans farkı olup olmadığını gösterecektir. Eğer model bir grup için belirgin derecede düşük performans veriyorsa, bu bir istenmeyen önyargı göstergesidir.

- **Karar ve Çıktı Aşaması (Post-Processing):** Model eğitilip kullanıma hazır hale geldikten sonra bile, üretilecek çıktıların adil olup olmadığını denetlemek ve

gerekiyorsa düzeltmek önemlidir. Standart, çıktı sonrası işlem (post-processing) tekniklerinin kullanılabileceğini belirtir. Bu yaklaşımlar, modelin orijinal çıktısını değiştirerek adaleti sağlamaya çalışır; örneğin, belirli bir gruba karşı sistematik bir olumsuzluk varsa sonuçlar yeniden ölçeklenebilir veya karar eşikleri gruplara özel ayarlanabilir. Basit bir örnek vermek gerekirse, işe alım skoru üreten bir model erkek adaylara kadınlara kıyasla sürekli daha yüksek skor veriyorsa, post-processing aşamasında kadın adaylar için belirli bir puan iyileştirmesi veya kota sistemi uygulanabilir (elbette bu, uygulama politikasına bağlı bir karardır). Bu tür düzeltme teknikleri, model çıktılarını adalet ilkelerine göre kalibre etme imkânı sunar. Ancak post-processing yöntemlerinin dikkatli kullanılması gerekir; çünkü aşırı düzeltme yapılırsa model performansı düşebilir veya yeni türde haksızlıklar doğabilir. Standart, mümkün olduğunca sorunu kaynağında (veride veya modelde) çözmenin, son aşamada palyatif düzeltmelere tercih edilmesini vurgular.

- **Devreye Alma ve İzleme:** Bir YZ sistemi saha ortamında çalışmaya başladıktan sonra da önyargı riskleri tamamen ortadan kalkmış olmaz. Gerçek dünya verisi ve kullanım koşulları, zamanla model davranışını değiştirebilir. Bu nedenle ISO/IEC TS 4213, model dağıtıma alındıktan sonra düzenli izleme ve bakım yapılmasını önermektedir.

Örneğin, modelin gerçek kullanımda topladığı yeni veriler, başlangıçta olmayan bir önyargıyı ortaya çıkarabilir (veri dağılımında kayma veya kullanıcı etkileşimlerinden doğan yeni önyargılar). Kuruluşlar, modelin çıktılarının sürekli olarak denetlenmesi, belirli gruplarda anormal birikimler olup olmadığının takip edilmesi ve gerektiğinde modele müdahale edilmesi konularında hazırlıklı olmalıdır.

Standart, bu izleme sürecini yaşam döngüsünün bir parçası olarak görür ve sürekli geribildirim döngüsü kurulmasını tavsiye eder. Yani, izleme sırasında tespit edilen önyargılar veya hatalar, tekrar geliştirme sürecine geri beslenerek modelin güncellenmesi veya yeniden eğitilmesi sağlanmalıdır. Bu döngüsel yaklaşım, YZ sisteminin öğrenen ve iyileşen bir sistem olmasını, dolayısıyla adalet açısından zamanla daha iyi hale gelmesini hedefler.

Yukarıdaki adımlardan anlaşılacağı üzere, standart yalnızca teknik model geliştirme konularını değil, aynı zamanda **işletimsel süreçleri** de kapsayan bütüncül bir yaklaşım sunmaktadır. Sonraki bölümlerde, standardın öngördüğü belli başlı önyargı azaltma teknikleri ve kurumsal önlemler daha detaylı incelenecektir.

Önyargının Tespiti ve Ölçülmesi

Bir sorunu çözebilmenin ilk adımı, onu tespit etmek ve ölçmekten geçer. ISO/IEC TS 4213 standardı, YZ sistemlerindeki istenmeyen önyargıyı belirleyebilmek için çeşitli yöntemler ve ölçütler sunar. Bu bölümde, önyargıyı görünür kılmak amacıyla kullanılabilecek yaklaşımlar ele alınmaktadır.

- **Veri ve Model İncelemesi:** Standart, geliştiricilerin ellerindeki veri kümesini ve model davranışını sistematik bir biçimde gözden geçirmelerini önerir. İlk olarak, veri incelemesi ile başlanmalıdır: Eğitim ve test verileri, demografik özellikler veya diğer hassas değişkenler açısından analiz edilerek bariz dengesizlikler veya eksiklikler tespit edilmelidir. Örneğin, bir kredi başvuru verisinde başvuranların gelir dağılımı, cinsiyet oranı veya bölgesel dağılımı incelenebilir; eğer veride belirgin bir eksiklik (örneğin çok az kadın başvuru olması gibi) gözlenirse, bu durum potansiyel bir önyargı kaynağı olabilir. Aynı şekilde, modelin yaptığı tahminler de çıktı analizi ile incelenmelidir: Model belirli bir kategorideki (ör. belirli bir etnik gruba ait) girdilere tutarlı bir şekilde olumsuz sonuç mu veriyor? Ya da belirli bir segmentte aşırı iyimser mi? Bu tür sorular, modelin önyargılı davranıp davranmadığını anlamak için kritiktir.
- **İstatistiksel Analiz ve Hipotez Testleri:** Önyargının ölçülmesinde istatistik önemli bir araçtır. Standarda göre, farklı gruplar arasındaki sonuç farklarının anlamlı olup olmadığını anlamak için istatistiksel testler kullanılabilir.

Örneğin, erkek ve kadın müşterilere kredi veren bir model düşünelim; modelin erkeklere kredi verme oranı %60, kadınlara %40 ise aradaki %20 fark rastlantısal mı yoksa sistematik mi? Bunu anlamak için khi-kare testi veya z-testi gibi hipotez testleri uygulanabilir. Eğer test sonucu, cinsiyet ile kredi kararı arasında istatistiksel olarak anlamlı bir ilişki olduğunu gösterirse, modelde cinsiyete dayalı istenmeyen bir önyargı olduğundan şüphelenmek gerekir.

Benzer şekilde, McNemar testi, t-test gibi yöntemler de iki modelin farklı gruplardaki performanslarını karşılaştırmak için kullanılabilir. Standart, bu tür tekniklerin önyargı tespitinde kullanılabileceğine işaret ederek geliştiricileri sayısal kanıtlar aramaya teşvik eder.

- **Adalet (Fairness) Metrikleri:** Günümüzde literatürde, makine öğrenmesi modellerinin adilliğini nicel olarak ifade etmek için birçok metrik tanımlanmıştır. ISO/IEC TS 4213, bu adalet metriklerinin uygun şekilde kullanılmasını tavsiye etmektedir. Bazı yaygın metriklerden yukarıda bahsettik (demografik parite, eşit fırsat gibi). Bir diğer örnek,

Kapsayıcı Hata Oranı (False Negative/False Positive Rate Balance) olabilir; bu metrik, modelin yanlış negatif veya yanlış pozitif hata oranlarının gruplar arasında ne derece farklılaştığını ölçer. Eğer bir yüz tanıma sistemi, belli bir etnik grupta diğerine göre çok daha fazla *false negative* (yanlış tanımama) yapıyorsa, bu o grup için adaletsiz bir durum yaratır. Bu farkın sayısal olarak ölçülmesi, problemin ciddiyetini ortaya koyar. Standart, uygun metriklerin seçilerek düzenli aralıklarla hesaplanmasını ve raporlanmasını önermektedir. Bu sayede, önyargı düzeyindeki değişimler takip edilebilir ve uygulanan düzeltici önlemlerin etkisi gözlemlenebilir.

- **Örnek Olay Analizi:** Sayılar kadar somut örnekler de önyargıyı gözler önüne sermede etkilidir. Standart, gerektiğinde gerçek vakaların derinlemesine incelenmesini de kapsamlı bir değerlendirme olarak görür.

Örneğin, bir bankanın kredi modelinden ret cevabı alan farklı gruplardan kişiler incelendiğinde, benzer gelir ve kredi puanına sahip olmasına rağmen bir grup üyelerinin diğerine kıyasla daha fazla reddedildiği ortaya çıkabilir. Bu tür bireysel örneklerin incelenmesi, istatistiksel olarak yakalanamayan ince önyargı biçimlerini dahi açığa çıkarabilir. Bu nedenle, standart kullanıcı deneyimlerinden veya geri bildirimlerden gelen kalitatif verilerin de hesaba katılmasını teşvik eder. Özellikle etki alanı insan hayatını doğrudan etkileyen (sağlık, adalet sistemi gibi) uygulamalarda, mağdur olabilecek grupların sesine kulak vermek de önyargının tespitinde önemli bir yöntemdir.

ISO/IEC TS 4213 standardı, önyargıyı **ölçülebilir ve izlenebilir** bir olgu haline getirmeyi amaçlar. Bu sayede, soyut etik kaygılar somut metriklere ve testlere dökülerek, teknik ekiplerin üzerinde çalışabileceği hedeflere dönüşür. Bir sonraki adım, tespit edilen önyargıları giderme yani önyargıyı azaltma tekniklerinin uygulanması olacaktır.

Önyargıyı Azaltma Teknikleri

Önyargının varlığını saptadıktan sonra, yapılması gereken düzeltici önlemleri hayata geçirmektir. ISO/IEC TS 4213, makine öğrenmesi projelerinde istenmeyen önyargıları azaltmaya yönelik çeşitli stratejiler sunar. Bu stratejiler, genel olarak üç aşamada ele alınabilir: veri ön işleme, model (algoritma) düzeyinde iyileştirmeler ve çıktı sonrası işlemler. Aşağıda, standartta vurgulanan başlıca önyargı azaltma teknikleri açıklanmaktadır:

1. **Veri Ön İşleme Teknikleri**

Verideki önyargıları gidermenin en etkili yollarından biri, daha model eğitilmeden önce veriye müdahale etmektir. Standart, veri ön işleme aşamasında kullanılabilecek birkaç temel yöntemi önerir:

- **Dengeleme (Resampling) ve Ağırlıklandırma (Reweighting):** Eğer veri setinde bazı gruplar az temsil ediliyorsa, modele girdi olarak verilen veriyi dengeli hale getirmek önyargıyı azaltabilir.

Örneğin, eğitim verisinde kadın hastaların sayısı erkeklere göre çok az ise, kadınlardan olan örnekleri aşırı örnekleme (oversampling) ile çoğaltmak veya erkek örneklerini alt örnekleme (undersampling) ile azaltmak tercih edilebilir. Alternatif olarak, her bir örneğe eğitim sırasında belirli ağırlıklar verilerek azınlık gruplarının model parametre güncellemelerine etkisi artırılabilir. Bu sayede model, az temsil edilen grubun özelliklerini öğrenmeye daha fazla odaklanır ve o grubun aleyhine oluşabilecek performans farkı azaltılır.

- **Sentetik Veri Üretimi (Data Augmentation):** Bazı durumlarda gerçek veriyi dengelemek zor veya yetersiz olabilir. Bu hallerde, eksik kalan gruplar veya özellik kombinasyonları için sentetik veriler üretmek bir çözüm olabilir.

Örneğin, yüz tanıma verisinde belirli bir yaş grubundan yeterli fotoğraf yoksa, mevcut veriler üzerinde dönüşümler (yaşlandırma efekti, farklı pozlar vs.) uygulayarak o yaş grubunu temsil edecek ek veriler oluşturulabilir. Sentetik verinin dikkatli kullanılması gerekir; ancak doğru yapıldığında modelin nadir görülen durumları daha iyi öğrenmesini sağlayarak önyargıyı azaltır.

- **Öznitelik (Feature) Dönüşümleri ve Filtreleme**: Bazı durumlarda verideki belirli özelliklerin çıkarılması veya dönüştürülmesi önyargıyı engelleyebilir.

Örneğin, model için aslında gerekli olmayan ancak önyargıya yol açabilecek bir özelliği (cinsiyet, ırk gibi) tamamen çıkarmak düşünülebilir. Bu yaklaşım, modelin ilgili bilgiye erişimini keserek dolaylı yoldan o faktöre göre ayrımcılık yapmasını engelleyebilir. Ancak bu her zaman kesin çözüm değildir; zira model, çıkarılan özelliği diğer korelasyonlu bilgilerden yine de tahmin edebilir (örneğin isimlerden cinsiyet tahmini gibi).

Yine de standart, hassas değişkenlerin kullanımına dair dikkatli bir değerlendirme yapılmasını önerir. Ayrıca, veri normalizasyonu, eksik değer atama gibi klasik ön işleme tekniklerinin de önyargıyı azaltıcı etkileri olabilir.

Örneğin, gelir dağılımı çok geniş aralıktaysa logaritmik dönüşüm gibi yöntemlerle ölçeklendirme yapmak, aşırı uçların modele etkisini azaltarak uç değerlerin temsil ettiği grupları koruyabilir.

- **Önyargılı Verilerin Çıkarılması:** Eğer tespit edilmiş bariz önyargılı bir veri altkümesi varsa (örneğin bir anket verisinde, belli bir şehirden gelen yanıtların tutarsız ve baskın olduğu fark edildiyse), bu kısmın veri kümesinden çıkarılması düşünülebilir. Ancak veri kaldırma işlemi dikkatli yapılmalıdır; zira veriyi azaltmak, modelin genel performansını düşürebilir veya başka bir önyargı yaratabilir. Standart, bu nedenle veri çıkarma gibi işlemlerin etki analizinin yapılmasını ve mümkünse yerine daha dengeli verinin koyulmasını tavsiye eder.

2. Algoritmik (Model İçi) Teknikler

Veri ön işleme kadar, modelin eğitimi sırasında uygulanan yöntemler de önyargı ile mücadelede kilit rol oynar. ISO/IEC TS 4213, algoritmik düzeyde önyargı azaltımı için şu yaklaşımları gündeme getirmektedir:

- **Adalet Kısıtları Uygulama:** Modelin eğitiminde, sadece hata oranını değil aynı zamanda adalet kriterlerini de optimize etmeye yönelik kısıtlar veya düzenlemeler eklenebilir.

Örneğin, bir sınıflandırma modelini eğitirken, "gruplar arasındaki doğru pozitif oran farkı 0'a yakın olsun" şeklinde bir ek kısıt konulabilir. Literatürde "eşit fırsat" veya "eşit hatalar" gibi kısıtlar olarak bilinen bu yöntemler, eğitim sırasında kayıp fonksiyonuna ceza terimleri ekleyerek modelin belirli bir grupta çok kötü performans göstermesini engellemeye çalışır. Bu sayede model, genel doğruluktan bir miktar feragat etse bile daha dengeli sonuçlar üretebilir.

- **Fairness-Aware (Adil) Algoritmaların Kullanımı**: Son yıllarda, makine öğrenmesi topluluğu tarafından adilliği gözeten özel algoritmalar geliştirilmiştir.

Örneğin, Adversary (Düşmancıl Temelli Adillik Modelleri, bir modelin çıktılarından hangi gruba ait olduğunu tahmin etmeye çalışan ikinci bir ağı eğiterek, orijinal modelin bu bilgiyi yansıtmasını engeller. Bu şekilde, modelin çıktıları ile hassas özellikler arasındaki ilişki minimize edilir (yani model "grup körü" davranmaya zorlanır).

Bir başka yaklaşım da karar ağaçları veya ormanları gibi algoritmalarda yapısal değişiklikler yaparak, veri bölünmesi sırasında adaletsiz dallanmaların önüne geçmektir. ISO/IEC TS 4213, proje ekiplerinin bu tür literatürde mevcut *önyargıyı azaltma odaklı algoritmaları* araştırıp kendi problem alanlarına uygulamalarını teşvik eder.

- **Üst (hiper) parametre Ayarlarında Denge:** Genellikle üst parametre optimizasyonu yapılırken ana kriter modelin genel doğruluğudur. Standart ise, üst parametre ararken adalet metriklerinin de bir parçası olabileceğini vurgular.

Örneğin, bir modelin karmaşıklığını belirleyen bir üst parametre (ağaç derinliği, düzenleme katsayısı vb.), farklı gruplardaki performans farklarını minimal yapan bir değere ayarlanabilir. Bu yaklaşım, belki genel doğruluğu bir miktar düşürebilir ancak adil bir model elde etmeyi kolaylaştırabilir.

Üst parametre aramasını birçok amaçlı optimizasyon problemi (doğruluk + adalet birlikte) olarak ele almak, dengeli bir sonuç verebilir. Bu teknik pratikte daha fazla hesaplama gerektirse de sonuçta istenmeyen önyargıyı azaltmaya katkı sağlar.

- **Model Mimarisinde Değişiklikler:** Eğer bir derin öğrenme modeli geliştiriliyorsa, mimari düzeyde bazı önlemler alınabilir.

Örneğin, çok görevli öğrenme (multi-task learning) kullanarak ana görevin yanında adaletle ilgili yardımcı bir görevi de optimize etmek mümkün olabilir. Diyelim ki elimizde etnik kökene dair istenmeyen önyargı riski var; modele hem asıl tahmin görevini yaptırıp hem de "çıktılarından etnik köken tahmin edilememe" şeklinde bir yardımcı görev verirsek, modelin bu bilgiyi nötralize etmesini sağlayabiliriz.

Bu stratejiler, doğrudan ISO standardında spesifik örneklerle geçmese de, standardın ruhuna uygundur: Geliştiricilerin yaratıcı ve problem-odaklı yaklaşımlarla önyargıyı sistemin kalbinde (modelin içinde) engellemeye çalışmaları önerilir.

3. Çıktı Sonrası (Post-Processing) Teknikler

Model eğitildikten ve doğrulandıktan sonra bile, elde edilen modelde bir miktar önyargı kalmış olabilir. Bu durumda, çıktı sonrası düzeltme teknikleri devreye girer. ISO/IEC TS 4213, model sonuçlarını düzenleyerek adaleti iyileştirme yolunda da tavsiyeler sunar:

- **Eşik Ayarlama:** Sınıflandırma modelleri için en yaygın post-processing yöntemlerinden biri, karar eşiklerini gruplara özel olarak ayarlamaktır.

Örneğin, bir kredi verme modeli için genel eşik kredi notu 600 iken, eğer görüldü ki belirli bir dezavantajlı grup için model bu eşikte çok fazla reddediyor, o grup için eşik 580'e indirilebilir. Böylece o gruptaki kabul oranı artırılarak demografik pariteye yaklaşılır. Tabii bu tür bir ayarlama, politik bir karardır ve her durumda uygulanamaz; ancak sağlık gibi kritik alanlarda farklı eşiğe göre müdahale yöntemleri (örn. farklı tarama testlerinde farklı risk eşiği) sıkça kullanılmaktadır ve YZ'de de benzer prensiplerle adalet sağlanabilir.

- **Skor Dönüştürme ve Yeniden Ölçeklendirme**: Modelin verdiği sürekli puanlar, bazı gruplar için sistematik olarak kaydırılmış olabilir. Bu durumda yeniden ölçeklendirme (calibration) yöntemleriyle gruplar arası denge sağlanabilir.

Örneğin, yüz tanıma skoru 0-1 arasında bir değer olsun; eğer modelin kadınlar için ortalama skoru erkeklere kıyasla hep düşükse, kadınların skor dağılımını yeniden ölçekleyerek ortalamayı eşitleyebilirsiniz.

Bu işlem, tüm skorları lineer bir şekilde dönüştürmek olabileceği gibi daha karmaşık bir eşleme de olabilir (örn. her grup için ayrı bir kalibrasyon eğrisi çıkarmak). Amaç, ham model çıktılarındaki sistematik farkları giderip olabildiğince tutarlı bir ölçek yakalamaktır.

- **Çıktı Filtreleme ve Kural Tabanlı Düzeltme:** Bazı durumlarda, model çıktılarına son aşamada mantıksal kurallar uygulamak gerekebilir.

Örneğin, bir otomatik özgeçmiş eleme sisteminde, belirli bir gruptan gelen adayların yeterlilik puanı belli bir seviyenin altındaysa bile, çeşitlilik hedefleri nedeniyle insan incelemesine yönlendirilmesi kuralı konabilir. Bu, bir nevi insan kontrolü ve düzeltmesi mekanizmasıdır.

Standart, bu tip karma yaklaşımları dışlamaz; tam tersine, YZ çıktılarının insan gözetiminde adil bir şekilde kullanılmasını teşvik eder. Sonuçta hedef, sistemin kararlarının hesap verebilir ve gerektiğinde düzeltilebilir olmasıdır.

- **Geribildirim Döngüsü:** Post-processing'in bir uzantısı olarak, canlı sistemden gelen geri bildirimlerin toplanıp modeli iyileştirmede kullanılması da önyargıyı zamanla azaltmanın bir yoludur.

*Örneğin, haksız kararlar karşısında kullanıcıların itiraz mekanizmaları varsa, bu itirazlar kayıt altına alınıp düzenli analiz edilerek modelin zayıf noktaları belirlenebilir. Standart, böyle geri bildirim mekanizmalarının oluşturulmasını ve **sürekli öğrenme süreçlerinin** işletilmesini önerir. Bu sayede, post-processing sadece bir seferlik bir düzeltme olmaktan çıkar ve sistem, kullanıcı etkileşimlerinden de öğrenerek daha adil hale gelir.*

Yukarıda sıralanan teknikler, istenmeyen önyargıyı azaltmak için bir araç kutusu sunmaktadır. Her projenin ihtiyaçları farklı olacağından, standart belirli bir yöntemi zorunlu tutmaz; bunun yerine, projenin bağlamına uygun teknik kombinasyonlarının bilinçli bir şekilde seçilip uygulanmasını tavsiye eder. Önemli olan, ekiplerin önyargının farkında olarak proaktif şekilde bu yöntemleri devreye alması ve etkilerini takip etmesidir.

Şeffaflık ve Hesap Verebilirlik

Bir YZ sisteminde önyargıyı ele almak sadece teknik bir mesele değil, aynı zamanda etik ve yönetişimle ilgili bir meseledir. ISO/IEC TS 4213, şeffaflık ve hesap verebilirlik ilkelerini önyargı azaltmanın vazgeçilmez bir parçası olarak görür. Bu kapsamda standardın getirdiği öneriler şunlardır:

- **Dokümantasyon:** Geliştirme sürecinin her aşamasında alınan önyargı azaltıcı önlemlerin ve yapılan analizlerin ayrıntılı biçimde belgelendirilmesi önerilir.

Örneğin, veri kümesindeki demografik dağılımın raporlanması, hangi ön işleme tekniklerinin uygulandığının not edilmesi, model eğitiminde kullanılan parametrelerin ve varsa adalet kısıtlarının kayıt altına alınması önemlidir. Bu dokümantasyon, hem ekip içi iletişim ve bilgi transferi için, hem de gerektiğinde dış denetimler için gereklidir. Standart, böyle bir şeffaflık raporu sayesinde, sistemin dışarıdan anlaşılabilirliğinin artacağını vurgular.

- **Raporlama ve Paydaşlarla İletişim:** Modelin önyargı ile ilgili performansı ve alınan önlemler, sadece geliştiriciler için değil, aynı zamanda yöneticiler, müşteriler, düzenleyici otoriteler gibi paydaşlar için de anlaşılır olmalıdır. ISO/IEC TS 4213, belirli

aralıklarla önyargı değerlendirme raporları hazırlanmasını ve bunların ilgili paydaşlarla paylaşılmasını önerir.

Örneğin, bir şirketin yönetim kuruluna veya denetleyici kuruma sunulan raporda "modelin farklı gruplar için hata oranları, tespit edilen önyargı türleri ve giderilmesi için atılan adımlar" özetlenebilir. Bu tür raporlamalar, hesap verebilirliği artırır ve dış paydaşların sisteme güven duymasına katkıda bulunur.

- **Açıklanabilirlik:** Şeffaflığın teknik boyutunda, modelin kararlarının açıklanabilir olması da yer alır. Bir modelin neden belli bir sonuca vardığı anlaşılabilirse, önyargı olup olmadığı da daha net değerlendirilebilir. Dolayısıyla standart, özellikle kritik kararlarda açıklanabilir YZ tekniklerinin kullanılmasını teşvik eder.

Örneğin, bir sinir ağının kararını Shapley değeri veya LIME gibi yöntemlerle açıklamak, hangi özelliklerin karara ne ölçüde etki ettiğini ortaya çıkarabilir. Eğer karar üzerinde beklenmedik bir şekilde yanlı bir özellik çok etkiliyse (örn. etnik kökenin vekili olabilecek posta kodu, vs.), bu bir alarm işareti sayılıp incelenebilir.

- **Paydaş Katılımı:** Özellikle istenmeyen önyargının söz konusu olduğu durumlarda, bundan etkilenen grupların veya temsilcilerinin sürece dahil edilmesi önemlidir. Standart, mümkün olduğunda sistemin hedef kitlesinden geri bildirim alınmasını ve onların endişelerinin dinlenmesini de şeffaflık kapsamına dahil eder.

Örneğin, bir işe alım algoritması geliştiriliyorsa, insan kaynakları departmanı kadar, şirketin çeşitlilik ve kapsayıcılık ekibinin de görüşleri alınmalıdır. Bu şekilde, teknik ekibin gözden kaçırabileceği toplumsal hassasiyetler sürece yansıtılabilir.

- **Hata ve Sorumluluk Mekanizmaları**: Hesap verebilirliğin bir diğer yönü de, bir hata veya haksızlık tespit edildiğinde bunun sorumluluğunu alacak mekanizmaların hazır olmasıdır. ISO/IEC TS 4213, kuruluşların bir sorumluluk zinciri oluşturmasını önerir.

Örneğin, model hataları için bir itiraz ve düzeltme prosedürü tanımlanmalı, belli periyotlarla bağımsız denetimler yapılmalı ve sorunlar tespit edildiğinde hızlıca ele alınmalıdır. Bu prosedürler, sadece teknik düzeltmeleri değil gerekirse etkilenenlere telafi sunmayı da içerebilir. Amaç, YZ sisteminin hatalarının sahiplenilmesi ve mağduriyetlerin mümkün mertebe giderilmesidir.

Şeffaflık ve hesap verebilirlik, özetle, güvenilir YZ kavramının temel taşıdır. ISO/IEC TS 4213, teknik önlemlerin yanında bu ilkelere de odaklanarak, geliştirilen sistemlerin toplum nezdinde kabul edilebilir ve denetlenebilir olmasını sağlamaya çalışır.

Kurumsal Önlemler ve En İyi Uygulamalar

Standart her ne kadar teknik konulara odaklansa da bunun başarılı bir şekilde hayata geçebilmesi için kurumsal düzeyde desteklenmesi gerektiğinin farkındadır. YZ sistemlerinde önyargıyı azaltmak, sadece birkaç mühendislik kararıyla sınırlı kalmayıp, organizasyonun genel kültür ve süreçlerine sirayet etmelidir. ISO/IEC TS 4213'ün önerdiği bazı kurumsal en iyi uygulamalar şunlardır:

- **Önyargı Politikaları ve Rehberleri:** Kuruluşlar, YZ ve veri bilimi projeleri için bir etik ve önyargı politikası oluşturmalıdır. Bu politika, projelerin en başından itibaren nelere dikkat etmesi gerektiğini belirtir.

Örneğin, hangi hassas özelliklerin asla modele dahil edilmeyeceği, veri toplamada çeşitlilik hedefleri, performans raporlarında hangi adalet metriklerinin yer alacağı gibi hususlar yazılı hale getirilmelidir. Böyle bir rehber doküman, ekipler için ortak bir referans sağlar ve standartta belirtilen ilkelerin günlük pratiğe dönüşmesini kolaylaştırır.

- **Eğitim ve Farkındalık:** Geliştirici ekipler ve proje paydaşları, önyargı ve adalet konularında eğitilmelidir. İnsanlar farkında olmadıkları önyargıları düzeltemezler. Bu nedenle, düzenli eğitim programları ile çalışanlara YZ'de önyargı örnekleri, bu standarttaki prensipler ve kullanılabilecek araçlar anlatılmalıdır.

Örneğin, veri bilimcilere hangi Python kütüphaneleriyle adalet metriklerinin hesaplanacağı uygulamalı olarak gösterilebilir. Aynı şekilde yöneticilere de bu konuların iş riskleri (itibar kaybı, yasal yaptırım riski gibi) anlatılarak konunun önemi vurgulanmalıdır.

- **Etik ve Uyumluluk İncelemeleri:** Standart, YZ sistemlerinin sadece teknik ekiple değil, etik kurul veya uyumluluk (compliance) birimleriyle birlikte değerlendirilmesini tavsiye eder. Özellikle büyük kuruluşlarda, her YZ projesinin bir etik gözden geçirme süreci olmalıdır. Bu süreçte, bağımsız uzmanlar veya şirket içi farklı departmanlardan kişiler modelin çıktısını, veri kümesini ve metodolojisini önyargı açısından incelemelidir. Tıpkı klinik araştırmalarda etik kurul onayı gerektiği gibi, yüksek etkiye

sahip YZ sistemleri de benzer bir onay mekanizmasından geçmelidir. Bu, önyargı tespit ve azaltma adımlarının yeterince uygulandığını doğrulayacak bir kontrol katmanı sağlar.

- **Sürekli Denetim ve İyileştirme**: Bir defalık değerlendirme yeterli değildir; standardın da belirttiği gibi, sürekli izleme ve düzenli denetimler kritik önem taşır. Kuruluşlar, belirli aralıklarla (örneğin yılda iki kez) YZ sistemlerini önyargı açısından denetlemeli, ISO/IEC TS 4213 ilkelerine uyumu değerlendirmelidir. Bu denetimler sonucunda bulunan eksiklikler, bir sonraki proje güncellemesinde giderilmelidir. Ayrıca şirketin farklı projelerinin performansı da karşılaştırılabilir; böylece öğrenilen dersler kurum içinde paylaşılır. Standart, bu yaklaşımı bir geri bildirim döngüsü olarak teşvik ederek, her döngüde kurumun YZ olgunluğunun artacağını öngörür.

- **Regülasyonlarla Uyumun Takibi:** ISO/IEC TS 4213, halihazırda var olan ve gelecekte çıkacak YZ düzenlemeleriyle uyumlu bir zeminde durur.

Örneğin, Avrupa Birliği YZ Tüzüğü taslağı veya ABD'de FDA'nın sağlık YZ sistemleri için kılavuzları, önyargı ve adalet konularını gündeme almaktadır. Kuruluşlar, bu standart sayesinde kendi iç kontrollerini güçlendirerek yasal uyumu proaktif biçimde sağlayabilirler. Bu da uzun vadede hem hukuki riskleri azaltır hem de topluma karşı sorumlu bir duruş sergilemelerine yardımcı olur.

- **Çeşitlilik ve İçinde Olma:** Son olarak, kurumsal düzeyde ekiplerin ve yöneticilerin çeşitliliği de YZ önyargılarının fark edilip azaltılmasında etkilidir. Farklı geçmişlere ve bakış açılarına sahip ekip üyeleri, veride veya modelde önyargı unsurlarını daha çabuk fark edebilir. Bu nedenle standardın dolaylı bir önerisi olarak, YZ geliştirme ekiplerinde çeşitlilik sağlanması da düşünülebilir. Bu, bir standart maddesi olmasa da sivil toplum kuruluşu (STK) raporu bakış açısıyla, iyi bir *yapısal önlem* olarak değerlendirilebilir.

Kurumsal önlemler, teknik çözümlerin üzerine inşa edildiği zemini hazırlar. ISO/IEC TS 4213, teknik standart olmasının ötesinde, aslında bir yönetim rehberi niteliği de taşıyarak organizasyonlara kendi YZ sistemlerini daha bilinçli ve sorumlu şekilde yönetme fırsatı sunar.

Sonuç

ISO/IEC TS 4213:2022 standardı, YZ sistemlerinde istenmeyen önyargıların tespiti ve giderilmesi konusunda kapsamlı bir rehberlik sağlamaktadır. Bu standart, makine öğrenmesi

modellerinin geliştirme sürecinden başlayarak kullanımına ve izlenmesine kadar her adımda adalet ilkesini gözeten bir yaklaşım benimsemeyi öğütler. Böylece, YZ sistemlerinin toplumdaki farklı kesimlere eşit fayda sunması ve hiçbir gruba haksızlık etmemesi hedeflenir.

Neden Önemlidir?

Çünkü YZ kararları, işe alım, kredi değerlendirmesi, sağlık teşhisi, adli kararlar gibi kritik alanlarda giderek daha fazla rol oynuyor. Eğer bu sistemler önyargılıysa, mevcut toplumsal eşitsizlikleri artırma potansiyeline sahiptir. ISO/IEC TS 4213, tam da bu nedenle, teknolojinin insanlara adil davranmasını sağlamak üzere hazırlanmıştır. Uluslararası bir standart oluşu, farklı ülkelerden uzmanların kolektif tecrübesini yansıtır ve en iyi uygulamaları derler. Bu standardın benimsenmesi, kuruluşlara sadece etik bir fayda sağlamakla kalmaz, aynı zamanda sistemlerinin güvenilirliğini ve kalitesini arttırır. Adil bir model genellikle genelleme konusunda da daha sağlıklıdır, çünkü çeşitli veriler üzerinde tutarlı performans göstermektedir.

Standart aynı zamanda sürekli gelişime açıktır. 2024 yılında ilk sürümü yayımlanan bu teknik spesifikasyon, ilerleyen yıllarda yeni teknikler ve tecrübeler ışığında güncellenebilir. YZ alanı hızla evrilirken, önyargı ile mücadele yöntemleri de gelişecektir. Ancak ISO/IEC TS 4213'ün ortaya koyduğu temel prensipler: şeffaflık, yaşam döngüsü entegrasyonu, ölçülebilirlik ve hesap verebilirlik; kalıcı değerler olarak varlığını sürdürecektir. Bu prensipler, yalnızca YZ projelerinde değil, genel olarak veri odaklı karar verme süreçlerinde de yol gösterici olabilir.

Kaynaklar

AI Standards Hub. (2024, July 18). *ISO/IEC TS 12791:2024 – Information technology — artificial intelligence — treatment of unwanted bias in classification and regression machine learning tasks* [Community page]. Retrieved from https://aistandardshub.org/ai-standards/information-technology-artificial-intelligence-treatment-of-unwanted-bias-in-classification-and-regression-machine-learning-tasks/

International Organization for Standardization & International Electrotechnical Commission. (2021). *Information technology — artificial intelligence (AI) — bias in AI systems and AI aided decision making* (ISO/IEC TR 24027:2021). Retrieved from https://www.iso.org/standard/77607.html

International Organization for Standardization & International Electrotechnical Commission. (2022). *Information technology — artificial intelligence — assessment of machine learning classification performance* (ISO/IEC TS 4213:2022). Retrieved from https://www.iso.org/standard/79799.html

International Organization for Standardization & International Electrotechnical Commission. (2024). *Information technology — artificial intelligence — treatment of unwanted bias in classification and regression machine learning tasks* (ISO/IEC TS 12791:2024). Retrieved from https://www.iso.org/standard/84110.html

Michalsons. (2021, November 8). *ISO/IEC TR 24027:2021 bias in YZ & YZ decision making* [Blog post]. Retrieved from https://www.michalsons.com/blog/iso-iec-tr-240272021-bias-in-ai-ai-decision-making/53507

Parchetalab, R. (2024, November 14). *New standards in YZ for addressing eliminating bias for fairer medical device outcomes*. LinkedIn. Retrieved from https://www.linkedin.com/pulse/new-standards-ai-addressing-eliminating-bias-fairer-ramin-parchetalab-pcifc

5.6.2. ISO/IEC TS 12791:2024 Teknik Şartname; YZ Sistemlerinde İstenmeyen Önyargının Ele Alınması

YZ sistemlerinin adil, tarafsız ve güvenilir bir şekilde çalışması, günümüzde hem toplumsal hem de ticari alanda kritik bir hedef haline gelmiştir. Özellikle makine öğrenmesi tabanlı algoritmalar, veri ve tasarım süreçlerinden kaynaklanan önyargıları (bias) istemeden öğrenerek kararlarına yansıtabilir. Bu durum, bazı grupların sistematik olarak dezavantajlı konuma düşmesine veya hatalı sonuçlara maruz kalmasına yol açabilir.

Örneğin, otomatik bir işe alım sisteminin geçmiş verilerden öğrenerek belirli bir cinsiyeti veya ırkı diğerine tercih etmesi, istenmeyen önyargı sonucunda adaletsiz kararlar üretir. Benzer şekilde, kredi tahsisinde kullanılan bir makine öğrenmesi modeli, eğitim verisindeki dengesizlikler yüzünden belirli bir bölgedeki başvuru sahiplerini haksız biçimde riskli görebilir. Bu tür vakalar, YZ'ye duyulan güveni sarsmakta ve "YZ sistemleri kime hizmet ediyor, kimi dışarıda bırakıyor?" sorularını gündeme getirmektedir.

YZ sistemlerindeki önyargı sorununu çözmek için standartlar ve rehberler büyük önem taşır. Uluslararası standartlar, geliştiricilerden politika yapıcılara kadar geniş bir kesime ortak bir çerçeve ve dil sunarak YZ'de etik ve adil uygulamaları teşvik eder. İşte bu bağlamda, 2024 yılında yayınlanan ISO/IEC TS 12791:2024 sayılı teknik şartname, YZ sistemlerinde istenmeyen önyargının tespitine ve azaltılmasına yönelik kapsamlı bir kılavuz niteliği taşımaktadır. Bu standart, makine öğrenmesi kullanarak sınıflandırma ve regresyon görevleri yürüten YZ sistemlerinde önyargıyı ele almak üzere geliştirilmiştir. Standart, yalnızca teknik çözümleri değil, aynı zamanda bir YZ projesinin yaşam döngüsü boyunca (başlangıç tasarımından kullanım ömrünün sonuna dek) alınması gereken organizasyonel ve yönetsel tedbirleri de içermektedir.

Yapay Zekâ Sistemlerinde Önyargı (Bias) Sorunu

Önyargı (bias) terimi, YZ ve makine öğrenmesi bağlamında, bir modelin sistematik hatalar yapmasına veya belirli bir yönde sapma göstermesine neden olan her türlü sapmayı ifade eder. Önyargı, veri toplama ve seçme süreçlerinden tutun da algoritmanın tasarımına ve hatta kullanıcı etkileşimlerine kadar pek çok kaynaktan ortaya çıkabilir.

Örneğin, eğitim verisi çoğunlukla genç bireylerden toplanmışsa, bir yüz tanıma sistemi yaşlı bireyleri tanımada sistematik olarak başarısız olabilir; bu durumda veri kaynaklı bir önyargı söz konusudur. Benzer şekilde, eğer bir karar destek sistemi tasarlanırken sadece dar bir uzman grubunun varsayımları dikkate alınmışsa, model kültürel veya sosyal bir önyargıyı yansıtabilir.

Önyargının her zaman tamamen kötü veya istenmeyen olmadığını not etmek gerekir. Bazı durumlarda, isteyerek eklenen önyargılar (desired bias) sistemin hedeflerine ulaşması için gereklidir.

Örneğin, toplumsal eşitsizlikleri dengelemek amacıyla bilinçli olarak belirli bir gruba pozitif ayrımcılık uygulayan bir model tasarımı, istenen (olumlu) bir önyargı örneği sayılabilir. Bir işe alım YZ'nin geçmiş verilerdeki kadın adayların azınlıkta olmasından kaynaklanan dengesizliği telafi etmek için kadın adaylara bir miktar avantaj sağlaması, sistem hedefi doğrultusunda kasıtlı bir önyargı barındırır ve olumlu etki yaratabilir.

Tarafsız etkili (nötr) önyargılar ise sistemin sonuçlarını ciddi şekilde etkilemeyen veya kimseye bariz bir zarar vermeyen sapmalardır.

Örneğin, otonom bir aracın görüntü işlemeye dayalı algı sisteminde, posta kutularını yangın musluğu olarak sınıflandıran bir hatanın olması bir önyargıyı gösterir; fakat eğer araç her iki nesneden de eşit şekilde kaçınacak biçimde programlanmışsa, bu önyargı gerçek dünyada nötr bir etki doğurur.

Ancak, YZ uygulamalarında en çok endişe edilen istenmeyen (olumsuz) önyargılar, sistemin amaçlarında özellikle belirtilmemiş ve sonuçları itibarıyla adaletsiz, ayrımcı veya hatalı olan önyargılardır. İstenmeyen önyargılar genellikle şu durumlarda karşımıza çıkar:

- **Eğitim verisinden kaynaklanan önyargılar**:

  Veri kümesindeki dengesizlikler, temsil eksiklikleri veya tarihsel ayrımcılık izleri modelin öğrenmesine yansıyabilir.

Örneğin, sadece belirli bir etnik grubun ağırlıklı olduğu bir veriyle eğitilen model, diğer gruplara dair hatalı genellemeler yapabilir (seçme yanlılığı, örneklem yanlılığı gibi). Tarihsel verilerde var olan toplumsal önyargılar da doğrudan modele sirayet edebilir; buna tarihsel önyargı denir. 2020 öncesine ait iş başvuru verilerinde teknik rollerde kadınların az temsil edilmesi, bir işe alım algoritmasının kadın adayları sistematik olarak düşük puanlamasına yol açan tarihsel bir önyargı yaratmış olabilir.

- **İnsan kararları ve etkileşimlerinden kaynaklanan önyargılar:**

  Veri etiketleme işlemini yapan görevlilerin bilinçli veya bilinçsiz önyargıları, etiket yanlılığı olarak veri kümelerine yansıyabilir.

Örneğin, bir içerik moderasyon YZ için eğitici veri toplayan insan çalışanlar, kendi toplumsal yargıları nedeniyle bazı ifadeleri "nefret söylemi" olarak işaretlemede tutarsız davranabilir. Ayrıca, kullanıcıların YZ ile etkileşim biçimleri de yeni önyargılar doğurabilir. Kullanıcı etkileşim önyargısı olarak adlandırılan bu durumda, bir arama motorunda en üstte görülen sonuçlara daha çok tıklanması ve popüler içeriklerin daha da görünür hale gelmesi, sistemin geri besleme yoluyla giderek dar bir perspektife kaymasına neden olabilir. Bu, arama algoritmaları ve öneri sistemlerinde sık rastlanan bir önyargı türüdür (popülarite önyargısı, sıralama önyargısı vb.).

- **Algoritma ve model tasarımından kaynaklanan önyargılar:**

  Bazen modelin kullandığı algoritma veya optimizasyon hedefi, veride olmasa bile kendi başına bir önyargı ekleyebilir; bu durumda algoritmik önyargı ortaya çıkar.

Örneğin, sağlık hizmetlerinde maliyet tahmini yapan bir YZ modeli, sağlık ihtiyacını değil de harcanan sağlık giderlerini optimize edecek şekilde tasarlandıysa, yoksul veya sağlık hizmetine erişimi kısıtlı kesimleri yanlış şekilde düşük riskli sınıfa sokabilir. Bu durumda model, veride açıkça bulunmayan ancak kullanılan hedef fonksiyon nedeniyle ortaya çıkan bir önyargı üretmiş olur.

Yine model mimarisindeki tercihler veya üst parametre ayarları, belirli örnekleri sistematik olarak avantajlı/dezavantajlı kılabilir (örneğin, bir sinir ağındaki aktivasyon işlevi seçiminden doğan uç değer duyarlılığı, bazı veri noktalarını dışlayıcı etki yapabilir).

- **Değerlendirme ve geribildirim süreçlerindeki önyargılar:**

Modelin başarımını ölçerken kullanılan metrikler veya test verileri de önyargı taşıyabilir. Bu durum değerlendirme önyargısı olarak bilinir.

Örneğin, yüz tanıma modellerini değerlendirmek için kullanılan yaygın bir karşılaştırma veri kümesi, çoğunlukla açık tenli bireylerin fotoğraflarından oluşuyorsa, modelin koyu tenli bireylerdeki hatası fark edilmeyebilir. Adil bir değerlendirme yapılmadığı için model gerçekte adaletsiz bir performansa sahip olduğu halde kabul edilebilir görünebilir.

Yukarıdaki örnekler, önyargının çok kaynaklı ve çok katmanlı bir sorun olduğunu gösteriyor. ISO/IEC TR 24027:2021 Teknik Raporu, YZ sistemlerindeki önyargıları kapsamlı bir şekilde sınıflandırarak bu duruma dikkat çekmiştir. Bu sınıflandırmaya göre önyargılar;

- İnsan kaynaklı bilişsel önyargılar,
- Veri kaynaklı istatistiksel önyargılar,
- Algoritmik tasarım önyargıları ve
- Çıktı/etkileşim önyargıları gibi kategorilere ayrılabilir.

Örneğin, doğrulama önyargısı (confirmation bias), geliştiricilerin veya veri bilimcilerin kendi hipotezlerini doğrulayacak şekilde deney tasarlamaları sonucu ortaya çıkan bir insan bilişsel önyargısıdır. Seçme yanlılığı (selection bias) ise veri önyargısı kategorisinde olup, toplanan veri örneklerinin gerçek dağılımı temsil etmemesinden kaynaklanır. Tarihsel önyargı, grup içi/dışı önyargılar, eksik veri önyargısı, çıktı sunum önyargısı gibi literatürde tanımlanmış pek çok önyargı türü mevcuttur ve standart, bunların taksonomik bir çerçevesini sunarak ortak terimler kullanılmasını sağlamaktadır. Böylece, geliştiriciler ve karar alıcılar, bir YZ sisteminde hangi tür önyargı sorunuyla karşı karşıya olduklarını net biçimde tanımlayabilir.

YZ alanında istenmeyen önyargılar hem etik hem de performans açısından ciddi riskler doğurur. Bu önyargılar, YZ sistemlerinin çıktılarında negatif etki yaratarak bireylerin fırsatlarına ket vurabilir, kurumsal kararlarda adaletsizliklere yol açabilir ve genel olarak teknolojiye güveni zedeleyebilir. Bu nedenle, önyargının yaşam döngüsü yaklaşımıyla ele alınması ve daha tasarım aşamasından itibaren sistemli biçimde azaltılması gerekliliği ortaya çıkmıştır. ISO/IEC TS 12791:2024 standardı da tam olarak bu noktada devreye girerek, YZ projelerinde önyargıyı baştan sona yönetmek için kılavuzluk sunmayı amaçlamaktadır.

5.6.3. ISO/IEC TS 12791:2024 Standardına Genel Bakış

ISO/IEC TS 12791:2024, YZ sistemlerinde istenmeyen önyargının tedavisi (azaltılması) konusunda uluslararası düzeyde hazırlanmış ilk ayrıntılı standartlardan biridir. ISO ve IEC'nin ortak teknik komitesi JTC 1/SC 42 (YZ) tarafından geliştirilmiş olup Ekim 2024'te Teknik Şartname (Technical Specification) statüsünde yayınlanmıştır. *Teknik Şartname (TS)*, bir alandaki en iyi uygulamaları hızla toplamak ve ilgili kesimlere yol göstermek için kullanılan bir standart türüdür; genellikle ilerleyen yıllarda deneyim kazanıldıkça uluslararası standart (IS) haline gelebilir. TS 12791:2024, tüm sektörler ile her ölçekteki kuruluş için uygulanabilir şekilde genel ilkeler ortaya koymaktadır. Yani büyük bir teknoloji şirketi de, küçük bir startup da bir kamu kurumu da bu standarttaki prensiplerden faydalanabilir.

Standart, geliştirilirken halihazırda mevcut ilgili çalışmalar temel alınmıştır. Özellikle ISO/IEC TR 24027:2021 sayılı Teknik Rapor'dan elde edilen birikim, TS 12791'in çatısını oluşturur. TR 24027, YZ sistemlerinde önyargı konusunu tanımlayan, farklı önyargı türlerini ve etkilerini anlatan bir kılavuz niteliğindeydi. TS 12791 ise bu teknik rapordaki önyargı taksonomisini resmileştirmiş ve onu bir adım öteye taşıyarak pratik azaltma teknikleri ve süreç önerileriyle donatmıştır. Ayrıca TS 12791, YZ sistemlerinin yaşam döngüsünü ele alırken ISO/IEC 22989 (YZ kavramları ve terminolojisi standardı) ile ISO/IEC 5338:2023 (YZ sistem yaşam döngüsü süreçleri standardı) tarafından tanımlanan aşamalara atıf yapar. Böylece, önyargı yönetimi için önerilen adımlar genel YZ proje yönetimi süreçleriyle uyumlu hale getirilmiştir.

Standardın kapsamı özellikle makine öğrenmesi kullanarak sınıflandırma veya regresyon görevleri gerçekleştiren YZ sistemleriyle sınırlıdır. Bunun nedeni, TR 24027'nin geliştirilmesi sırasında uzmanlardan gelen geri bildirimlerin büyük ölçüde denetimli öğrenme (sınıflandırma/regresyon) sistemlerine odaklanmış olmasıdır. Bu alan, önyargıların en somut ve ölçülebilir etkilerinin görüldüğü, aynı zamanda geniş uygulama yelpazesine sahip bir alan

olduğu için önceliklendirilmiştir. Sınıflandırma görevleri (örneğin bir resimdeki nesnenin tipini belirleme, bir e-postanın spam olup olmadığını saptama) veya regresyon görevleri (örneğin bir kişiye verilecek kredi limitini sayısal olarak tahmin etme) günlük hayatta pek çok kritik karara temel teşkil eder. Bu nedenle standart, bu tip sistemlerde önyargıyı azaltmaya dönük yöntemleri derinlemesine ele alarak odaklı bir rehber sunar.

ISO/IEC TS 12791:2024, bağlamdan bağımsız (context-agnostic) bir yaklaşım benimser. Yani önyargı kavramını belirli bir alan veya tek bir "fairness/adalet" tanımıyla sınırlamaz. Örneğin, bazı düzenlemeler veya kurumlar önyargıyı özellikle *ayrımcılık yasağı* bağlamında ele alırken, kimileri *istatistiksel tutarlılık* olarak tanımlayabilir. Bu standart ise önyargıyı genel anlamda istenmeyen bir sapma olarak çerçeveler ve sunulan çözüm teknikleri, hangi adalet veya etik çerçevenin benimsendiğinden bağımsız olarak uygulanabilecek şekilde tasvir edilir. Bu sayede standart, ister "etiketli gruplar arası adalet" gibi belirli bir fairness ölçütünü benimseyen kuruluşlar olsun, ister sadece doğruluk ve hata payı açısından önyargıyı ele alan teknik ekipler olsun, hepsine uygun olabilecek *genel geçer* yöntemler önerir. Elbette uygulamada her kuruluş, içinde bulunduğu bağlamın ihtiyaçlarına göre bu teknikleri adapte edecektir; ancak TS 12791'in dili ve kapsamı, belirli yasa veya sektör kısıtlarına atıf yapmadan evrensel iyi uygulamaları tariflemeyi amaçlar.

Standart, YZ sisteminin yaşam döngüsü boyunca atılması gereken adımları tarif ederken, her aşamada karşılaşılan önyargı risklerine dikkat çekmekte ve uygun azaltma stratejilerini önermektedir. Geliştirmenin en başından (ihtiyaçların tanımlanması ve veri toplama) başlayarak, modelin tasarımı, doğrulama testleri, kullanım sırasındaki izleme ve hatta model kullanım dışı bırakılırken alınacak önlemlere dek kapsamlı bir plan sunar. Bu yaklaşım, yalnızca belirli bir teknik yönteme odaklanmak yerine, sürekli bir iyileştirme ve gözetim süreci tariflediği için önemlidir. Nitekim standart, önyargıyı bir seferlik düzeltilebilir bir hata değil, *yönetilmesi gereken bir risk* olarak ele alır.

YZ Yaşam Döngüsü Boyunca Önyargının Ele Alınması

ISO/IEC TS 12791:2024, bir YZ sisteminin yaşam döngüsünü çeşitli aşamalara ayırarak her aşamada önyargı ile mücadele yöntemlerini ayrı ayrı tartışmaktadır. Bu yaklaşım, önyargının yalnızca model eğitimi esnasında değil, proje fikrinin oluşmasından sistemin emekliye ayrılmasına kadar her noktada ortaya çıkabileceği gerçeğine dayanır. Aşağıda, standartta tanımlanan ana yaşam döngüsü adımları ve bu adımlarda yapılması tavsiye edilen önyargı azaltma faaliyetleri açıklanmaktadır:

1. Proje Başlangıcı (Inception) Aşamasında Önyargı Yönetimi

Bir YZ projesinin en başında, yani inception aşamasında, proje ekibinin ve paydaşların önyargı risklerini göz önünde bulundurarak planlama yapması gerekir. ISO/IEC TS 12791, inception aşamasında aşağıdaki adımların üzerinde durur:

- **Paydaşların ve Etkilenecek Grupların Belirlenmesi:** Projenin çıktılarından etkilenecek tüm paydaşlar (nihai kullanıcılar, karar vericiler, ilgili toplum kesimleri vb.) erkenden tanımlanmalıdır. Bu sayede, herhangi bir grubun ihtiyaç veya endişeleri göz ardı edilmez.

Örneğin, bir sağlık tanı sistemi geliştiriliyorsa, sadece doktorlar ve hastane yönetimi değil, hastalar, farklı demografik gruplar ve hatta sağlık otoriteleri de paydaş listesinde yer almalıdır. Standart, önyargı riskinin genellikle dışlanan veya görünmez kalan paydaş gruplarında yoğunlaştığını vurgular. Bu nedenle kapsayıcı bir paydaş analizi yapmak, daha başlangıçta "kimi için adil?" sorusunu sorarak önyargının hedefini daraltmaya yardımcı olur.

- **Paydaş İhtiyaç ve Gereksinimlerinin Tanımlanması:** Her paydaş grubunun beklentileri ve gereksinimleri adalet ve önyargı açısından değerlendirilmelidir.

Örneğin, bir kamu hizmeti için geliştirilen makine öğrenmesi sistemi, yasal olarak belirli gruplara karşı tarafsız olmak zorundadır; bu durumda hukuki gereklilikler sistem gereksinimlerine dahil edilmelidir. Aynı şekilde, projenin etik ilkeleri belirlenirken "sistem hangi düzeyde hata/önyargı yaparsa kabul edilemez olur?" sorusu yanıtlanmalıdır. Standart, bu aşamada açıkça "adalet kriterleri" veya "kabul edilebilir önyargı düzeyi" tanımlamayı önerir. Bu kriterler daha sonra sistemin başarısını değerlendirirken kullanılacak kabul şartlarını oluşturacaktır.

- **Çapraz Disiplinli Ekip Oluşturma:** Önyargı genellikle farklı disiplinlerin (teknik, sosyal, hukuki) kesişiminde ortaya çıkan bir sorundur. Bu nedenle inception aşamasında, ekibin sadece veri bilimciler veya mühendislerden değil, aynı zamanda etik uzmanları, alan uzmanları, sosyologlar veya hukukçulardan da oluşması tavsiye edilir. Standart, projenin başında *"uygun alanlar arası uzmanlığa sahip olunduğundan emin olunması"* gerektiğini vurgular.

Örneğin, bir yüz tanıma sistemi geliştiriliyorsa, ekibin içinde farklı etnik kökenlerden insanlar veya insan hakları konusunda bilgili uzmanlar bulunması, olası önyargıları erkenden fark etmeye yardımcı olabilir.

Veri Kaynaklarının Belirlenmesi ve Değerlendirilmesi

Bu aşamada kullanılacak veri kaynakları önceden planlanmalı ve önyargı potansiyeli açısından değerlendirilmelidir. Hangi veri kümelerinin kullanılacağı, verinin hangi popülasyonu temsil ettiği, güncelliği ve toplanma şekli önemlidir. Standart, verinin içeriğine dair meta-verilerin ve kaynak bilgisinin kaydedilmesini önerir; böylece ileride veride önyargı tespiti kolaylaşır. Ayrıca, bazı kritik verilerin toplanmasının yasal/etik nedenlerle kısıtlı olabileceğini (örneğin ayrımcılık yasağı nedeniyle iş başvurularında adayın cinsiyet bilgisinin toplanmaması gibi) belirtir. Bu tür durumlar, daha sonra modelde belirli gruplara dair performans analizi yapmayı zorlaştırabilir. Dolayısıyla, inception safhasında veri kaynakları planlanırken eksik veya kısıtlı veri alanlarının nasıl yönetileceği de kararlaştırılmalıdır. Gerekirse, ileride *ek veri toplama* veya *sentetik veri oluşturma* opsiyonları göz önünde bulundurulmalıdır.

- **Tedarik ve Dış Kaynak Kullanımı (Procurement):** Eğer proje kapsamında harici bir YZ aracı, model veya veri satın alınacaksa, tedarik sürecine önyargı kriterlerinin dahil edilmesi gerekir. ISO/IEC TS 12791, YZ sistemleri tedarik edilirken *"önyargı ile ilgili beklentilerin sözleşmelere yansıtılması"* gerektiğini belirtir.

Örneğin, bir kurum dışarıdan bir yüz tanıma yazılımı alacaksa, tedarik sözleşmesinde bu yazılımın farklı demografik gruplarda belli bir doğruluk seviyesini garanti etmesi talep edilebilir veya ürünün bir önyargı değerlendirme raporu sunması istenebilir. Bu, tedarikçilerin de sorumluluk üstlenmesini sağlayarak önyargı riskini azaltır. Benzer şekilde, dışarıdan sağlanan veri kümeleri için sağlayıcılardan veri toplama süreçlerinde önyargıyı minimize edecek tedbirleri almış olması beklenebilir.

- **Risk Yönetimi ile Entegrasyon:** Önyargı konusu, genel YZ risk yönetimi çerçevesinin bir parçası haline getirilmelidir. ISO 31000 gibi risk yönetimi standartlarına aşina kuruluşlar, bu standartlara uygun olarak bir *YZ risk matrisi* oluştururken önyargıyı da bir risk faktörü olarak eklemelidir. Standart, özellikle yüksek riskli YZ uygulamalarında (örneğin adli karar destek sistemleri, kredi skorlama sistemleri gibi insanların hayatını ciddi etkileyebilen uygulamalarda)

önyargının ayrı bir risk kalemi olarak ele alınmasını önerir. Risk yönetimi entegrasyonu kapsamında, önyargı olasılığı ve etkisi değerlendirilir, bir *risk iştahı* belirlenir ve gerekli kontrol mekanizmaları planlanır.

Örneğin, belirli bir önyargı türünün (cinsiyet yanlılığının) tamamen önlenmesi mümkün değilse, bunun etkisini azaltacak ek denetimler (insan onayı, ikincil gözden geçirme vb.) bir risk kontrolü olarak tanımlanabilir.

- **Kabul Kriterlerinin Belirlenmesi:** Projenin başarı kriterleri arasında, önyargı ile ilgili kabul edilebilirlik limitleri tanımlanmalıdır. Bu kriterler, paydaş gereksinimleri ve risk değerlendirmesi ışığında şekillenir.

Örneğin, "Modelin pozitif tahmin değerinde kadın ve erkek adaylar arasında en fazla %5'lik bir fark olabilir" veya "Sistem hiçbir etnik grup için genel hata oranını %10'un üzerine çıkarmamalıdır" gibi nicel eşikler konulabilir. Bu kabul kriterleri, ileride sistem test edilirken referans noktaları olacaktır.

Standart, bu kriterlerin mümkün olduğunca ölçülebilir ve net olmasını tavsiye eder. Ayrıca sadece istatistiksel eşiklerin değil, aynı zamanda süreçle ilgili kriterlerin de tanımlanabileceğini belirtir.

Örneğin, "Model çıktıları önemli kararlar için kullanılmadan önce mutlaka bir insan tarafından incelenmelidir" gibi bir şart da kabul kriterinin parçası olabilir. Bu tür kriterler, önyargı kaynaklı hatalı bir kararın direkt uygulanmasını engelleyen bir son kontrol mekanizması işlevi görür.

Başlangıç (Inception) aşaması, önyargıyla mücadelede proaktif olmanın en önemli adımıdır. Bu aşamada yapılan kapsamlı planlama ve farkındalık çalışması, daha sonraki teknik adımlarda çıkabilecek sorunların pek çoğunu başlamadan önleyebilir. ISO/IEC TS 12791, bu nedenle yaşam döngüsünün başlangıç noktasına özellikle vurgu yaparak, *"En baştan doğru hedefleri koy, doğru ekibi kur, doğru soruları sor"* ilkesini benimsetmeye çalışır. Bir projede önyargı risklerinin baştan kabul edilmesi ve planlanması, o projenin sonucunda ortaya çıkacak YZ sisteminin daha adil ve güvenilir olacağının teminatı gibidir.

2. Tasarım ve Geliştirme (Design & Development) Aşamasında Önyargı Azaltma

Projenin planlaması yapıldıktan ve gereksinimler belirlendikten sonra, YZ sisteminin tasarım ve geliştirme aşamasına geçilir. ISO/IEC TS 12791, bu aşamada geliştiricilerin

bilinçli kararlarla önyargıyı en aza indirmesi için çeşitli yöntemler önerir. Başlıca odak noktaları şunlardır:

- **Özellik (Feature) Temsili ve Seçimi**: Modelin öğreneceği özelliklerin (değişkenlerin) nasıl temsil edileceği, önyargının kaynağı olabilir. Standart, özellik tasarımı sırasında dikkatli olunmasını ve özellikle *korunan özellikler* veya *proxy (vekil) özellikler* konusunda özen gösterilmesini tavsiye eder.

Örneğin, iş başvurusu değerlendiren bir modelde "üniversite mezunu olup olmama" bilgisi masum görünebilir ancak aslında sosyo-ekonomik geçmişin bir temsilcisi olabilir ve dolaylı olarak etnik/demografik ayrımcılığa neden olabilir. Geliştiriciler, her bir özelliğin modele nasıl bir bilgi kattığını ve olası önyargıları temsil edip etmediğini analiz etmelidir. Eğer belirli bir özellik, istenmeyen bir önyargının taşıyıcısı ise, o özelliğin çıkarılması veya farklı şekilde kodlanması (örneğin kategorik bir değeri sayısal skala yerine etiket olarak kullanmak gibi) düşünülmelidir. Ayrıca, özellik değerlerinin ölçeklendirilmesi ve normalizasyonu da önem taşır; zira yanlış ölçekleme, bir grubun verilerini yapay olarak genişletebilir veya daraltabilir (örneğin, gelir verisini log dönüşümü yapmadan kullanmak, yüksek gelirlilerin etkisini abartabilir).

- **Veri Kümesi Yeterliliği ve Meta-Veri (Metadata) Sağlanması**: Modelin geliştirilmesinde kullanılan veri kümeleri için mümkün olduğunca zengin meta-veri tutulmalıdır. Meta-veri, veri kümesinin hangi koşullarda toplandığı, örneklerin hangi alt gruplardan geldiği, veride eksiklikler olup olmadığı, etiketleme süreçlerinin nasıl yürütüldüğü gibi bilgileri içerir. Standart, bu bilgilerin toplanmasının önyargı tespiti ve ileride yapılacak düzeltici işlemler için kritik olduğunu vurgular.

Örneğin, bir görüntü veri kümesinde her bir resmin çekildiği cihaz, ışık koşulları veya coğrafi konum gibi meta-veriler mevcutsa, modelin belirli koşullarda neden hatalı davrandığı daha kolay anlaşılabilir. Ayrıca meta-veri, veri çeşitliliğinin izlenmesine de olanak tanır; geliştiriciler veri kümesinin alt gruplarının dağılımını bu sayede inceleyerek (ör. cinsiyet, yaş, coğrafya bazında) bariz dengesizlikleri fark edebilirler.

- **Veri Etiketleme ve Açıklama Süreci**: Eğer yeni veri toplanıyor veya mevcut veriye yeni etiketler ekleniyorsa, etiketleme sürecinin önyargıya karşı korunması gerekir. Bu, bir yandan etiketi veren kişilerin önyargılarının asgari düzeyde

kalmasını sağlamayı, diğer yandan da etiket kategorilerinin dikkatli tanımlanmasını içerir. Standart, etiketleyicilere eğitim vermeyi, mümkünse birden fazla etiketleyici kullanarak tutarlılığı ölçmeyi ve *"etiketleyici önyargısı"* olup olmadığını kontrol etmeyi önerir.

Örneğin, insan duygularını tanıtan bir veri kümesi oluşturuluyorsa, "öfke", "üzüntü" gibi etiketlerin kültürel bağlamlara göre değişebileceği hatırlanmalı ve etiket tanımları netleştirilmelidir. Gerekirse, çoklu etiketleme yapılarak her örnek için konsensus tabanlı bir doğru etiket belirlenebilir. Ayrıca, etiketleme yapılırken veri anonimleştirmesi de önyargı riskini azaltabilir; etiketleyiciler, örneğin bir özgeçmişi değerlendirirken adayın ismini veya cinsiyetini görmezse daha tarafsız etiketleyebilir.

- **Veri Düzeltme ve Ayarlama (Adjusting Data):** Geliştirme aşamasında, model eğitimine geçmeden önce veri üzerinde düzeltici işlemler uygulamak gerekebilir. Bu, önyargı tespit edilen alanlarda verinin düzenlenmesi anlamına gelir.

Örneğin, veri kümesinde belirli bir gruba ait örnek sayısı çok azsa, yeniden örnekleme (resampling) veya ağırlıklandırma (reweighting) teknikleri ile bu dengesizlik giderilebilir. Standart, bu tür ön-işleme tekniklerinin (pre-processing) sıklıkla uygulanmasını ve eğitim verisindeki bariz önyargıların modele verilmeden önce düzeltilmesini önerir. Yine örnek vermek gerekirse, yüz tanıma verisinde kadınların oranı %20, erkeklerin %80 ise, ya erkeklerden bir kısmı atılarak ya da kadın fotoğrafları artırılarak denge sağlanabilir.

Ancak bu işlemlerin dikkatli yapılması gerekir; zira aşırı örnekleme veya silme işlemleri, modelin genel başarımını etkileyebilir. Bu noktada, standart *veri artırma (data augmentation)* yöntemlerinin de önyargı azaltmak için yaratıcı biçimde kullanılabileceğini belirtir.

Örneğin, az temsil edilen bir sınıfın verilerini çoğaltmak için bu verilere küçük gürültüler eklemek, dönüştürmek veya sentetik veriler üretmek bir çözüm olabilir.

- **Önyargı Riskleri İçin Çözüm Yöntemleri Belirleme:** Tasarım sırasında, inception aşamasında belirlenen potansiyel önyargı risklerine karşı hangi teknik yöntemlerin uygulanacağı planlanmalıdır. ISO/IEC TS 12791, geliştiricilerin bu aşamada literatürdeki önyargı azaltma algoritmalarını göz önünde bulundurmalarını ve uygun olanları seçmelerini önerir. Bu, bir nevi *"önyargı tedavi planı"* hazırlamak gibidir.

Örneğin, proje ekibi, sistemde ırksal önyargı riski tespit etmiş olsun. Tasarım aşamasında, eğitim sırasında kullanılmak üzere bir düzenleme terimi (regularizer) eklemeye karar verilebilir (örneğin, farklı gruplar için hata oranlarını eşitlemeye çalışan bir ceza terimi). Ya da model çıktıları sonrasında bir post-processing *yöntemi uygulanarak kararların adil dağılımı sağlanabilir. Bu tür yöntemlere raporun ilerleyen kısmında daha ayrıntılı değineceğiz. Burada önemli olan, tasarım aşamasında bu tekniklerin envanterinin çıkarılması ve ekibe gerekli bilgi/araç desteğinin sağlanmasıdır. Geliştirme başlamadan, hangi kütüphanelerin veya araçların kullanılacağı, ekibin bu teknikleri uygulamak için eğitim gereksinimi olup olmadığı da değerlendirilmelidir.*

- **Algoritma ve Model Seçimi:** Bazı makine öğrenmesi algoritmaları, doğası gereği daha şeffaf veya kontrol edilebilir olabilir. Standart, tasarım aşamasında hedefe uygun algoritma seçiminde önyargı açısından da düşünülmesi gerektiğini ima etmektedir.

Örneğin, karar ağaçları veya lineer modeller gibi açıklanabilirlik düzeyi yüksek modeller, önyargının izini sürmek için daha elverişlidir. Derin öğrenme modelleri çok güçlü olsa da kara-kutu yapıları nedeniyle önyargı tespiti ve düzeltmesi zor olabilir. Bu nedenle, bir miktar doğruluk pahasına daha basit ama izlenebilir bir model tercih etmenin gerekip gerekmediği kararını kuruluşlar bu aşamada vermelidir. Eğer kompleks bir model zorunluysa, o zaman daha sonra gelecek doğrulama aşamasında ekstra özen göstermek ve belki de model içi açıklanabilirlik araçları kullanmak planlanmalıdır.

Tasarım ve geliştirme aşaması, teknik anlamda önyargıyla *mücadeleye başlandığı* noktadır. Yukarıdaki önlemler, model ortaya konmadan evvel önyargıyı minimuma indirmeyi hedefler. Unutmamak gerekir ki, bu aşamada yapılan her iyileştirme ileride karşılaşılabilecek sorunları orantısız şekilde azaltır: Önyargılı bir modeli sonradan düzeltmek, önyargısız bir model tasarlamaktan çok daha zordur. Bu yüzden ISO/IEC TS 12791, "önlemek, düzeltmekten iyidir" prensibiyle tasarım ve geliştirme adımlarının hakkıyla yerine getirilmesini teşvik eder.

3. Doğrulama ve Geçerleme (Verification & Validation) Aşamasında Önyargı Kontrolü

Model geliştirildikten sonra, sıra onun test edilmesine, doğrulanmasına ve geçerliğinin onaylanmasına gelir. Doğrulama ve geçerleme (V&V) aşaması, modelin belirlenen gereksinimleri karşılayıp karşılamadığını değerlendirdiğimiz kritik bir süreçtir. ISO/IEC TS

12791, bu aşamada önyargının tespitine ve ölçümüne yönelik özel faaliyetleri vurgular. Başlıca öneriler şöyledir:

- **Veri Kümelerinin Statik Analizi (Static Testing of Data):** Modelin eğitildiği, doğrulandığı ve test edildiği veri kümeleri, önyargı belirtileri açısından istatistiksel olarak incelenmelidir. Bu, modele girdi olmadan, verinin kendi içindeki dağılımlara bakarak potansiyel sorunların ortaya çıkarılması anlamına gelir.

Örneğin, test veri kümesindeki her bir alt grubun (farklı cinsiyet, yaş aralığı, bölge vb.) oransal temsiline bakılabilir. Eğer test verisi gerçek dünyayı yansıtmıyor veya belirli gruplar yeterince temsil edilmiyorsa, modelin bu gruplardaki performansı güvenilir olmayacaktır. Aynı şekilde, eğitim verisinde belirli özellik kombinasyonlarının eksik olması (veri boşlukları) önyargı kaynağı olabilir.

Standart, geliştiricilerin veri üzerinde çeşitli farklılık ölçümleri yapmasını önerir.

Örneğin, her bir duyarlı özellik (protected attribute) için dağılım eşitliği kontrolü, özellikler arası korelasyon analizleri, varsa etiket dağılımlarında adaletsizlik (örneğin "olumlu" olarak işaretlene bir etiketin bir grupta çok az, diğerinde çok fazla olması) gibi. Simpson paradoksu gibi istatistiksel tuzaklara karşı da uyanık olunmalıdır; bir modeli parçalara göre adil görünüp tüm veri birleştiğinde adaletsiz sonuç verebilir durumlar için uyarılar mevcuttur. Bu tip analizler, henüz model devreye girmeden verideki olası önyargıları belgelemeyi sağlar.

- **Model Çıktılarının Dinamik Testi (Dynamic Testing of Model):** Model çalışır hale geldikten sonra, doğrulama aşamasında dinamik testler ile modelin önyargı davranışı ölçülmelidir. Bu kapsamda, standart çeşitli *fairness* metriğinin kullanılmasını önerir.

Örneğin, "eşit hata oranı", "pozitif tahmin değeri paritesi", "demografik parite" veya "eşit fırsat" gibi ölçütler, modelin farklı gruplara karşı nasıl davrandığını sayısal olarak ortaya koyar. Örneğin, bir kredi onay modelinde erkek ve kadın başvuru sahipleri için onay oranları ve hata oranları ayrı ayrı hesaplanıp karşılaştırılabilir. Eğer arada anlamlı bir fark varsa, bu modelin cinsiyet bazlı bir önyargı içerdiğine işaret eder.

Standart, tek bir metrik yerine mümkün olduğunca birden fazla ölçütle değerlendirme yapmanın resmin tamamını görmek açısından yararlı olduğunu belirtir. Zira bir metrikte adil görünen model, başka bir açıdan hala sorunlu olabilir.

Örneğin, model "demografik parite" sağlasa bile "eşit fırsat" ilkesini çiğneyebilir. Bu nedenle, projenin başında belirlenen kabul kriterlerine uygun olarak, çeşitli metriklerin hesaplanıp bunların kabul eşikleriyle karşılaştırılması gerekir. Ayrıca dinamik test sürecinde, counterfactual test (karşı-olgusal test) gibi yöntemler de önerilir: Yani bir girdi veride hassas bir özelliği değiştirip (örneğin aynı özgeçmişte sadece ismi erkekten kadına çevirerek) modelin çıktısının nasıl değiştiğini gözlemlemek. Eğer küçük bir değişiklik, model kararını dramatik şekilde değiştiriyorsa, model muhtemelen o hassas özelliği dolaylı yoldan kullanıyor ve önyargılı davranıyor demektir.

- **Dil ve Görüntü Modellerinde Özel Testler:** Eğer model bir doğal dil işleme veya görüntü tanıma sistemi ise, önyargı testleri buna göre uyarlanmalıdır. Standart her ne kadar spesifik uygulama detaylarına girmese de, bu alanlarda bilinen önyargı testlerine işaret edilebilir.

Örneğin, dil modelleri için önceden tanımlı önyargı test cümleleri kullanarak modelin belirli gruplar hakkında olumsuz çağrışımlar yapıp yapmadığı test edilebilir. Görüntü sistemleri için, farklı ten rengi, cinsiyet veya yaş gruplarına ait görüntülerde model performansının tutarlı olup olmadığı incelenebilir. Bu tür testler, genellikle bir performans matrisi şeklinde sunulur: Modelin her bir alt grup için doğruluk, duyarlılık, özgüllük gibi ölçümleri listelenir ve karşılaştırılır.

- **Onay ve Kabul Süreci:** Doğrulama aşamasının sonunda, modelin önyargı kriterlerine göre kabul edilip edilmeyeceğine karar verilmelidir. Eğer model, inception aşamasında belirlenen kabul kriterlerini karşılamıyorsa (örneğin önyargı metriği limitlerin dışında ise), standart burada *iteratif bir süreç* öngörür. Yani modele tekrar geliştirme aşamasına dönülerek düzeltmeler uygulanır ve V&V tekrar gerçekleştirilir. Bu döngü, model kabul edilebilir seviyeye gelene kadar sürer. Bu belki proje takviminde gecikmeye yol açabilir, ancak uzun vadede hatalı veya adaletsiz bir sistemi sahaya sürmenin getireceği maliyetlerden kaçınmak için gereklidir. Eğer model hiçbir şekilde istenen kriterleri sağlayamıyorsa, proje yöneticilerinin alternatif çözümleri değerlendirmesi (farklı bir model türü, farklı veri, insan müdahalesinin arttırılması vb.) gerekir. ISO/IEC TS 12791, bu kararı destekleyecek biçimde belgelendirme yapılmasını da önerir: Yani test sonuçları, uygulanan önyargı ölçümleri, bulunan problemler ve alınan aksiyonlar detaylı

şekilde kayıt altına alınmalıdır. Bu hem ileride yapılacak denetimler için kanıt teşkil eder, hem de ekip içinde öğrenmeyi sağlar.

- **Kullanıcı veya Uzman Geri Bildirimi ile Test:** Bazı durumlarda, teknik ölçütlerin ötesinde, hedef kitleden veya konu uzmanlarından geribildirim almak da önyargı tespiti için değerlidir.

Örneğin, model pilot olarak kullanıma sokulup gerçek kullanıcıların tepkileri ölçülebilir. Eğer belirli bir grup kullanıcı sistemin önerilerinden bariz şekilde memnun değilse veya şikâyet iletiyorsa, bu da kalitatif bir önyargı göstergesi olabilir.

Standart, kullanıcı kabul testlerinde çeşitliliğe önem verilmesi gerektiğini not edebilir; farklı geçmiş ve özelliklere sahip test kullanıcıları seçmek, önyargının pratikteki etkisini gözlemlemeye yardımcı olur.

Doğrulama ve geçerleme aşaması, *"ayıbın örtüden kalktığı"* yerdir. Eğer önceki aşamalarda bir şey gözden kaçmışsa, burada ortaya çıkar. ISO/IEC TS 12791'nin katkısı, bu aşamada önyargı tespitini sistematik hale getirmesidir. Gerek veri kümesi analizleri gerek model metrikleri sayesinde, önyargıyı nicel olarak ortaya koymak mümkün olur. Bu da subjektif tartışmalar yerine somut veriler üzerinden aksiyon almayı kolaylaştırır.

Örneğin, "modelimizde biraz cinsiyet yanlılığı olabilir" demek yerine, "modelin erkeklere pozitif karar oranı %60, kadınlara %40; hedefimiz %50-%50'ye çekmek" diyebilir hale geliriz. Böylelikle tartışma, sorunun varlığı noktasından çözüm stratejisine kayar.

4. Kullanım, İzleme ve Sürekli İyileştirme (Deployment & Monitoring) Aşamasında Önyargının Takibi

Model başarıyla doğrulanıp kabul edildikten sonra gerçek dünyada kullanım (deployment) safhasına geçer. Ancak standart burada durmaz; tam aksine, modelin canlı kullanımı sırasında ve sonrasında da önyargının takip edilmesi gerektiğinin altını çizer. Bu aşamada önemli olan noktalar şunlardır:

- Sürekli İzleme ve Periyodik Değerlendirme: Bir YZ modeli saha ortamında çalışmaya başladıktan sonra, zaman içinde performansı ve önyargı davranışı değişebilir. Veri dağılımları kayabilir (concept drift), kullanıcı profilleri değişebilir veya modelin kararları beklenmedik geri beslemelere yol açabilir. ISO/IEC TS 12791, bu nedenle bir sürekli izleme mekanizması kurulmasını önerir. Bu mekanizma, modelin önemli

metriklerini belirli aralıklarla (ör. aylık, çeyreklik) yeniden hesaplamalı ve özellikle önyargı ile ilgili göstergeleri kontrol etmelidir.

Örneğin, bir kredi skorlama modeli ilk başta kabul edilebilir bir sapma ile çalışırken, altı ay sonra ekonomik koşulların değişmesiyle bir grup için hatalı sonuçlar üretmeye başlayabilir. Eğer kurum periyodik olarak modelin farklı gelir gruplarındaki isabet oranlarını izlerse, bu tür bir kaymayı erkenden tespit edip modeli güncelleyebilir.

Standart bu yaklaşımı "sürekli doğrulama (continuous validation)" olarak tanımlar ve ISO/IEC 5338'de de bahsedildiği üzere, YZ sistemlerinin yaşam döngüsünün doğrusal değil döngüsel (sürekli geri beslemeli) olduğuna dikkat çeker.

- **Dışsal Değişikliklerin Takibi (External Change):** Model ortamında meydana gelen büyük değişiklikler, önyargı profilini etkileyebilir. Bu değişiklikler dışsal (harici) faktörler olabilir: Yeni bir yasa yürürlüğe girebilir (örneğin belirli bir özelliği karar verme sürecinde kullanmak yasaklanır), toplumsal eğilimler değişebilir, rakip sistemlerin ortaya çıkışı kullanıcı beklentilerini dönüştürebilir. ISO/IEC TS 12791, böyle durumlar için bir alarm sistemi kurulmasını salık verir.

Örneğin, eğer işe alım alanında cinsiyet eşitliğine dair yeni bir regülasyon geldiyse, mevcut modelin bu regülasyona uygunluğu derhal gözden geçirilmelidir. Ya da modelin eğitildiği veri artık güncelliğini yitirdiyse (pandemi öncesi veriler pandemi sonrası dönemi iyi temsil etmeyebilir), bu bir dışsal değişiklik olarak kabul edilip model yeniden eğitilmelidir.

Standart, organizasyonların bir *değişiklik yönetimi* süreci içinde YZ modellerini de ele almasını önerir. Yani önemli dış olayların bir listesi tutulur ve her olay gerçekleştiğinde model üzerinde bir değerlendirme tetiklenir.

- **Geribildirim Toplama ve İnsan Gözetimi:** Model çalışırken, gerçek kullanıcılardan veya modelin kararlarını değerlendiren insan operatörlerden geribildirim toplamak, önyargı tespitinin devamlılığı için değerlidir. ISO/IEC TS 12791, YZ sistemlerinin mümkün olduğunca insan gözetimi altında işletilmesi gerektiğini, özellikle yüksek riskli kararlarda bir insan onay mekanizmasının yer almasının önyargı kaynaklı hataları yakalamaya yardımcı olacağını belirtir.

Örneğin, banka kredi reddi alan müşterilerin itirazlarını dinleyen bir ekip kurmak, modelin haksızlık yaptığı durumları ortaya çıkarabilir. Bu itirazların sistematik analizi, modelin hangi durumlarda potansiyel önyargı sergilediğine dair kalitatif içgörü sunar.

Standart, kuruluşların bu tür geribildirimleri kayıt altına alıp belirli periyotlarla analiz etmesini önerir. Belki de, gelen şikayet ve itirazlar bir sonraki model güncellemesinde dikkate alınacak yeni gereksinimler doğuracaktır.

- **Model Güncellemeleri ve Yeniden Eğitim**: İzleme sürecinde bir önyargı tespit edilirse, modelin yeniden eğitilmesi veya güncellenmesi gündeme gelir. Bu, yaşam döngüsünde tekrar tasarım/geliştirme ve doğrulama aşamalarına dönmek demektir. Standardın getirdiği bir kavram da budur: YZ projeleri *bitmez*, sürekli evrilir.

Örneğin, periyodik değerlendirmede modelin belirli bir etnik gruba %15 daha az doğrulukla çalıştığı görüldüyse, bu kabul kriterlerinin dışına çıkıyorsa model yeniden ele alınmalıdır. Belki bu grup için ek veri toplanıp veri kümesine eklenebilir, belki modelin üst parametreleri güncellenebilir veya bambaşka bir algoritmaya geçilebilir. Kritik olan nokta, kuruluşun bu değişikliği yapmaya hazır olmasıdır.

ISO/IEC TS 12791, bu amaçla bir model bakım planı oluşturulmasını ve gerektiğinde müdahale edecek sorumluların önceden belirlenmesini önerir.

Örneğin, model geliştiricilerinden oluşan küçük bir "görev gücü" aktif hizmet esnasında da modelden sorumlu tutulabilir ve metrikler bozulduğunda aksiyon alması beklenir.

- **Kullanım Kılavuzu ve Sınırlar:** Standardın üzerinde durduğu başka bir husus da, modeli kullananların sistemin sınırlarını ve önyargı risklerini bilmesi gerektiğidir. Yani YZ sisteminin bir *kullanım kılavuzu* veya *etiket bilgisi* olmalıdır.

Örneğin, bir yüz tanıma yazılımı piyasaya sunuluyorsa, üretici firma yazılımın koyu tenli bireylerde %X daha yüksek hata payı olduğunu, bu nedenle kritik güvenlik uygulamalarında dikkatli olunması gerektiğini belgelerinde belirtmelidir. Bu şeffaflık, kullanıcıların modele aşırı güven duyarak hatalı kullanım yapmasının önüne geçer (otomasyon yanlılığı olarak bilinen, insanların YZ çıktısına gereğinden fazla güvenmesi durumu böylece azaltılabilir). Eğer bir modelin bilinen bazı önyargıları varsa, bunlar açıkça ifade edilmeli ve kullanıcı tarafında telafi edici prosedürler konmalıdır (örneğin, model bir grubun lehine hatalı eğilimliyse, kullanıcı kararı verirken bunu dengeleyici manuel kontroller yapabilir).

Özetlemek gerekirse, kullanım ve izleme aşaması, YZ sisteminin *gerçek dünya sınavı*dır. ISO/IEC TS 12791, bu sınavın sürekli gözetim ve bakım ile geçilmesini salık verir. Zira hiçbir model durağan bir dünyada yaşamaz; toplumsal dinamikler, veri akışları, hatta sistemin kendi etkileşimleri zamanla değişir ve yepyeni önyargı türleri ortaya çıkabilir.

Örneğin, modelin kullanımı belli bir grubu dışladığı için o grup sistemi kullanmayı bırakabilir, bu da veride yeni bir dengesizlik yaratır (kendini gerçekleştiren önyargı döngüsü). Bu tür durumları ancak dikkatli bir izleme ve esnek bir güncelleme yaklaşımıyla engelleyebiliriz.

Standart, YZ sistemlerinin yaşam döngüsü boyunca sorumluluk gerektiren varlıklar olduğunu, tıpkı bir ürünün bakım gerektirmesi gibi model bakım ve önyargı bakımının da işin parçası olduğunu hatırlatır.

5. Sistem Sonlandırma (Disposal) Aşamasında Dikkat Edilecekler

Her YZ sistemi sonsuza dek çalışmaz; teknolojinin yenilenmesi, ihtiyaçların değişmesi veya performansın yetersiz kalması gibi nedenlerle bir model kullanımdan kaldırılabilir (decommissioning). ISO/IEC TS 12791, yaşam döngüsünün bu son aşamasında da bazı hususlara dikkat çekmektedir:

- **Bilgi Birikiminin Aktarılması**: Bir model emekliye ayrılırken, onunla ilgili öğrenilen derslerin ve özellikle önyargı ile ilgili deneyimlerin kaydedilip yeni projelere aktarılması değerlidir. Standart, projeden elde edilen bulguların (hangi önyargılarla karşılaşıldığı, nasıl çözüldüğü, nelerin işe yarayıp yaramadığı gibi) kurumsal hafızaya işlenmesini önerir. Bu sayede, yeni geliştirilecek YZ sistemlerinde benzer hataların tekrarı önlenebilir.

Örneğin, bir konuşma tanıma sistemi belirli aksanlarda başarısız olduğu için rafa kaldırıldıysa, yeni alınacak sistemde bu aksan çeşitliliğine özel önem verilmesi gerektiği not edilmelidir.

- **Veri ve Modelin Sorumlu İmhası veya Arşivlenmesi:** Bir sistem devreden çıkarılırken, onun eğitildiği verinin ve model parametrelerinin güvenli bir şekilde saklanması veya imha edilmesi de önyargı perspektifinden önemlidir. Çünkü ileride aynı veri veya model parçaları başka projelerde kullanılabilir. Eğer eski model önyargılı ise ve bu durum bilinmezse, yeni projeye de bu önyargının taşınması riski vardır. Dolayısıyla, standart bir model sonlandırılırken etiketleme yapılmasını (bu model neden sonlandırıldı, bilinen problemleri neydi) tavsiye edebilir. Ayrıca, yasal açıdan da eğer modelin kararlarından etkilenen kişiler varsa, belirli bir süre kayıtların tutulması gerekebilir (mesela haksızlığa uğradığını düşünen bir birey, süreçle ilgili inceleme talep edebilir). Bu nedenle modelin çıktıları, karar kayıtları vs. belirli bir

süre için arşivlenmelidir. Bu arşivleme sürecinde kişisel verilerin korunması, anonimleştirme gibi konular da göz önünde bulundurularak, gelecekteki analizler için yeterli verinin tutulması sağlanmalıdır.

- **Alternatif Süreçlere Geçiş:** Bir YZ sistemi sonlandırıldığında, onun yerine manuel bir süreç veya farklı bir sistem gelebilir. Standart, bu geçiş döneminde önyargı açısından bir *geri tepme (backfire)* olmaması için uyarıda bulunabilir.

Örneğin, bir şirket YZ tabanlı bir seçme sisteminden tekrar insan değerlendirmesine dönüyorsa, insan kaynakları ekiplerine önyargı eğitimi verildiğinden emin olmak gerekir; aksi halde YZ sisteminde bir nebze kontrol altına alınmış önyargı, insan değerlendirmesinde daha da şiddetli dönebilir. Yani, insan veya başka sistemler de kusursuz değildir; dolayısıyla "YZ'yi kaldırdık, sorun çözüldü" diye düşünmek yerine, yerine gelen süreçte de adalet prensiplerinin uygulandığından emin olunmalıdır.

- **Kamuya Açıklama ve Hesap Verebilirlik:** Eğer ilgiliyse, bir YZ sisteminin kullanımdan kaldırılması kamuoyuna veya düzenleyici kurumlara açıklanabilir. Özellikle kamusal alanda kullanılan veya tartışma konusu olmuş sistemlerde, sonlandırma kararının nedenleri (örneğin önyargı endişeleri) şeffafça paylaşılabilir. Bu, kurumun hesap verebilirliğini ve toplumun güvenini artırır. Standart doğrudan bunu dikte etmese de, şeffaflık ve hesap verebilirlik ilkelerini genel hatlarıyla destekler.

Sonlandırma aşaması, YZ etiği tartışmalarında sık ele alınan bir konu olmasa da, TS 12791 bu boşluğu doldurarak yaşam döngüsünün son halkasını tamamlar. Bu da bize gösteriyor ki, *önyargı ile mücadele bir süreçtir ve sistemin ömrüyle sınırlıdır*. Sistem hayatını tamamladığında, mücadele de biter ancak ondan alınan dersler devrede kalır. Standart, bu bütüncül bakış açısıyla kurumların her proje döngüsünde bir adım daha öğrenmesini ve olgunlaşmasını teşvik ediyor.

Önyargı Azaltma Teknikleri ve Yöntemleri

ISO/IEC TS 12791:2024, sadece süreç bazında ne yapılacağını anlatmakla kalmaz, aynı zamanda pratikte kullanılabilecek belli başlı **önyargı azaltma tekniklerine** de değinir. Standart içerisinde özel bir bölüm, önyargıyı ele almak için uygulanabilecek algoritmik ve veri odaklı yöntemleri özetlemektedir (Bölüm 6: Techniques to address unwanted bias). Burada amaç, proje ekiplerine ellerindeki potansiyel araçlar hakkında fikir vermek ve gerekiyorsa

derinlemesine araştırma yapmaları için yön göstermektir. Bu bölümde, standartta bahsi geçen teknik kategorilerini ve örnek yöntemleri inceleyeceğiz.

Makine öğrenmesi modellerinde önyargı azaltma stratejileri, genellikle ön işlem, iç işlem (eğitim sırasında) ve son işlem olmak üzere üç aşamada uygulanır. Yukarıdaki diyagram bu üç kategoriye ait yaygın teknikleri özetlemektedir.

Örneğin, ön işlem aşamasında veri kümesine yeniden örnekleme veya ağırlıklandırma uygulanarak dengesizlikler giderilebilir (sarı blok); iç işlem aşamasında modele bir düzenleme terimi eklemek veya çekişmeli (adversaryal) ağlar kullanmak yoluyla modelin korunan bir özelliği öğrenmesi engellenebilir (beyaz blok); son işlem aşamasında ise model çıktılarına eşik ayarlamaları yapılarak veya sonuçlar yeniden etiketlenerek adalet metriksleri sağlanabilir (yeşil blok).

Bu teknikler, literatürde pre-processing, in-processing ve post-processing algoritmaları olarak da anılmaktadır ve TS 12791'de sunulan yaklaşım da bu genel çerçeveyi yansıtır.

- **Veri Odaklı Teknikler (Pre-Processing):** Standart, verideki önyargıyı azaltmak için uygulanabilecek yöntemleri vurgular (Bölüm 6.3 Data techniques). Bu kapsamda en yaygın yaklaşım, *veri dengeleme (balancing)* veya *yeniden ağırlıklandırma (reweighting)* yöntemleridir. Yukarıda da bahsettiğimiz gibi, eğer veri kümesinde bazı gruplar az temsil ediliyorsa örnek sayıları eşitlenebilir. Reweighing tekniği, her bir veri noktasına ait olduğu (grup, etiket) kombinasyonuna göre bir ağırlık vererek model eğitimi sırasında adaletsizliği giderir.

Örneğin, azınlık gruptan gelen veriler eğitim sırasında daha fazla ağırlık alır, böylece modelin hata fonksiyonunda o verilerin etkisi artar ve model onların üzerine daha fazla eğilir.

Örnek kaldırma (deletion) veya ekleme: Bazı durumlarda, verideki önyargının kaynağı belirli veriler olabilir.

Örneğin, geçmişte ayrımcı kararlar içeren kayıtlar veride bulunuyorsa, bunların çıkarılması düşünülebilir. Ya da veride her birey için "yaş" bilgisi varsa ve yaşa dayalı bir önyargı istenmiyorsa, yaş verisi tamamen silinebilir veya genelleştirilebilir (örn. yaş yerine yaş aralığı kullanmak).

Veri sentezi ve veri artırma: Temsil eksikliği olan gruplar için yapay veriler üretmek de bir yoldur.

Örneğin, açık kaynak bazı araçlar kullanılarak farklı ten renklerine sahip sentetik yüz görüntüleri oluşturulup veri kümesine eklenebilir.

Öznitelik gizleme veya dönüştürme: *Disparate impact remover* gibi yöntemlerle, veri içindeki hassas özellikler modifiye edilir. Mesela, bir özelliğin dağılımı tüm gruplar için benzer olacak şekilde dönüşümler uygulanabilir. Ya da hassas özelliklerin veri kümesinden tamamen kaldırılması ("algılanamaz kılma") sağlanabilir. Ancak, veri odaklı tekniklerin dezavantajı, verinin yapısını değiştirdikleri için bazen modelin performansını da düşürebilmeleridir. Bu nedenle, ISO/IEC TS 12791, bu yöntemlerin dikkatli ve test edilerek uygulanmasını, gerekirse birden fazla yöntemin kombinasyonundan faydalanılmasını tavsiye eder. Son olarak, anlam kaybı ve yan etki riskine de değinilir: Veriyi değiştirirken, orijinal problemin anlamını çok bozmamaya dikkat etmek gerekir (örneğin tüm ekstrem değerleri atmak, belki de önemli bir alt grubun tamamen kaybına yol açabilir).

- **Algoritmik ve Eğitim Sırasındaki Teknikler (In-Processing):** Önyargıyı modelin öğrenme sürecinde engellemeye yönelik çeşitli algoritmalar mevcuttur. En bilinen yöntemlerden biri, adversaryal (düşmancıl) ağların kullanımıdır. *Adversarial debiasing* olarak bilinen teknikte, modele asıl görevini yaptırırken paralel bir ağ da modelin hassas özelliği tahmin etmesine çalışır; ana model, bu düşmancıl ağı şaşırtacak şekilde eğitilir.

Örneğin, bir metin sınıflandırma modeline eş zamanlı olarak metindeki yazarın cinsiyetini tahmin etmeye çalışan bir alt model eklenir. Ana model, cinsiyet bilgisini veremeyecek şekilde optimize edilirken asıl görevindeki performansını korumaya çalışır. Bu sayede model istemeden cinsiyet bilgisini kodlamamış olur.

Bir diğer yaklaşım, düzenlileştirici (regularizer) eklemektir.

Örneğin prejudice remover (önyargı giderici) düzenlileştirici, kayıp fonksiyonuna bir ceza terimi ekleyerek belirli bir ayrımcılık ölçütünün azalmasını sağlar. Diyelim ki elimizde "erkek – kadın" grupları için doğru pozitif oran farkı gibi bir metrik var; bu fark büyüdükçe ceza artar, model eğitimi bu cezayı minimize etmeye çalışırken adil bir çözüme yakınsar.

Bunun dışında, eşitlik kısıtları (fairness constraints) uygulamak da popüler bir yaklaşımdır: Model eğitimi bir optimizasyon problemi olduğundan, çözüme bazı adalet kısıtları eklenebilir.

Örneğin "Gruplar arasındaki pozitif karar oranları arasındaki fark 0'a yakın olmalıdır" şeklinde bir kısıt konulup, model bu kısıt altında en iyi doğruluğu sağlamaya çalışabilir. ISO/IEC TS 12791, teknik detaya girmeden bu tür yöntemlerin varlığına atıf yaparak ekiplerin literatürü takip etmesini önerir. Bu sayede, geliştiriciler bu alandaki açık kaynak kütüphaneleri (IBM'in YZ Fairness 360'ı gibi) veya akademik yayınları standardın referansıyla keşfetmeye yönlendirilir.

In-processing tekniklerin avantajı, model performansını çok düşürmeden adalet sağlayabilmeleridir; dezavantajı ise genellikle daha karmaşık bir eğitim süreci ve uzmanlık gerektirmeleridir.

- **Çıktı Üzerindeki Teknikler (Post-Processing):** Model eğitimi tamamlandıktan ve bir tahmin ürettikten sonra, bu tahmin üzerine uygulanacak işlemlerle önyargı giderilmeye çalışılabilir. TS 12791'de doğrudan bu kavram adıyla geçmese de, Bölüm 6.2'nin devamında ve örneklerde bu yaklaşımdan da bahsedildiği anlaşılmaktadır. Post-processing yöntemlerinin en tanınanlarından biri eşitlenmiş fırsatlar (equalized odds) sonrası işleme yöntemidir. Bu teknik, modelin çıktısını, her bir grup için hataları eşitleyecek şekilde ayarlamayı hedefler. Pratikte, model çıktı skoru belirli aralıklardaysa ve grup avantajı gözleniyorsa, bazı skorlar bilinçli olarak pozitiften negatife (veya tersi) çevrilebilir. Calibrated equalized odds ise benzer bir yaklaşımı kalibre edilmiş skorlara uygular, yani modelin olasılık çıktıları üzerinde optimizasyon yaparak gruplar arasında denge sağlar. Bir başka ilginç yöntem de "reject option" (reddetme opsiyonu) sınıflandırmasıdır. Bu yöntemde, modelin en kararsız kaldığı alanlarda (örneğin çıktının olasılığı %45-55 arasıysa) pozitif kararı genelde dezavantajlı gruba, negatif kararı avantajlı gruba vererek bir dengeleme yapılır.

Örneğin, tam sınırda kalan kredi başvurularında, sistemik olarak dezavantajlı olduğu bilinen gruptaki başvuru sahiplerine onay verilirken, avantajlı gruptakilere ret verilir. Bu şekilde, genel başarıyı çok düşürmeden, dezavantajlı grubun lehine bir düzeltme yapılır. Post-processing yöntemlerin en büyük artısı, modeli "kara kutu" haline getirmeden dışarıdan düzeltme olanağı sunmalarıdır. Yani elde var olan bir model, herhangi bir değişiklik yapılmadan üzerine bir katman olarak adalet düzeltmesi konabilir. Bu, özellikle modelin değiştirilmesinin zor veya maliyetli olduğu durumlarda (örn. hazır bir API kullanılıyorsa) işe yarar.

Dezavantajı ise, tekil karar bazında tutarsızlığa yol açabilmesi ve karar mantığını biraz daha az şeffaf hale getirebilmesidir (çünkü kullanıcılar ya da yöneticiler, neden bazı durumlarda modelin çıktısının değiştirildiğini tam anlamayabilir; bunu iyi ifade etmek gerekir).

- **Önceden Eğitilmiş Modellerin Kullanımı**: ISO/IEC TS 12791, günümüzde sıkça başvurulan önceden eğitilmiş modellerin (pre-trained models) kullanımına ayrı bir başlık ayırmaktadır. 6.2.2 maddesinde belirtildiği üzere, eğer bir proje kapsamında hazır bir makine öğrenmesi modeli, özellikle de büyük ölçekli bir model (ör. büyük dil modelleri, yüz tanıma modelleri vb.) kullanılacaksa, bunun taşıyabileceği önyargılara dikkat edilmelidir. Önceden eğitilmiş modeller genellikle devasa veri kümeleri üzerinde eğitilir ve bu veri kümelerinin içerdiği önyargılar, modelin içine işlenmiş olabilir.

Örneğin, belli bir dil modeli, internetteki metinleri tarayarak öğrenmişse, toplumdaki cinsiyetçi veya ırkçı eğilimleri de kapmış olabilir.

Standart, bu nedenle hazır modelleri kullanmadan önce mümkünse bir önyargı denetiminden geçirmeyi önerir. Bazı büyük modeller için akademik çalışmalar, o modelin bilinen önyargı profilini yayınlamaktadır; bunlardan faydalanılabilir. Eğer bir hazır model kullanılacaksa, proje ekibi bu modelin geliştirici belgelerini incelemeli ve orada adalet/önyargı ile ilgili açıklamalar var mı bakmalıdır. Gerekirse, hazır modele bir *ince ayar (fine-tuning)* yaparak, kendi veri kümesine uygun hale getirirken önyargı azaltma teknikleri entegre edilebilir. Hatta bazı durumlarda, önceden eğitilmiş modelin son katmanlarına eklenen bir adversaryal bileşen ile bu modelden gelen önyargı arındırılabilir. Eğer hazır model çok karmaşık ve değiştirilemez haldeyse, o zaman çıktılarına *post-processing* yöntemlerini uygulamak kalan tek çare olabilir (mesela bir bulut servisi olarak alınan modelin çıktılarını izleyip, önyargı tespit edilen yerlerde otomatik düzeltme yapmak gibi).

Yukarıdaki tekniklerin her birinin uygulanabilirliği, eldeki probleme, veri yapısına ve ekibin uzmanlığına bağlıdır. ISO/IEC TS 12791, bir reçete kitabı şeklinde spesifik algoritmalar dayatmaz; bunun yerine genel kategorileri anlatıp bunların yaşam döngüsündeki yerlerini gösterir. Böylece, ekipler kendi araç setlerini ve bilgi birikimlerini kullanarak uygun kombinasyonu seçebilirler.

Standart ayrıca, bu tekniklerin kısıtları ve olası yan etkileri konusunda da uyarılarda bulunur (özellikle Annex/Ek B'de, farklı kullanıcı tiplerine olası etkilerden bahsedildiği belirtilmiştir). Bir tekniğin bir gruba fayda sağlarken başka bir gruba zarar vermesi ihtimali, veya genel doğruluğu düşürmesi gibi trade-off'lar her zaman değerlendirilmelidir. Bu noktada,

kuruluşların önceliklerini netleştirmesi gerekir: Adalet mi daha öncelikli, yoksa toplam performans mı? Bu tercih, kullanılacak tekniklerin dozunu belirler. Örneğin, bir sağlık teşhis modelinde toplam doğruluk kritiktir ama aynı zamanda etnik eşitlik de önemlidir; bu ikisini dengelemek için belki çok agresif bir önyargı azaltma uygulanmaz (çünkü doğruluk da hayati olabilir), ama göz ardı da edilmez.

Son olarak, TS 12791'deki teknik önerilerin esnek ve gelişime açık olduğunu vurgulamak gerekir. YZ alanında önyargı ve adalet konusu çok aktif bir araştırma alanıdır; sürekli yeni yöntemler, yeni metrikler ortaya çıkmaktadır. Standart, 2024 itibarıyla bilinen iyi uygulamaları kapsasa da, kullanıcıların gelişmeleri takip etmesi önemlidir.

Dağıtık YZ Sistemlerinde Önyargı Yönetimi

Standardın dikkat çekici bölümlerinden biri de "Dağıtık YZ sistemi yaşam döngüsünde önyargının ele alınması" konusudur. Burada kastedilen, bir YZ sisteminin geliştirilmesi ve kullanımının tek bir elde olmadığı, birden fazla organizasyon, ekip veya paydaş arasında paylaştırıldığı durumlardır. Modern YZ projelerinde sıkça, veri bir kaynakta, model başka bir elde, uygulama ise başka bir platformda olabilir. Bu dağıtık yapı, önyargı yönetimini karmaşıklaştırır çünkü sorumluluklar ve kontrol mekanizmaları dağılmıştır. ISO/IEC TS 12791, bu durumlar için de yol gösterici ilkeler sunar:

- **Sorumlulukların Belirlenmesi**: Dağıtık bir sistemde, önyargı yönetimi sorumlulukları açıkça tanımlanmalıdır. Örneğin, bir şirket dış bir servis sağlayıcıdan makine öğrenmesi modeli alıp kendi verisinde kullanıyorsa, verideki önyargılardan kim sorumlu, modeldeki önyargılardan kim sorumlu önceden mutabık kalınmalıdır. Standart, olası önyargı risklerinin paydaşlar arasında iletişimle paylaşılmasını ve herkesin kendi katkı alanında gerekli tedbirleri almasını önerir. Bu, hukuki açıdan da önemlidir; zira önyargı nedeniyle bir yaptırımla karşılaşıldığında, tedarikçi mi yoksa kullanıcı mı sorumlu tartışması yaşanmamalıdır. En ideali, proaktif olarak birlikte çalışıp sorunu baştan engellemektir.
- **Kalite ve Etik Standartlarının Uyumlaştırılması**: Farklı organizasyonların projeye dahil olduğu senaryolarda, herkesin aynı kalite standartlarına ve etik ilkelere bağlı olması gerekir. ISO/IEC TS 12791, dağıtık ortamlarda bir çeşit ortak sözleşme/çerçeve oluşturulmasını tavsiye edebilir.

Örneğin, eğer bir YZ ürünü konsorsiyum şeklinde geliştiriliyorsa, konsorsiyum üyeleri proje başlangıcında bir etik sözleşme imzalayarak, veri paylaşımı, önyargı denetimi, şeffaflık gibi konularda asgari standartları kabul edebilir. Bu, tüm tarafların beklentilerini hizalar ve zayıf halkaların oluşmasını önler.

- **Veri Paylaşımı ve Birleştirme (Federated / Distributed Learning):** Dağıtık sistemlerden bahsederken, federatif öğrenme veya dağıtık öğrenme gibi yaklaşımlar da akla gelir. Bu yaklaşımlarda veri merkezi bir yerde toplanmaz; model, verinin ayağına giderek yerinde öğrenir ve sadece gerekli parametreler merkezde birleşir. Bu, gizlilik için faydalı olsa da, önyargı izlemini zorlaştırabilir. Her düğüm (node) kendi lokal verisinin önyargısını taşıyabilir ve merkeze sadece model parametresi geldiği için, parametrelerdeki bir kaymayı doğrudan önyargıya bağlamak güç olabilir. TS 12791, dağıtık öğrenim senaryolarında her bir düğümde önyargı kontrolü yapılması ve merkezde de bunların bütünleştirilerek incelenmesi gerektiğini belirtebilir.

Örneğin, federatif bir sağlık modeli düşünelim: Farklı hastanelerde lokal modeller eğitiliyor ve merkezde birleştiriliyor. Her hastane, kendi verisinde cinsiyet dağılımını ve modele yansıyan hataları kontrol etmeli, bu veriyi merkezle paylaşmalı.

Merkez de modelin global performansında bir grubun bariz kötüleştiğini görürse, hangi hastaneden kaynaklandığını saptayabilmeli. Bu, dağıtık ortamda önyargı tespiti için şeffaflık gerektirir. Standart, taraflar arasında bu tip veri/metrik paylaşım protokollerinin oluşturulmasını teşvik eder.

- **İletişim ve Eğitim:** Birden fazla kurumun ortaklaşa çalıştığı projelerde, ekipler arası iletişim de kritik hale gelir. Herkesin önyargı konusunda aynı bilinç ve bilgi düzeyine sahip olması beklenemez; bu yüzden, lider konumdaki kurumlar veya proje koordinatörleri, diğerlerine önyargı eğitimi verebilir, atölyeler düzenleyebilir. ISO/IEC TS 12791, belki doğrudan bunu söylemez ancak *farkındalık yaratma* konusunda genel mesajlar verir. Özellikle tedarik zinciri boyunca (supply chain) bir YZ sistemi geliştiriliyorsa, zincirin her halkasının hassas olması gerekir.

Örneğin, bir firma bir YZ modelini bir alt yükleniciye geliştirtip sonra son ürüne entegre ediyorsa, alt yüklenicinin süreçlerini denetlemeli veya en azından eğitimli olduğundan emin olmalıdır.

- **Dağıtık Karar Alma ve Çakışmalar:** Dağıtık sistemler bazen birden fazla modelin veya bileşenin birlikte çalışması anlamına da gelir.

> *Örneğin, bir karar süreci, biri bir kurumun geliştirdiği, diğeri başka bir kurumun geliştirdiği iki farklı YZ modülünün sonuçlarına dayanarak ilerliyor olabilir. Bu durumda, modüllerden birinin önyargılı olması genel sonucu da etkiler.*

Standart, bu gibi durumlarda bütünsel bir değerlendirme yapılmasını önerir. Yani her modül tek tek adil olsa bile, birleşik sistem adil olmayabilir veya tersi, modüllerden biri sorunluysa bütünü bozabilir. Bu nedenle entegrasyon testlerinde, tüm sistemi uçtan uca simüle edip önyargı analizi yapmak gerekir. Dağıtık modüllerin arayüz tanımlarında, hassas değişkenlerin geçişine dair kurallar konabilir (belki bir modül diğerine cinsiyet bilgisini hiç geçirmez gibi).

- **Regülasyonlara Uyum:** Birden çok ülke veya bölge aktörlerinin bulunduğu projelerde, farklı yasal gereksinimler devreye girebilir. Standart, bu konuda spesifik olmasa da, dağıtık yapıda en sıkıntılı konulardan birinin de bu olduğunu bilir.

> *Örneğin, AB içinde geliştirilen bir model, ABD'de kullanılıyorsa, AB'nin YZ Tüzüğü beklentileri ile ABD'nin eşitlik yasaları arasında farklar olabilir. Proje, en yüksek standartları hedefleyerek herkesi buna uydurmalıdır ki zayıf halka oluşmasın.*

Kısaca, dağıtık YZ sistemlerinde önyargı yönetimi, kolektif bir çaba gerektirir. ISO/IEC TS 12791, bu bölümde belki de bir nevi "işbirliği manifestosu" sunarak, önyargı ile mücadeleyi ortak bir sorumluluk olarak tanımlar. Tek bir aktörün mükemmel olması yetmez; zincirin tüm halkaları güçlü olmalıdır. Bu aynı zamanda politikacılar ve düzenleyiciler için de bir mesaj içerir: Uluslararası veya sektörler arası işbirliklerinde, standartlara ortak uyum teşvik edilmelidir. Zira bir firma kendi içinde ne kadar düzgün çalışsa da, tedarik ettiği teknoloji önyargılıysa sonuç değişmez. Bu nedenle, TS 12791 gibi standartlar, tüm ekosistemi kapsayan bir çerçeve sunarak farklı aktörlerin aynı sayfada buluşmasını sağlar.

Sonuç ve Değerlendirmeler

ISO/IEC TS 12791:2024, YZ sistemlerinde istenmeyen önyargıyla mücadele konusunda kapsamlı ve yapısal bir yaklaşım ortaya koymaktadır. Raporda incelediğimiz üzere, standart bir YZ projesinin başlangıcından sonuna dek her evresinde alınması gereken önlemleri, dikkat edilmesi gereken noktaları ve uygulamaya konabilecek teknik yöntemleri detaylandırmıştır. Bu bütüncül yaklaşım, YZ etiği ve güvenilirliği alanında önemli bir boşluğu

doldurmaktadır: İlk kez uluslararası düzeyde, önyargının **nasıl azaltılacağına** dair somut adımlar tanımlanmıştır.

Standarttan çıkan temel mesajları birkaç ana başlıkta özetleyebiliriz:

- **Proaktif Planlama:** Önyargının ele alınması, daha proje fikri oluşurken başlıyor. Paydaş analizi, gereksinim tanımı ve risk değerlendirmesi aşamalarında önyargıyı masaya yatırmak, sonradan telafisi zor hataların önüne geçiyor. Kurumlar, YZ projelerinde adeta bir "*etik kontrol listesi*" ile yola çıkmalı; TS 12791 bu listeyi büyük ölçüde sağlamış durumda.
- **Yaşam Döngüsü Boyunca Süreklilik**: Önyargı yönetimi, tek seferlik bir müdahale değil, sürekli bir süreç. Model geliştirilirken de test edilirken de kullanılmaya başladıktan sonra da bitmeyen bir ödev. Bu bakış açısı, YZ sistemlerine adeta *canlı bir organizma* gibi yaklaşmayı gerektiriyor – büyümesi, adaptasyonu ve bazen emekliliği var. Standart, bu bakımdan kurum kültüründe bir dönüşüm öneriyor: "Yap, bitir" yerine "yap, izle, düzelt, yinele".
- **Teknik ve Yönetsel Önlemlerin Birleşimi:** Sadece algoritmalarla veya sadece politika belgeleriyle önyargıyı yenemeyiz; ikisinin entegre çalışması gerek. TS 12791, teknik araç kutusunu (yeniden ağırlıklandırma, adversaryal öğrenme, vs.) yönetsel araç kutusuyla (sorumluluk atama, tedarik şartları, vb.) bir araya getiriyor. Özellikle politika yapıcılar ve yöneticiler için, teknik ayrıntıları anlaşılır ilkelere dönüştürerek sunması önemli. Mesela bir CEO, bu standart sayesinde teknik ekibine "modeli adil yapın" demek yerine, hangi somut adımları beklediğini ifade edebilir; bir kamu düzenleyicisi, firmalardan ne tür dokümantasyon ve önyargı test raporları isteyeceğini bu referansa dayanarak belirleyebilir.
- **Uyum ve Hesap Verebilirlik:** Standardın vurguladığı bir diğer nokta, önyargı konusunda şeffaflığın ve hesap verebilirliğin artmasıdır. Eğer kurumlar TS 12791'in önerilerini takip ederse, her kararda "neden böyle oldu" sorusuna cevap verebilecek verilere ve gerekçelere sahip olacaklar. Bu da kamuoyuna veya denetleyici kurumlara karşı güven yaratır. Bir ürün yöneticisi, ürününün önyargı analiziyle ilgili sorular geldiğinde, "uluslararası standartlara göre geliştirdik, işte kanıtı" diyebilme rahatlığına kavuşacaktır. Bu hem yasal riskleri azaltır (örneğin bir ayrımcılık davasında savunma imkanı verir) hem de etik sorumluluk bilincini gösterir.
- **Esneklik ve Güncellik:** Her ne kadar standardın yayın tarihi 2024 olsa da içinde barındırdığı yaklaşım geleceğe dönüktür. Hızla gelişen YZ dünyasında, TS 12791'in

yaşam döngüsü metodolojisi ve genel prensipleri yeni teknolojilere de uygulanabilir niteliktedir. Örneğin, 2025 ve sonrasında yaygınlaşan büyük dil modelleri veya YZ destekli karar sistemleri için de bu standarttan ilham almak mümkün olacaktır. Gerektiğinde, standardın teknik içerikleri güncellenebilir veya belirli alt alanlar için ek standartlar çıkabilir; ancak TS 12791'in omurgası olan *"bias risk management"* anlayışı kalıcı değerdedir.

Politika yapıcılar için, bu standart açıkça gösteriyor ki, YZ'deki önyargı sorunu çözümsüz değildir. Yeter ki sistematik bir çerçeve uygulansın. Kanun koyucular, şirketlerden ve kurumlarından bu standartta belirtilen uygulamaları takip etmelerini bekleyebilir, hatta bunu bir uyumluluk şartı haline getirebilirler.

Örneğin, bir kamu ihalesinde "YZ sistemleri ISO/IEC TS 12791'e uygun bias risk yönetimi ile geliştirilmelidir" gibi bir şart konulması, sektörde ciddi bir farkındalık ve seviye yükselmesi yaratacaktır.

Özel şirket yöneticileri için de standart büyük bir fırsattır: Rehber niteliğinde bir doküman olarak, karmaşık bir konuyu adım adım açıklıyor. Şirket içi YZ geliştirme süreçlerini bu standarda göre revize etmek, uzun vadede hem itibar hem de ürün kalitesi açısından getiriler sağlayacaktır. Özellikle müşteri odaklı sektörlerde, YZ kararlarının adil olması şirket markasının parçası haline gelebilir; TS 12791, bu hedefe ulaşmak isteyen yöneticiler için somut bir eylem planı sunuyor.

STK'lar ve sivil toplum için ise bu standart, ellerinde savunuculuk yapabilecekleri güçlü bir araç demektir. YZ sistemlerinin adil olmamasıyla mücadele eden bir STK, şirketlere veya devlet kurumlarına bu standardı hatırlatarak daha sorumlu davranmaları yönünde baskı yapabilir. Aynı şekilde, akademik veya bağımsız denetim kuruluşları, bir YZ sistemini değerlendirirken TS 12791'in kriterlerini kontrol listesi olarak kullanabilirler.

Sonuç olarak, ISO/IEC TS 12791:2024, YZ'de istenmeyen önyargıyı yönetilebilir bir çerçeveye oturtarak önemli bir eşiği temsil etmektedir. Önyargı, her zaman tamamen yok edilemeyebilir; ancak istenmeyen önyargıyı tanımlamak, ölçmek ve azaltmak elimizdedir. Bu standart bize şunu hatırlatıyor: YZ sistemleri, insan yapısıdır ve insan değerlerini yansıtmalıdır. Adalet, eşitlik ve kapsayıcılık gibi değerlerin teknolojiye entegre edilmesi tesadüfe bırakılamayacak kadar mühimdir. Bunun için de planlı bir çaba, disiplinler arası işbirliği ve doğru araçların kullanımı gerekir. ISO/IEC TS 12791 tam da bu nedenle var. YZ'yi *daha iyi bir güç haline getirmek için*.

Kaynaklar:

AI Standards Hub. (2024, 18 Temmuz). *Information technology — Artificial intelligence — Treatment of unwanted bias in classification and regression machine learning tasks (ISO/IEC TS 12791:2024)*. Erişildi: AI Standards Hub web sitesi: https://aistandardshub.org/ai-standards/information-technology-artificial-intelligence-treatment-of-unwanted-bias-in-classification-and-regression-machine-learning-tasks

American National Standards Institute. (2024). *ISO/IEC TS 12791:2024 – Information technology – Artificial intelligence – Treatment of unwanted bias in classification and regression machine learning tasks* [Standart ön izlemesi]. ANSI Webstore. Erişildi: https://webstore.ansi.org/standards/iso/ISOIECTS127912024

Price, A. (2021, 8 Kasım). Standards help address bias in artificial intelligence technologies. *IEC e-tech*. Erişildi: IEC e-tech dergisi (Çevrimiçi), Sayı 06/2021.

Michalsons. (2021, 8 Kasım). *ISO/IEC TR 24027:2021 – Bias in AI & AI decision making*. Erişildi: Michalsons (AI Law Blog) web sitesi: https://www.michalsons.com/blog/iso-iec-tr-24027-2021-bias-in-ai-ai-decision-making/53507

Hasanzadeh, F., Josephson, C. B., Waters, G., Adedinsewo, D., Azizi, Z., & White, J. A. (2025). Bias recognition and mitigation strategies in artificial intelligence healthcare applications. *NPJ Digital Medicine, 8*, 154. https://doi.org/10.1038/s41746-025-01503-7

Kumar, A. (2018, 20 Kasım). Machine learning models: Bias mitigation strategies. *DZone*. Erişildi: https://dzone.com/articles/machine-learning-models-bias-mitigation-strategies

Hasanzadeh, F., Josephson, C. B., Waters, G., Adedinsewo, D., Azizi, Z., & White, J. A. (2025). *Bias mitigation strategies across AI model life cycle* [Şekil/Figure]. “Bias recognition and mitigation strategies in artificial intelligence healthcare applications” başlıklı makalede, *NPJ Digital Medicine, 8*, 154 içinde yer almaktadır. (Erişim: ResearchGate Şekil bağlantısı)

International Organization for Standardization & International Electrotechnical Commission. (2021). *Information technology — Artificial intelligence (AI) — Bias in AI systems and AI aided decision making* (ISO/IEC TR 24027:2021, Teknik Rapor). Cenevre, İsviçre: ISO.

Yapay Zekâ Sistemleri ve Kalite Uzmanlığı

6. YAPAY ZEKÂ SİSTEMLERİ YAŞAM DÖNGÜSÜ SÜRECİNDE KALİTE UZMANLIK ALANININ DEĞERLENDİRİLMESİ

Dr. Hatice Merve Karataş[a], Soner Karataş[b]

[a] *TBD Ankara Şubesi Yönetim Kurulu Üyesi, TSE, merveolmez@gmail.com*
[b] *TBD YZ Standartları ÇG Başkanı, sonerkaratas@gmail.com*

Bu bölümün amacı YZ sistemleri yaşam döngüsü sürecinin kalite uzmanlık alanı açısından değerlendirilmesidir. Bu bağlamda YZ sistemleri yaşam döngüsü süreçlerini tarifleyen ISO/IEC 5338 (Bilgi teknolojisi — YZ Sistem Yaşam Döngüsü Süreçleri) Standardının kalite uzmanlık alanıyla ilgili olan maddeleri incelenmiştir. Öncelikle ilgili maddelerin detaylı olarak açıklaması ISO/IEC/IEEE 15288 (Sistem ve yazılım mühendisliği- Sistem yaşam döngüsü süreçleri) ve ISO/IEC/IEEE 12207 (Bilgi teknolojisi- Yazılım yaşam döngüsü süreçleri) standartları da dikkate alınarak verilmiştir. YZ sistemlerine yönelik olan ilave hususların açıklaması bu bilgilerle birlikte sağlanmıştır. Her bir madde içerisinde konuyla ilgili örnek vakalar verilmiştir. İlgili maddenin işletilmesi esnasında kullanılan standartlar ve hangi amaçlarla kullanıldığının bilgisi paylaşılmıştır. Son bölümde ise YZ sistemleri yaşam döngüsü sürecinde kalite uzmanlık alanının mevcut durumuna ilişkin genel değerlendirme ve gelecekteki değişimi konusunda öngörüler verilmiştir.

6.1. Kalite Yönetim Süreci (ISO/IEC 5338 Madde 6.2.5)

ISO/IEC 5338:2024-03 standardının 6.2.5 “Kalite yönetim süreci” maddesi, YZ sistem yaşam döngüsü boyunca kalite yönetimini kurumsal düzeyde güvence altına almak için çerçeve sunmaktadır. Bu süreç organizasyonel düzeyde politika geliştirme, kalite hedeflerinin belirlenmesi, kaynak tahsisi, sorumlulukların atanması ve sürekli iyileştirme mekanizmalarının işletilmesini içermektedir. YZ sistemlerinde veri bağımlılığı, model şeffaflığı, etik riskler ve

dinamik çevresel değişkenlik kaliteyi doğrudan etkiler. Kalite yönetim süreçlerinin bir bileşeni olan kalite güvencesi, YZ sistemlerini barındıran süreçlerde daha kritik bir rol üstlenebilmektedir.

Bu süreçte gerçekleştirilen faaliyetler yalnızca kalite planlaması ile sınırlı değildir. Aynı zamanda kalite göstergelerinin (Key Quality Indicators – KQIs) belirlenmesi, bu göstergelerin periyodik ölçülmesi, uygunsuzlukların yönetilmesi ve düzeltici-önleyici faaliyetlerin uygulanması da kalite yönetiminin önemli unsurlarıdır. Bu faaliyetler ISO/IEC 42001 ve ISO 9001'den alınan kalite yönetim sistemi ilkeleriyle güçlendirilir. ISO/IEC 42001 ve ISO 9001, süreç yaklaşımı, sürekli iyileştirme ve müşteri gereksinimlerine odaklanma ilkeleri ile kalite yönetiminin temelini sağlar. Risk temelli bakış açısının kaliteye entegrasyonu için ISO 31000'den faydalanılır; bu sayede kalite hedeflerini tehdit eden riskler tanımlanır, analiz edilir ve kontrol altına alınır (ISO 31000, 2018). Veri kalitesinin güvence altına alınmasında ISO/IEC 25012 veri kalite modeli, sistem niteliklerinin tanımlanmasında ise ISO/IEC 25010 kalite özellikleri modeli kullanılır (ISO/IEC 25012, 2008; ISO/IEC 25010, 2011). Ayrıca, bilgi güvenliği ve gizlilik gereksinimlerinin kalite hedefleriyle uyumlu şekilde yönetilebilmesi için ISO/IEC 27001 de referans alınır (ISO/IEC 27001, 2022).

Bu madde kapsamında öncelikle veri bağımlılığına karşı kalite yönetimi sürecinde veri kaynaklarının güvenilirliği, bütünlüğü ve güncelliği kalite planına dahil edilir; ISO/IEC 25012 ve 25024 veri kalitesi standartları bu noktada referans olarak kullanılır. Model şeffaflığı sorunu için kalite hedeflerine açıklanabilirlik kriterleri eklenir; modellerin nasıl karar verdiğini açıklayan dokumanlar, görselleştirmeler kalite güvence (QA) raporlarının bir parçası olur. Etik riskler bakımından kalite yönetim süreci, ISO/IEC 23894'e dayalı olarak adalet, tarafsızlık ve insan hakları ile uyumu kalite hedeflerinin içine alır; bu riskler için kontrol adımları ve ölçümler kalite planında tanımlanır.

Bu maddeye ilişkin örnek şu şekilde verilebilir; Bir sağlık kurumunda YZ tanı sistemi geliştirildiğinde kalite yönetimi sürecinde, eğitim verilerinin güncel klinik rehberlere göre seçildiği, modelin kararlarının doktorlara anlaşılır raporlarla sunulduğu, hasta grupları arasında adaletin gözetildiği ve yeni hastalık varyantları ortaya çıktığında sistemin performansının düzenli olarak yeniden değerlendirildiği kayıt altına alınır. Böylece kalite yönetimi, ISO/IEC 5338'in 6.2.5 maddesi uyarınca yalnızca teknik uygunluğu değil, aynı zamanda veri kalitesi, şeffaflık, etik uygunluk ve değişken çevresel faktörleri kapsayan çok boyutlu bir güvence mekanizması haline gelir.

6.2. Risk Yönetim Süreci (ISO/IEC 5338 Madde 6.3.4)

ISO/IEC 5338:2024-03 standardının 6.3.4 “Risk Yönetim Süreci” maddesi, YZ sistem yaşam döngüsünde risklerin tanımlanması, değerlendirilmesi, kontrol edilmesi ve izlenmesine yönelik süreçleri kapsar. Risk yönetimi, güvenlik veya güvenilirlik riskleriyle birlikte kalite hedeflerini tehdit eden unsurların sistematik biçimde ele alınmasını gerektirir. Kalite güvencesi süreçlerinin etkinliği, risklerin erken dönemde tanımlanmasına ve yönetilmesine bağlıdır.

Risk yönetimi süreci genellikle dört aşamadan oluşur. İlk aşamada risklerin tanımlanması yapılır; bu tanımlama teknik, organizasyonel ve çevresel boyutları kapsar. İkinci aşamada risklerin olasılık ve etki düzeyleri değerlendirilir, risk matrisi oluşturulur. Üçüncü aşamada risk kontrol önlemleri geliştirilir; riskin kabul edilebilir seviyelere çekilmesi için iyileştirme faaliyetleri planlanır. Son aşamada risklerin sürekli izlenmesi ve değerlendirilmesi yapılır. Bu döngü, kalite güvence süreciyle paralel olarak işletilir.

Riskler yalnızca teknik tehditlerle sınırlı değildir; veri önyargısı, etik uyumsuzluklar, modelin öngörülemez davranışları ve sürekli değişen çevresel koşullar da risk yönetiminin parçasıdır. Standard, risk yönetiminde ISO/IEC 23894 gibi YZ’ye özel rehberleri kullanarak etik ve sosyal risklerin kalite hedefleriyle birlikte değerlendirilmesini öngörür. Ayrıca risklerin yaşam döngüsü boyunca sürekli izlenmesi, “data drift” veya “concept drift” nedeniyle kalite kayıplarının önlenmesi için kritik görülür.

Bu madde ile ilgili uygulama örneği olarak banka kredi onay sistem verilebilir. Bir bankada kredi onay sistemi geliştirilirken veri setinde kadın ve erkek müşteriler arasında önyargılı sonuçlar ortaya çıkabilir. Risk uzmanı bu önyargıyı tanımlar, olasılık ve etkiyi değerlendirir; veri kalite uzmanı bu riskin azaltılması için veri dengelenmesi yöntemini önerir. Burada ISO/IEC 23894 standardı etik ve adalet risklerini ele almak için kullanılır.

Bu süreçte kullanılan uluslararası standartlar önemli destek sağlarlar. ISO 31000 risk yönetimi için genel çerçeveyi verir; risklerin tanımlanması, değerlendirilmesi ve kontrol edilmesi için metodolojik bir yaklaşım sunar (ISO 31000, 2018). ISO/IEC 27005 bilgi güvenliği risk yönetimini detaylandırır ve özellikle YZ sistemlerinin siber güvenlik risklerinin ele alınmasında kullanılır (ISO/IEC 27005, 2018). ISO/IEC 23894 YZ risk yönetimi için özel olarak geliştirilmiş bir rehberdir; etik, sosyal ve teknik risklerin yönetimini YZ bağlamında tanımlar (ISO/IEC 23894, 2023). Ayrıca ISO/IEC 42001 ve ISO 9001’in risk temelli düşünme ilkesi, kalite yönetim sistemiyle entegrasyon sağlar (ISO/IEC 42001,2023), (ISO 9001, 2015). Bu standartlar, ISO/IEC 5338’de tanımlanan risk yönetimi sürecinin daha güçlü ve bağlama özgü şekilde uygulanmasını mümkün kılar.

6.3. Ölçüm Süreci (ISO/IEC 5338 Madde 6.3.7)

ISO/IEC 5338:2024-03 standardının 6.3.7 “Ölçüm süreci” maddesi, YZ sistem yaşam döngüsü boyunca performans, kalite ve uygunluk göstergelerinin, veri ve bilgilerin sistematik biçimde tanımlanması, toplanması, analiz edilmesi ve raporlanmasını amaçlar. Ölçüm süreci, yalnızca teknik doğruluk ve verimlilik göstergelerinin değil, aynı zamanda kalite yönetim hedeflerinin, risklerin ve paydaş gereksinimlerinin de nicel verilerle takip edilmesini sağlar.

Ölçüm sürecinin işleyişi, ölçüm hedeflerinin belirlenmesiyle başlar. Bu hedefler, kalite politikaları, proje hedefleri ve regülasyon gereksinimleri ile uyumlu şekilde tanımlanır. Sonrasında ölçülecek metrikler seçilir; bunlar teknik performans göstergeleri (ör. doğruluk, gecikme süresi), kalite göstergeleri (ör. güvenilirlik, kullanılabilirlik) veya süreç göstergeleri (ör. hata düzeltme süresi, uygunsuzluk kapanış oranı) olabilir. Metriklerin tanımlanmasının ardından veri toplama ve doğrulama yöntemleri belirlenir. Toplanan veriler analiz edilerek raporlanır; sonuçlar kalite güvence, risk yönetimi ve sürekli iyileştirme süreçlerine girdi sağlar.

YZ sistemlerinde performans ve kalite göstergeleri yalnızca doğruluk oranlarıyla sınırlı değildir; adalet, açıklanabilirlik, güvenilirlik ve sürekli iyileştirme gibi boyutlar da ölçüm kapsamına girer. Ayrıca ölçüm süreci, değişen veri dağılımlarına karşı sistemin güncel performansını izlemek için sürekli olarak uygulanmalıdır. Bu bağlamda ISO/IEC 25010 kalite nitelikleri ve ISO/IEC 23894 YZ risk yönetimi rehberi ölçüm hedeflerine entegre edilir. Örneğin bir e-ticaret platformunda ürün öneri algoritmasının performansı ölçülürken yalnızca önerilerin doğruluk oranı değil, aynı zamanda önerilerin çeşitliliği ve müşteri grupları arasında adaletli dağılımı da kalite göstergeleri olarak takip edilir.

Bu süreçte ISO/IEC 5338 çerçeveyi belirlerken, başka standartlar da destekleyici rol oynar. ISO/IEC 42001, ISO 9001, ölçüm ve analiz faaliyetlerinin kalite yönetim sistemi içinde sürekli iyileştirme aracı olarak kullanılmasını sağlar (ISO/IEC 42001,2023), (ISO 9001, 2015). ISO/IEC 15939 (Yazılım Ölçüm Süreci)) ölçüm sürecinin metodolojik detaylarını (ölçüm planı, metrik seçimi, veri toplama ve analiz) tanımlar ve ISO/IEC 5338’deki ölçüm sürecine dayanak oluşturur. ISO/IEC 25010 kalite özellikleri modeli, hangi niteliklerin ölçüleceğini tanımlamada kullanılır (ISO/IEC 25010, 2011). Ayrıca ISO/IEC 27004 bilgi güvenliği ölçümleri standardı, güvenlik performansının ölçülmesinde QA kapsamına entegre edilir (ISO/IEC 27004, 2016).

6.4. Kalite Güvence Süreci (ISO/IEC 5338 Madde 6.3.8)

ISO/IEC 5338:2024-03 standardının 6.3.8 maddesi “Kalite Güvence Süreci”, YZ sistem yaşam döngüsü boyunca yürütülen süreçlerin ve elde edilen çıktıların tanımlı kalite standartlarına uygunluğunu güvence altına almayı amaçlar.

Kalite Güvencesi (QA), kalite şartlarının yerine getirildiğine dair güven oluşturmayı hedefler. Proje yaşam döngüsü süreçleri ve bunların çıktıları, ürün yada hizmetlerin hedeflenen kalite düzeyine ulaşması amacıyla proaktif biçimde analiz edilir. Kalite Güvence Süreci’nin işleyişinde birkaç faaliyet öne çıkar. Öncelikle, kalite güvence planı hazırlanır; burada hangi süreçlerin hangi ölçütlerle izleneceği, hangi kanıtların toplanacağı ve kabul kriterlerinin nasıl uygulanacağı tanımlanır. Daha sonra, süreç gözetimleri ve iç denetimler yürütülür; süreçlerin ve ürünlerin kalite kriterlerine uyumlu olup olmadığı kanıta dayalı şekilde belgelenir. Uygunsuzluklar tespit edildiğinde bunlar kayıt altına alınır, kök neden analizi yapılır ve düzeltici-önleyici faaliyetler uygulanır. Son aşamada, kalite güvence sonuçları yönetim gözden geçirmesine sunularak sürekli iyileştirme için bilgi sağlanır.

YZ sistemlerinde verinin güvenilirliği, modelin şeffaflığı, etik uygunluğu ve sürekli performans istikrarı da kalite güvencesinin bir parçasıdır. Bu nedenle kalite güvence (QA) süreci, düzenli testlerle birlikte veri önyargısı kontrolleri, açıklanabilirlik değerlendirmeleri ve sürekli validasyon mekanizmalarıyla desteklenir. Ayrıca QA faaliyetleri, ISO/IEC 23894 gibi YZ risk yönetimi standartlarıyla entegre edilerek etik ve sosyal boyutların da güvence altına alınmasını sağlar. Örneğin bir hastanede YZ destekli tanı sisteminin kalite güvencesinde sistemin doğru teşhis yapıp yapmadığıyla birlikte, farklı hasta grupları arasında tutarlı ve adil sonuçlar üretip üretmediği de düzenli olarak kontrol edilir.

Bu sürece yönelik ISO/IEC 42001 ve ISO 9001 standartları kalite yönetiminin genel ilkelerini ve uygunsuzluk-düzeltici faaliyet mekanizmasını sağlar (ISO/IEC 42001,2023), (ISO 9001, 2015). ISO 19011 kalite denetimlerinde kullanılan yöntemler, denetçi tarafsızlığı ve kanıt toplama teknikleri için referansı sağlar (ISO 19011, 2018). ISO/IEC 12207 yazılım yaşam döngüsündeki doğrulama, geçerleme ve konfigürasyon yönetimi gibi süreçlerin YZ QA bağlamına uyarlanmasında yol gösterici görevi görür (ISO/IEC/IEEE 12207, 2017). ISO/IEC 25010 kalite özellikleri modeli, özellikle güvenilirlik, kullanılabilirlik ve güvenlik gibi niteliklerin QA değerlendirmelerinde kullanılmasını sağlar (ISO/IEC 25010, 2011). ISO/IEC 27001 ise bilgi güvenliği ve gizlilik kısıtlarının QA değerlendirmelerine dahil edilmesini sağlar (ISO/IEC 27001, 2022).

6.5. Bilgi Edinme Süreci (ISO/IEC 5338 Madde 6.4.7)

ISO/IEC 5338:2024-03 standardının 6.4.7 “Bilgi Edinme Süreci” maddesi, YZ sistem yaşam döngüsünde kullanılacak bilginin elde edilmesi, belgelenmesi, doğrulanması ve kalite kriterlerine göre güvence altına alınmasını kapsamaktadır. Bu süreç bilgi kaynaklarının güvenilirliği, doğruluğu, güncelliği ve sistem gereksinimleriyle uyumunu hedeflemektedir. Kalite perspektifinden bakıldığında, bilgi edinme süreci, sonraki veri hazırlama ve model geliştirme süreçlerinin temeli olduğu için kritik önem taşır. Yanlış veya eksik bilgi kaynakları, sistemin çıktılarında ciddi kalite sorunlarına ve risklere yol açabilir.

Sürecin icrasında belirli faaliyetler öne çıkar. İlk aşamada bilgi ihtiyacının tanımlanması yapılır; bu ihtiyaçlar paydaş gereksinimlerine ve kalite hedeflerine göre şekillenir. İkinci aşamada bilgi kaynakları belirlenir; bu kaynakların güvenilirliği ve doğruluğu kalite uzmanları tarafından test edilir. Üçüncü aşamada bilgi belgelenir ve sistem yaşam döngüsü boyunca izlenebilir olacak şekilde kayıt altına alınır. Dördüncü aşamada kalite güvence ekibi, bilginin güncelliğini ve tutarlılığını periyodik olarak değerlendirir.

YZ sistemlerinde bilgi edinme sürecinin klasik yazılım projelerine göre daha karmaşık olduğu değerlendirilir. YZ sistemleri için edinilen bilginin kaynağı, güvenilirliği, güncelliği ve önyargılardan arındırılmış olması kritik kalite faktörleridir. Ayrıca edinilen bilginin etik, hukuki ve bağlamsal uygunluk açısından da değerlendirilmesi gerekir. Bu nedenle kalite güvencesi sürecinde, bilgi kaynaklarının izlenebilirliği ve periyodik doğrulaması kalite yönetim planının ayrılmaz bir parçası olarak ele alınır. Örneğin bir tıp fakültesinde geliştirilen YZ tabanlı tanı sistemi için edinilen klinik rehberlerin yalnızca güncel bilimsel otoritelerden alınması değil, aynı zamanda farklı hasta gruplarını kapsayıcı ve tarafsız içerikler sunması da kalite kontrolünün bir parçası olur.

Bu sürecin uygulanmasında farklı standartlar tamamlayıcı rol oynar. ISO/IEC 25012 veri kalite modeli, bilgi kalitesini ölçmede kullanılır; doğruluk, eksiksizlik ve güncellik kriterleri bu modele göre değerlendirilir. ISO/IEC 42001, YZ yönetim sistemi ve ISO 9001 kalite yönetim sistemi standartları, bilgi edinme sürecinin planlı, dokümante edilmiş ve sürekli iyileştirmeye açık olmasını sağlar (ISO/IEC 42001,2023), (ISO 9001, 2015). ISO/IEC 27001 bilgi güvenliği yönetim sistemi standardı, elde edilen bilginin güvenliğini, gizliliğini ve bütünlüğünü garanti altına almak için uygulanır (ISO/IEC 27001, 2022). Ayrıca ISO/IEC 23894 YZ risk yönetimi rehberi, bilgi edinme sürecinde etik veya sosyal risklerin (örneğin önyargılı bilgi kaynakları) ele alınmasına yardımcı olur (ISO/IEC 23894, 2023).

6.6. YZ Veri Mühendisliği Süreci (ISO/IEC 5338 Madde 6.4.8)

ISO/IEC 5338:2024-03 standardının 6.4.8 "Yapay Zekâ Veri Mühendisliği Süreci" maddesi, YZ sistemlerinin yaşam döngüsünde kullanılan verilerin işlenmesi, dönüştürülmesi ve kalite standartlarına uygun hale getirilmesini kapsamaktadır. Bu süreç verinin doğruluk, bütünlük, güvenilirlik ve önyargılardan arındırılmış olması gibi kalite kriterlerinin karşılanmasını hedefler. Kalite boyutuyla ele alındığında, veri mühendisliği süreci, YZ modelinin güvenilirlik, adalet ve performans açısından sürdürülebilir olmasını sağlayan temel aşamalardan biridir.

Sürecin faaliyetleri arasında veri temizleme (hatalı, eksik veya yinelenen kayıtların düzeltilmesi), veri dönüştürme (format, birim ve yapı uyumlaştırmaları), veri entegrasyonu (farklı kaynaklardan elde edilen verilerin birleştirilmesi) ve veri önyargısının analizi bulunmaktadır. Kalite açısından bu faaliyetlerin amacı, model eğitimine giren verinin mümkün olduğunca temsili, hatasız ve dengeli olmasını sağlamaktır.

YZ modellerinin başarısı, kullanılan verinin kalitesi, temsiliyet düzeyi ve önyargılardan arındırılmış olmasına doğrudan bağlıdır. Bu nedenle veri mühendisliği sürecinde verilerin teknik olarak temizlenmesiyle birlikte etik uygunluk, çeşitlilik ve yasal düzenlemelere uyumu da gözetilmelidir. Örneğin bir bankada kredi başvurularını değerlendiren YZ modeli için müşteri verileri işlenirken yalnızca eksik ya da hatalı veriler düzeltilmez, aynı zamanda farklı sosyo-ekonomik grupların dengeli biçimde temsil edilmesi sağlanarak modelin adil ve güvenilir sonuçlar üretmesi güvence altına alınır.

Bu süreçte çeşitli uluslararası standartlardan yararlanılır. ISO/IEC 25012 veri kalite modeli, verilerin doğruluk, eksiksizlik, tutarlılık ve güvenilirlik açısından değerlendirilmesini sağlar (ISO/IEC 25012, 2008). ISO/IEC 25024 veri kalite ölçüm standardı, bu boyutların nicel olarak ölçülmesine imkan verir (ISO/IEC 25024, 2015). ISO/IEC 23894 YZ risk yönetimi standardı, veri mühendisliği sürecinde önyargı, adalet ve etik risklerin ele alınmasını destekler (ISO/IEC 23894, 2023). ISO/IEC 27001 bilgi güvenliği yönetimi, özellikle kişisel verilerin gizliliği ve bütünlüğünün korunması için sürece entegre edilir (ISO/IEC 27001, 2022). Ayrıca ISO/IEC 42001 YZ yönetim sistemi ve ISO 9001 kalite yönetim sistemi, veri mühendisliği sürecinde planlama, izleme ve sürekli iyileştirme mekanizmasının işletilmesine katkı sağlar (ISO/IEC 42001,2023), (ISO 9001, 2015).

6.7. Doğrulama Süreci (ISO/IEC 5338 Madde 6.4.11)

ISO/IEC 5338:2024-03 standardının 6.4.11 "Verifikasyon Süreci" maddesi, YZ sistemlerinde geliştirilen bileşenlerin tanımlanmış gereksinimlere, tasarım belgelerine ve teknik spesifikasyonlara uygun olarak inşa edilip edilmediğinin sistematik biçimde test edilmesini amaçlar. Verifikasyon süreci kalite açısından kritik bir noktada yer alır; çünkü burada yapılan kontroller, sistemin "doğru şekilde yapılıp yapılmadığını" ortaya koyar. Böylelikle validasyon sürecine geçmeden önce tasarım uyumsuzluklarının erken aşamada tespit edilmesi ve düzeltilmesi sağlanır.

Verifikasyon sürecinin uygulanışında üç temel faaliyet öne çıkar. İlk olarak, doğrulama kriterleri ve kabul koşulları tanımlanır; bu koşullar genellikle tasarım belgelerinden ve gereksinim analizlerinden türetilir. İkinci olarak, test senaryoları uygulanır; sistem bileşenleri üzerinde fonksiyonel, performans ve entegrasyon seviyesinde doğrulama testleri yürütülür. Üçüncü olarak, test sonuçları dokümante edilir ve uygunsuzluk raporları hazırlanır.

YZ sistemlerinde verifikasyon yalnızca kodun tasarım gereksinimlerine uygun olup olmadığını test etmekle sınırlı değildir. Aynı zamanda modelin eğitildiği veriye uygun davranıp davranmadığı, performansının farklı senaryolarda istikrarlı olup olmadığı ve önyargılı sonuçlar üretip üretmediği de değerlendirilmelidir. Verifikasyon süreci, veri çeşitliliği, belirsizlik ve olasılıklı çıktılar nedeniyle deterministik sistemlere kıyasla daha fazla belirsizlik içerir. Bu nedenle kalite güvencesi kapsamında, test senaryoları geniş veri kümeleri ve farklı kullanım koşullarıyla çeşitlendirilerek yürütülür.

Verifikasyon sürecine örnek olarak sesli asistan komutu verilebilir. Bir sesli asistanın **"Işığı kapat"** komutunu doğru algılayıp uygulayıp uygulamadığını test etmek verifikasyon sürecine girer. Eğitim setinde olmayan farklı ses tonları ve aksanlarla deneme yapılır. %95 doğruluk oranı yakalanırsa sistem onaylanır ve modelin farklı koşullarda doğru tepki verip vermediği doğrulanmış olur.

Bu süreçte destekleyici rol oynayan uluslararası standart bulunmaktadır. ISO/IEC/IEEE 12207 yazılım yaşam döngüsünde doğrulama sürecini detaylandırır; doğrulama planlarının hazırlanması, test faaliyetlerinin yürütülmesi ve sonuçların belgelenmesi için metodolojik bir temel sunar (ISO/IEC/IEEE 12207, 2017). ISO/IEC 25010 kalite özellikleri modeli, doğrulama sırasında hangi kalite boyutlarının kontrol edileceğini tanımlar; örneğin güvenilirlik, kullanılabilirlik veya güvenlik gibi (ISO/IEC 25010, 2011). ISO 9001 kalite yönetimi standardı, doğrulama faaliyetlerinin belgelenmesi, izlenebilirlik ve sürekli iyileştirme çerçevesinde

yürütülmesini sağlar (ISO 9001, 2015). ISO/IEC 27001 ise özellikle bilgi güvenliği gereksinimlerinin doğrulama kriterlerine entegre edilmesi için kullanılır (ISO/IEC 27001, 2022).

6.8. Geçerleme Süreci (ISO/IEC 5338 Madde 6.4.13)

ISO/IEC 5338:2024-03 standardının 6.4.13 "Geçerleme (Validasyon) Süreci" maddesi, YZ sisteminin geliştirilme amacıyla uyumlu olup olmadığını ve paydaşların gereksinimlerini karşılayıp karşılamadığını değerlendirmeyi hedefler. Bu süreç, verifikasyon sürecinden farklı olarak, "doğru sistemin geliştirilip geliştirilmediğini" test eder. Kalite açısından validasyon, sistemin kullanıcı ihtiyaçlarını, regülasyonları, etik ilkeleri ve organizasyonel hedefleri karşılayıp karşılamadığını güvence altına alır.

Validasyon sürecinde üç ana aşama öne çıkar. İlk aşamada, paydaş gereksinimlerine ve kalite hedeflerine dayalı validasyon kriterleri belirlenir. İkinci aşamada, bu kriterler üzerinde sistem test edilir; bu testler genellikle kullanım senaryolarına dayalıdır ve gerçekçi ortam koşullarını simüle eder. Üçüncü aşamada, test sonuçları raporlanır ve kalite güvence ekibi tarafından doğrulanır. Eğer sistem paydaş ihtiyaçlarını karşılamıyorsa, düzeltici faaliyetler başlatılır.

YZ sistemlerinde validasyon teknik gereksinimlerle birlikte paydaş beklentileri, etik uygunluk ve düzenleyici gereklilikleri de kapsamalıdır. YZ modelleri deterministik olmadığı için, geçerleme sürecinde performansın farklı kullanıcı gruplarında ve değişen koşullar altında tutarlı olup olmadığı da incelenir. Ayrıca sistemin amaçlanan işlevi gerçekten yerine getirip getirmediği, kullanıcıya fayda sağlayıp sağlamadığı ve adalet ilkeleriyle uyumlu çalışıp çalışmadığı validasyon kapsamında değerlendirilir. Bu nedenle klasik yazılım projelerinde olmayan sosyal etki ve önyargı kontrolleri burada kritik hale gelir. Örneğin bir sağlık kurumunda geliştirilen YZ tabanlı tanı sistemi, doğru tanılar üretip üretmediğiyle birlikte farklı yaş ve cinsiyet gruplarındaki hastalara eşit doğrulukla hizmet verip vermediği açısından da validasyon testlerine tabi tutulur.

Validasyon sürecinin uygulanmasında çeşitli standartlar destekleyici rol oynar. ISO/IEC 25010 kalite özellikleri modeli, validasyon sırasında kontrol edilecek kalite boyutlarını tanımlar (örneğin güvenilirlik, kullanılabilirlik, doğruluk) (ISO/IEC 25010, 2011). ISO/IEC 23894 YZ risk yönetimi standardı, validasyon sürecine etik, sosyal ve güvenlik risklerinin entegrasyonunu sağlar (ISO/IEC 23894, 2023). ISO/IEC/IEEE 12207 validasyon adımlarını yazılım yaşam

döngüsüne göre yapılandırır (ISO/IEC/IEEE 12207, 2017). ISO/IEC 27001, bilgi güvenliği kontrollerinin validasyon kapsamında değerlendirilmesini mümkün kılar (ISO/IEC 27001, 2022).

6.9. Sürekli Geçerleme Süreci (ISO/IEC 5338 Madde 6.4.14)

ISO/IEC 5338:2024-03 standardının 6.4.14 "Sürekli Validasyon Süreci" maddesi, YZ sistemlerinin devreye alındıktan sonra da sürekli olarak tatmin edici performans göstermeye ve paydaş ihtiyaçlarını karşılamaya devam edip etmediğini değerlendirir. YZ sistemlerinde çevresel koşulların, veri dağılımlarının ve kullanıcı davranışlarının zamanla değişmesi nedeniyle kalite yalnızca geliştirme aşamasında sağlanamaz; kullanım süresince de doğrulanmalıdır. Girdi verilerindeki sapmaların (data drift) ya da hedeflenen çıktılardan sapmaların (concept drift) test verileriyle ölçülmesi ve izlenmesi bu nedenle önem kazanmaktadır. Bu madde, özellikle concept drift ve data drift gibi problemlerin kaliteyi olumsuz etkilemesini önlemeyi hedefler. Validation sürecinin bir uzantısı olarak düşünülebilecek bu süreç, "sistem zaman içinde doğru çalışmaya devam ediyor mu?" sorusunu yanıtlar. Kalite açısından sürekli validasyon, YZ sistemlerinin güvenilirliğini, etik uygunluğunu ve kullanıcı memnuniyetini korumak için kritik bir mekanizmadır.

Sürekli validasyon süreci, YZ sistemlerinde performansın zamanla sabit kalmadığını ve sürekli olarak gözden geçirilmesi gerektiğini vurgular. Veri dağılımlarındaki değişim (data drift) veya kavramsal kaymalar (concept drift) nedeniyle sistemin ilk validasyon sonuçları geçerliliğini yitirebilir. Bu nedenle sürekli validasyon, sistemin güncel verilerle yeniden test edilmesini, kalite hedefleriyle uyumunun korunmasını ve risklerin erken aşamada tespit edilmesini sağlar. Ayrıca etik ve regülasyon boyutlarının da düzenli aralıklarla tekrar değerlendirilmesi sürece dahildir. Örneğin bir bankada kredi risk modeli, ekonomik koşulların değişmesiyle birlikte müşterileri olduğundan daha riskli veya daha güvenli sınıflandırmaya başlayabilir; bu durumda sürekli validasyon süreci, modelin adil ve doğru kararlar üretmeye devam edip etmediğini kontrol eder.

Bu süreçte birçok standart destekleyici rol oynar. ISO/IEC 25010 kalite özellikleri modeli, sürekli validasyonda hangi kalite boyutlarının izleneceğini tanımlar (ör. güvenilirlik, kullanılabilirlik, güvenlik) (ISO/IEC 25010, 2011). ISO/IEC 27004 bilgi güvenliği metrikleri standardı, güvenlik performansının sürekli ölçülmesini sağlar (ISO/IEC 27004, 2016). ISO/IEC 23894 YZ risk yönetimi rehberi, veri veya algoritma önyargılarının zamanla artıp artmadığını

izleme konusunda yol gösterir (ISO/IEC 23894, 2023). ISO/IEC 42001 YZ yönetim sistemi ve ISO 9001 kalite yönetimi standardları ise sürekli iyileştirme döngüsü sayesinde validasyon sürecini organizasyonel kalite hedefleriyle uyumlu hale getirir (ISO/IEC 42001,2023), (ISO 9001, 2015).

6.10. İşletme Süreci (ISO/IEC 5338 Madde 6.4.15)

ISO/IEC 5338:2024-03 standardının 6.4.15 "İşletme Süreci" maddesi, YZ sistemlerinin geliştirme aşamasından sonra işletilmesi sürecini düzenler. Sistemi işletmekle görevli personelin gereksinimleri belirlenir ve ilgili personel görevlendirilerek sunulan hizmetlerin performansı değerlendirilir. Operasyonel anormallikler tespit ve analiz edilir. Bu madde, YZ sisteminin beklenen işlevleri sürekli, güvenilir ve kullanıcı ihtiyaçlarına uygun bir biçimde yerine getirmesini amaçlar. Kalite açısından işletme süreci, hizmet sürekliliğini, kullanıcı deneyimini, regülasyon uyumunu ve operasyonel risklerin kontrolünü güvence altına alır.

İşletme sürecinde öne çıkan faaliyetler : sistemin canlı ortama alınması, operasyonel prosedürlerin uygulanması, kullanıcı desteği sağlanması, performans ve güvenlik metriklerinin sürekli izlenmesi, uygunsuzlukların raporlanması ve düzeltici faaliyetlerin başlatılmasıdır. Bu süreç, kalite yönetim sisteminin "müşteri odaklılık" ve "süreklilik" ilkeleriyle bağlantılıdır.

YZ sistemleri dinamik veriyle çalıştığı için, operasyon sürecinde performans kaybı, önyargı veya güvenlik zafiyetleri ortaya çıkabilir. Bu nedenle kalite yönetimi, sistemin çalışırlığıyla birlikte, güvenilirliğini, etik uygunluğunu ve mevzuata uyumunu da hedefler. Operasyon sırasında sürekli validasyon ve risk yönetimiyle entegrasyon, sistemin zamanla değişen koşullarda da kalite hedeflerini karşılamasını sağlar. Bu sürecin uygulanmasına örnek olarak öneri sistemleri verilebilir. Örneğin bir e-ticaret platformunda çalışan öneri sistemi, operasyon sürecinde müşterilere ürün önermeye devam ederken kalite uzmanları düzenli olarak algoritmanın adaletli, güvenli ve müşteri memnuniyetini artıracak şekilde çalışıp çalışmadığını izler.

Bu süreçte destekleyici rol oynayan standartlar şu şekilde verilebilir. ISO/IEC 20000 hizmet yönetimi standardı, operasyonel hizmet kalitesinin izlenmesini sağlar (ISO/IEC 20000, 2018). ISO/IEC 27001 bilgi güvenliği yönetimi standardı, operasyon sırasında gizlilik, bütünlük ve erişilebilirliğin korunmasını garanti eder (ISO/IEC 27001, 2022). ISO/IEC 25010 kalite modeli, operasyon sırasında kullanılabilirlik, performans verimliliği ve güvenilirlik gibi niteliklerin ölçülmesini sağlar (ISO/IEC 25010, 2011). ISO/IEC 42001, ISO 9001 kalite

yönetimi standardı, operasyonun planlı, dokümante edilmiş ve sürekli iyileştirmeye açık şekilde yürütülmesine katkı sağlar (ISO/IEC 42001,2023), (ISO 9001, 2015). Ayrıca ISO/IEC 23894 YZ risk yönetimi rehberi, operasyon sırasında etik ve sosyal risklerin değerlendirilmesinde kullanılır (ISO/IEC 23894, 2023).

6.11. Bakım Süreci (ISO/IEC 5338 Madde 6.4.16)

ISO/IEC 5338:2024-03 standardının 6.4.16 "Bakım Süreci" maddesi, YZ sistemlerinin devreye alınmasından sonra sürekliliğini, güvenilirliğini ve kalite hedefleriyle uyumunu korumak için yürütülen faaliyetleri kapsar. YZ sistemleri dinamik ortamlarda çalıştığından, veri dağılımları, kullanıcı beklentileri, yasal düzenlemeler ve teknolojik altyapı zamanla değişir. Bu nedenle bakım süreci, sistemin işlevselliğini ve kalitesini koruyabilmek için düzeltici, uyarlayıcı ve iyileştirici müdahaleleri içerir. Kalite açısından bakım süreci, teknik sürekliliklе birlikte etik uygunluk, yasal uyum ve müşteri memnuniyetini de garanti altına alır.

Bakım sürecinde üç ana faaliyet grubu öne çıkar. Düzeltici bakım, sistemdeki hataların ve uygunsuzlukların giderilmesine odaklanır. Uyarlayıcı bakım, değişen çevresel koşullara veya düzenlemelere uyum sağlamak için sistemin güncellenmesini kapsar. Önleyici/iyileştirici bakım ise performansı artırmak ve olası kalite sorunlarını önlemek amacıyla yapılan iyileştirme faaliyetlerini içerir. Bu üç faaliyet, sistemin uzun vadeli kalite güvenliğini sağlar.

YZ sistemlerinde bakım faaliyetleri, klasik yazılım bakımından farklı olarak dinamik veri, değişen regülasyonlar ve etik gereklilikler doğrultusunda sürekli güncellemeler gerektirmektedir. YZ sistemlerinde teknik hataların düzeltilmesi, veri setlerinin güncellenmesi, önyargıların giderilmesi ve model performansının yeniden dengelenmesi bakım süreçlerine dahil olmaktadır. Örneğin bir sağlık kurumunda kullanılan tanı destek sistemi için yeni tıbbi protokoller yayımlandığında, bakım süreci bu güncellemeleri sisteme entegre eder ve farklı hasta gruplarında doğruluğun korunmasını sağlar.

Bu süreçte bir dizi standart tamamlayıcı rol oynar. ISO/IEC/IEEE 14764 yazılım bakım standardı, bakım süreçlerinin sınıflandırılmasını ve uygulanmasını metodolojik olarak tanımlar (ISO/IEC/IEEE 14764, 2006). ISO/IEC/IEEE 12207 yazılım yaşam döngüsü standardı, bakımın sistem yaşam döngüsündeki yerini ve kalite entegrasyonunu ortaya koyar (ISO/IEC/IEEE 12207, 2017). ISO/IEC 42001 YZ yönetim sistemi ve ISO 9001 kalite yönetim sistemi standartları, bakım süreçlerinin dokümante edilmesini, izlenmesini ve sürekli iyileştirme döngüsüne dahil edilmesini sağlar (ISO/IEC 42001,2023), (ISO 9001, 2015).

ISO/IEC 27001 bilgi güvenliği standardı, bakım sırasında sistem güvenliği ve veri gizliliğinin korunmasını güvence altına alır (ISO/IEC 27001, 2022). ISO/IEC 23894 YZ risk yönetimi standardı, bakım faaliyetlerinde etik, sosyal ve güvenlik risklerinin dikkate alınmasına katkı sağlar (ISO/IEC 23894, 2023).

6.12. Sonuç ve Nihai Değerlendirme

Bu bölümde incelenen süreçler (6.2.5, 6.3.4, 6.3.7, 6.3.8, 6.4.7, 6.4.8, 6.4.11, 6.4.13, 6.4.14, 6.4.15, 6.4.16), YZ sistemleri yaşam döngüsünde kalitenin, yalnızca test ve ürün doğruluğu ile sınırlı kalmadığını; bilgi/veri kalitesi, etik ve adalet, açıklanabilirlik, güvenlik-gizlilik, risk yönetimi ve operasyonel süreklilik boyutlarını birlikte ele alan bir disipline dönüştüğünü göstermektedir. ISO/IEC 5338 bu disipline süreçsel omurga sağlarken; ISO/IEC/IEEE 15288–12207 (yaşam döngüsü), ISO/IEC 25010–25012–25024 (kalite nitelikleri ve veri kalitesi), ISO/IEC 15939 (ölçüm), ISO/IEC 27001–27004–27005 (BGYS ve ölçümler), ISO/IEC 42001 (YZ Yönetim sistemi), ISO 9001 (KYS), ISO 19011 (denetim) ve ISO/IEC/IEEE 14764 (bakım) gibi standartlar kalite uzmanının kanıta dayalı uygulamasını etkinleştirmektedir. Uygulamada, kalite yönetimi; KQI'ların tanımlanması/izlenmesi, veri ve bilgi tedarik zincirinin doğrulanması, doğrulama–geçerleme–sürekli geçerleme döngüsünün işletilmesi ve uygunsuzlukların kök-nedenleriyle kapatılması şeklinde bütüncül bir yapıya ulaşmaktadır.

Geleceğe yönelik olarak YZ sistemlerinin ISO/IEC 42001 tabanlı YZ Yönetim Sistemi entegrasyonunun belirleyici olacağı değerlendirilmektedir. Bu çerçevede kalite uzmanının rolü, proje bazlı kontrolden kurumsal yönetişime evrilerek risk, veri/bilgi kalitesi, güvenlik ve uyum fonksiyonlarıyla birlikte, sürdürülebilir değer üretimini güvence altına alan çok-disiplinli bir pozisyonuna dönüşeceği değerlendirilmektedir.

6.13. Kaynakça

ISO 9001:2015 Quality management systems—Requirements. Geneva, Switzerland: ISO.

ISO 31000:2018 Risk management—Guidelines. Geneva, Switzerland: ISO.

ISO 19011:2018 Guidelines for auditing management systems. Geneva, Switzerland: ISO.

ISO/IEC 15939:2007 Systems and software engineering—Measurement process. Geneva, Switzerland: ISO/IEC.

ISO/IEC 25012:2008 Software engineering—Software product Quality Requirements and Evaluation (SQuaRE)—Data quality model. Geneva, Switzerland: ISO/IEC.

ISO/IEC 25010:2011 Systems and software engineering—Systems and software Quality Requirements and Evaluation (SQuaRE)—System and software quality models. Geneva, Switzerland: ISO/IEC.

ISO/IEC 25024:2015 Systems and software engineering—Systems and software Quality Requirements and Evaluation (SQuaRE)—Measurement of data quality. Geneva, Switzerland: ISO/IEC.

ISO/IEC 27004:2016 Information technology—Security techniques—Information security management—Monitoring, measurement, analysis and evaluation. Geneva, Switzerland: ISO/IEC.

ISO/IEC 20000-1:2018 Information technology—Service management—Part 1: Service management system requirements. Geneva, Switzerland: ISO/IEC.

ISO/IEC 27005:2018 Information technology—Security techniques—Information security risk manage ment. Geneva, Switzerland: ISO/IEC.

ISO/IEC 27001:2022 Information security, cybersecurity and privacy protection—Information security management systems—Requirements. Geneva, Switzerland: ISO/IEC.

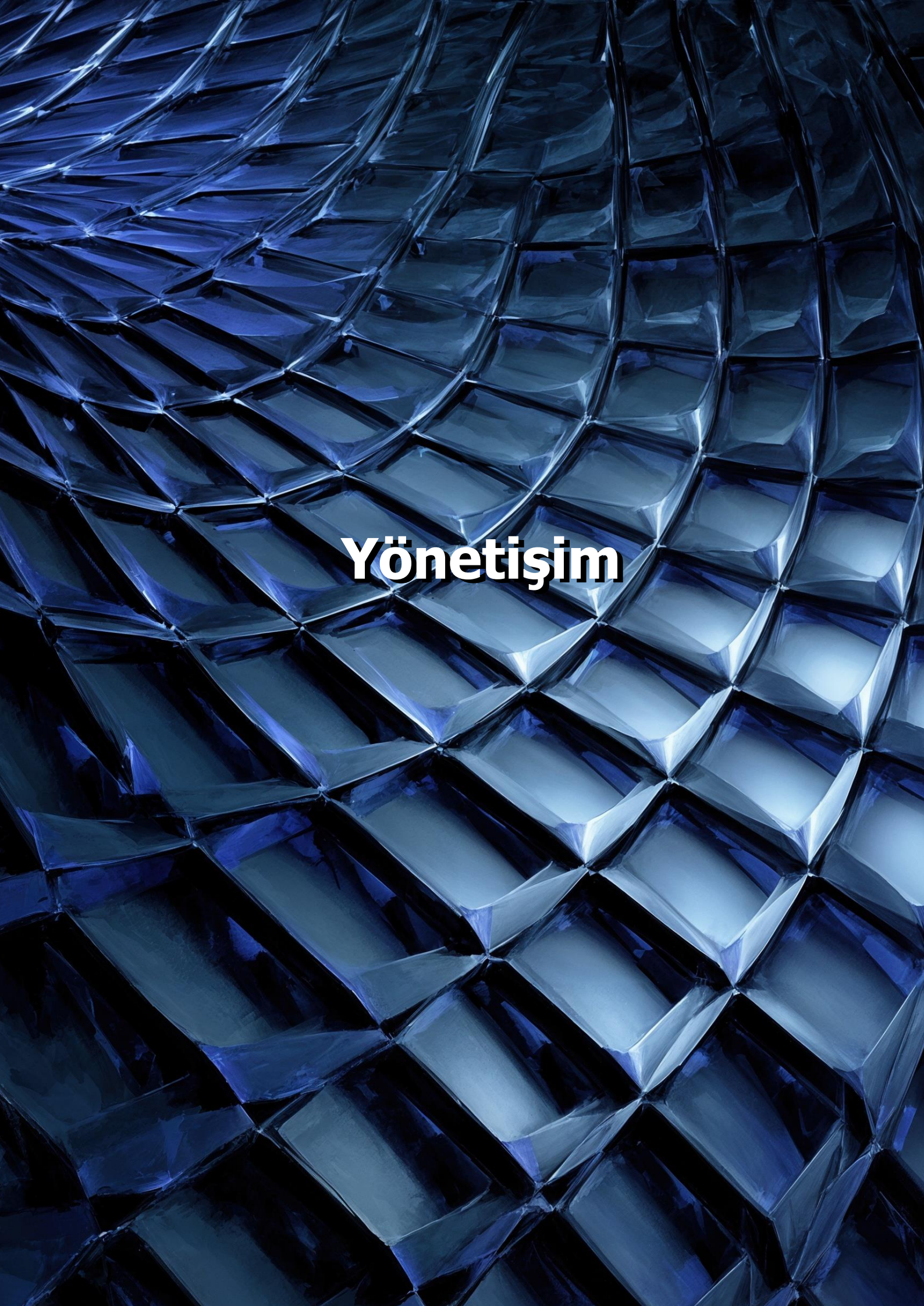
ISO/IEC 23894:2023 Information technology—Artificial intelligence—Guidance on risk management. Geneva, Switzerland: ISO/IEC.

ISO/IEC 42001:2023 Artificial intelligence—Management system (AIMS)—Requirements. Geneva, Switzerland: ISO/IEC.

ISO/IEC 5338:2024 Artificial intelligence—AI system life cycle processes. Geneva, Switzerland: ISO/IEC.

ISO/IEC/IEEE 14764:2006 Software engineering—Software life cycle processes—Maintenance. Geneva, Switzerland: ISO/IEC/IEEE.

ISO/IEC/IEEE 12207:2017 Systems and software engineering—Software life cycle processes. Geneva, Switzerland: ISO/IEC/IEEE.


Yönetişim

7. YÖNETİŞİM

Mehmet Ali Günay[a], Dr. Merve Ayyüce Kızrak[b],
Ersin Tufan Yalvaç[c]

[a] *Çevre Şehircilik ve İklim Değişikliği Bak., YZ ve Yenilikçi Teknolojiler D. Bşk.*
[b] *Kurucu & CEO, HUX AI, ayyucekizrak@gmail.com*
[c] *TBD Ankara Şubesi Denetleme Kurulu Başkanı, ersinyalvac@gmail.com*

7.1. Kamu Kurumları için Yapay Zekâ Standartları

7.1.1. Stratejik Çerçeve ve Uluslararası Bağlam

Bu bölüm, gelişmiş ve özellikle genel amaçlı YZ sistemlerinin kamu yönetimi için sunduğu fırsatları ve riskleri yönetmek amacıyla hazırlanmıştır. Kamu kurumlarının YZ teknolojilerini edinim, geliştirme ve kullanım süreçlerinde güvenlik, şeffaflık, hesap verebilirlik ve sürdürülebilirlik ilkelerini somutlaştırmak hedeflenmektedir. Çalışmanın temeli, iki ana uluslararası referansa dayanmaktadır: teknik süreçler için ISO/IEC 5338:2023 YZ Yaşam Döngüsü Standardı ve yönetişim/güvenlik politikaları için Birleşik Krallık tarafından yayımlanan Uluslararası YZ Güvenlik Raporu (IASR-2025).

IASR-2025, genel amaçlı YZ'nin hızlı ilerleyişinin politika yapıcılar için bir "kanıt ikilemi" (evidence dilemma) yarattığını vurgulamaktadır. Erken atılacak adımlar yetersiz kalabilirken, geç kalmak toplumu hazırlıksız bırakabilir. Bu nedenle rapor, risk yönetiminin sistem yaşam döngüsünün tamamına yayılmasını ve birden fazla katmanda önlem alınmasını önermektedir. Öne çıkan temel risk alanları şunlardır:

- Pazar yoğunlaşması ve tek hata noktası riski.
- Enerji, su ve hammadde tüketimine bağlı çevresel etkiler.
- Gizlilik, telif hakları ve ikili kullanım (biyogüvenlik, siber saldırılar vb.) tehditleri.

1. **Mevcut Standartlar ve Politika Boşlukları: Ne Var, Ne Yok?**

Mevcut uluslararası belgeler, yapbozun farklı parçalarını sunmaktadır. Bütüncül bir ulusal politika için bu iki parçanın birleştirilmesi kritiktir:

- ISO/IEC 5338:2023 Standardı → "**Nasıl**" Yapılacağını Tanımlar: Bu standart, bir YZ sisteminin gereksinim analizinden imhasına kadar uçtan uca tüm teknik yaşam döngüsü süreçlerini detaylı bir şekilde ortaya koyar. Veri mühendisliği (6.4.8), model geliştirme, doğrulama (6.4.11), sürekli geçerleme (6.4.14), bakım (6.4.16) ve kalite güvence (6.3.8) gibi maddeler, kararname için sağlam bir teknik iskelet sunmaktadır.
- IASR-2025 Raporu → "**Ne Kadar Güvenli Olmalı?**" Sorusunu Cevaplar:

Bu rapor ise, teknik süreçlerin üzerine inşa edilecek politika ve yönetişim mekanizmalarını keskinleştirir. Önerdiği temel mekanizmalar şunlardır:

- **Güvenlik Çantası (Safety Case):** Kanıt yükünü geliştiriciye veren bir risk ispatı yaklaşımı.
- **ALARP (As Low As Reasonably Practicable):** Kabul edilebilir risk tolerans seviyelerinin belirlenmesi.
- **Üç Savunma Hattı:** Kurumsal risk yönetimini katmanlara ayıran bir model (ürün ekipleri, bağımsız risk birimi, iç denetim).
- **Zorunlu Olay Bildirimi ve İhbarcı Koruması:** Şeffaflık ve hesap verebilirliği artıran mekanizmalar.
- **Açık Ağırlık (Open-Weight) Modeller İçin Özel Kontroller**: Kötüye kullanım ve "yükseltme riski" analizleri.
- **Çevresel Etki Raporlaması**: Enerji ve su tüketiminin şeffaflaştırılması.

Temel Öneri: Hazırlanacak kararnamede ISO/IEC 5338 standardı "zorunlu teknik süreç standardı" olarak referans alınmalı;

IASR-2025'in yukarıda belirtilen politika ve yönetişim gereklilikleri ise **"bağlayıcı yasal hükümler"** haline getirilmelidir.

2. **Kararnameye Eklenmesi Gereken Maddeler (Fark Analizi)**

Aşağıdaki başlıklar, IASR-2025'ten alınan politika gereksinimlerinin ISO 5338'deki ilgili süreçlerle nasıl ilişkilendirileceğini ve kararname maddesine nasıl dönüştürülebileceğini göstermektedir.

a. Güvenlik Çantası (Safety Case) Zorunluluğu

- Ne Eklenmeli? Yüksek riskli ve genel amaçlı YZ sistemlerini geliştiren veya kamuya sunan kurumlar, sistemin "maksimum risk eşiklerini aşmadığını" kanıtlayan (tehdit senaryoları, kanıt setleri, azaltım önlemleri içeren) bir güvenlik davası sunmadan sistemi canlıya alamaz. Bu, kanıt yükümlülüğünü geliştiriciye devreder.
- ISO 5338 Bağlantısı: Doğrulama (6.4.11), Geçerleme (6.4.13), Sürekli Geçerleme (6.4.14).

b. Risk Eşikleri, ALARP ve Marjinal Risk İzleme

- Ne Eklenmeli? Düzenleyici kurum tarafından tanımlanan "kabul edilemez risk" seviyeleri ve ALARP prensibi benimsenmelidir. Geliştiriciler, sistemdeki marjinal risk artışlarını sürekli olarak izlemek ve kademeli artışların gözden kaçmaması için tanımlanmış eşik değerlerine yaklaşılması halinde bildirimde bulunmakla yükümlü olmalıdır.
- ISO 5338 Bağlantısı: Risk Yönetimi (6.3.4), Ölçüm (6.3.7).

c. Kurumsal Yönetişim: Üç Savunma Hattı, Olay Bildirimi ve İhbarcı Koruması

- Ne Eklenmeli? Kurumlarda üç savunma hattı modeli zorunlu kılınmalıdır: 1. Hat: Ürün ve geliştirme ekipleri, 2. Hat: Bağımsız risk/güvenlik birimleri, 3. Hat: İç denetim. Ciddi güvenlik olayları 72 saat içinde ulusal otoriteye bildirilmeli ve merkezi bir olay siciline kaydedilmelidir. Kurum içi ve dışı ihbarcılar için yasal koruma ve teşvik mekanizmaları oluşturulmalıdır.
- ISO 5338 Bağlantısı: Kalite Güvence (6.3.8), Bilgi Yönetimi (6.3.6).

d. Çok Katmanlı Güvenlik: Dağıtım Öncesi Test ve Dağıtım Sonrası İzleme

- Ne Eklenmeli?
- Dağıtım Öncesi: Kırmızı ekip (red-teaming) testleri ve biyogüvenlik, siber saldırı kapasitesi, manipülasyon gibi alanlarda hedefli değerlendirmeler zorunlu olmalıdır.
- Dağıtım Sonrası: Canlı sistemlerde sürekli geçerleme mekanizmaları (giriş/çıkış filtreleri, anomali tespiti) kurulmalı ve performans düşüşü ya da tehlikeli davranış durumunda sistemi otomatik olarak önceki güvenli sürüme döndürecek "rollback" eşikleri tanımlanmalıdır.

- ISO 5338 Bağlantısı: Sürekli Geçerleme (6.4.14), Operasyon (6.4.15), Bakım (6.4.16).

e. Açık Ağırlıklı (Open-Weight) Modeller İçin Ek Kontroller

- Ne Eklenmeli? Açık ağırlıklı modellerin fayda-risk dengesi dikkatle yönetilmelidir. Model ağırlıklarını yayınlamadan önce, kötüye kullanım potansiyelini (ikili kullanım tehditleri) ve mevcut modellere kıyasla yarattığı marjinal riski analiz eden bir değerlendirme raporu zorunlu olmalıdır. Yüksek risk durumunda, modelin yayınlanması yerine kısıtlı API veya lisanslı kullanım gibi yöntemler tercih edilmelidir. Ayrıca, bu modellerde toplu güncelleme (patch) imkânı olmadığından, sözleşmelere güncelleme takibi ve kapatma senaryoları eklenmelidir.
- ISO 5338 Bağlantısı: Risk Yönetimi (6.3.4), Tedarik (6.1.2).

f. Çevresel Etki (Enerji-Su-Emisyon) Raporlaması

- Ne Eklenmeli? YZ modelinin hem eğitim hem de çıkarım (inference) aşamaları için harcanan enerji tüketimi, su kullanımı ve lokasyon bazlı emisyon yoğunluğu raporlanmalıdır. Özellikle büyük eğitim süreçleri için düzenleyici kuruma ön bildirim yapılması ve yeşil enerji kullanımı, verimli soğutma gibi azaltım planları sunulması zorunlu kılınmalıdır. Bu, YZ'nin veri merkezi enerji tüketimindeki %10-28'lik payı ve artan su ayak izi göz önüne alındığında kritik bir adımdır.
- ISO 5338 Bağlantısı: Operasyon (6.4.15), Altyapı Yönetimi (6.2.2).

g. Tedarik ve Sözleşme Hükümleri

- Ne Eklenmeli? Kamu alımlarında yapılan sözleşmelere; veri hakları, tedarikçinin güvenlik davası sunma yükümlülüğü, çevresel raporlama ve olay bildirim sorumlulukları gibi maddeler açıkça eklenmelidir. Kavram kanıtlama (PoC) aşamaları, operasyonel risklerin yönetimi şartına bağlanmalıdır. Pazar yoğunlaşmasını önlemek için çoklu sağlayıcı stratejileri teşvik edilmelidir.

3. Sonuç

ISO/IEC 5338 standardını temel süreç çerçevesi olarak benimsemek doğru ve güçlü bir yaklaşımdır. Ancak bu teknik iskelet tek başına yeterli değildir. Eksik halka, IASR-2025 raporunda özetlenen politika ve yönetişim araçlarının (güvenlik davası, olay bildirimi, ALARP, çevresel raporlama vb.) yasal olarak bağlayıcı maddelere dönüştürülmesidir. Bu bütüncül

yaklaşım, Türkiye'nin YZ teknolojilerinden güvenli, şeffaf ve sürdürülebilir bir şekilde faydalanmasını sağlayacaktır.

7.1.2. Türkiye'deki Mevcut YZ Politikaları ve Yasal Düzenlemeler

YZ konusunda Türkiye'nin son beş yılı gözden geçirilecek olunursa; yalnızca teknolojik bir yenilikçilik konusu olarak değil, aynı zamanda stratejik kalkınma, veri egemenliği ve ulusal güvenlik boyutlarıyla da ele alındığı görülmektedir. 2020'den bu yana yayımlanan politika belgeleri, etik rehberler ve yasal düzenlemeler, Türkiye'nin küresel ölçekte gelişen bu alanda kendi yolunu inşa etmeye çalıştığını göstermektedir. Raporun bu bölümünde 2020–2025 arasında atılan adımları kronolojik olarak incelemekte ve geleceğe dönük planlara ışık tutmaktadır.

1. ***Yapay Zekâ Alanında İlk Kurumsal Adım, 2020***

Türkiye'de YZ alanındaki kurumsal yapılanma TÜBİTAK BİLGEM bünyesinde Yapay Zekâ Enstitüsü'nün Kasım 2020'de kuruluşuyla somutlaşmaya başladı. Bu enstitü, ulusal ölçekte YZ araştırmalarının odak noktası olmayı hedefledi. Bu yapı, yalnızca teknik araştırmalar yapmakla kalmayıp aynı zamanda akademi, sanayi ve kamu kurumları arasında işbirliğini geliştirmeyi amaçlamıştır. Enstitü, finans, e-ticaret, sağlık, tarım ve eğitim alanlarına kadar geniş bir yelpazede Ar-Ge projeleri yürüterek Türkiye'nin YZ'de kendi teknolojilerini geliştirme kapasitesini güçlendirmeyi hedeflemiştir. Böylelikle Türkiye, YZ ekosistemini sistematik bir şekilde destekleyen ilk kurumsal platformuna kavuşmuş oldu [1].

2. ***Strateji ve Etik Çerçevenin Oluşturulması, 2021***

2021 yılı, Türkiye'de YZ politikaları açısından bir dönüm noktası olmuştur. Bu dönemde hem etik ilkelere yönelik bağlayıcı olmayan rehberler hem de ülkenin ilk Ulusal YZ Stratejisi yayımlandı.

2.1. KVKK Rehberi

Kişisel Verileri Koruma Kurumu (KVKK), 15 Eylül 2021'de "YZ Alanında Kişisel Verilerin Korunmasına Dair Tavsiyeler" başlıklı bir rehber yayımladı. Bu belge, YZ sistemlerinde şeffaflık, ayrımcılık yapmama, insan müdahalesi imkânı ve veri minimizasyonu gibi temel ilkeleri vurguladı. OECD ve Avrupa Konseyi'nin benzer belgelerinden esinlenen rehber, bağlayıcı olmasa da uygulayıcılar için güçlü bir referans noktası oluşturdu. Özellikle kamu kurumları ve özel sektör, YZ projelerinde bu ilkeleri dikkate alarak veri koruma konusundaki riskleri azaltmaya yönlendirildi [2].

2.2. Ulusal Yapay Zekâ Stratejisi 2021–2025

Aynı yıl, Türkiye'nin ilk Ulusal Yapay Zekâ Stratejisi 20 Ağustos 2021'de Resmî Gazete'de yayımlandı. Cumhurbaşkanlığı Dijital Dönüşüm Ofisi ve Sanayi ve Teknoloji Bakanlığı koordinasyonunda hazırlanan strateji, *"Müreffeh bir Türkiye için çevik ve sürdürülebilir YZ ekosistemiyle küresel ölçekte değer üretmek"* vizyonuyla ortaya çıktı [3]. Belge, insan kaynağı yetiştirme, Ar-Ge ve girişimcilik, kaliteli veri ve altyapı erişimi, sosyo-ekonomik uyum, uluslararası işbirlikleri ve yapısal dönüşüm olmak üzere altı stratejik önceliğe odaklandı. Bu çerçevede 24 amaç ve 119 tedbir tanımlandı. Tedbirlerin uygulanması için Cumhurbaşkanı Yardımcısı başkanlığında bir Yönlendirme Kurulu oluşturuldu. 2025 yılına kadar Ulusal YZ Stratejisi Sekreteryası Cumhurbaşkanlığı Dijital Dönüşüm Ofisi (Mülga) ve Sanayi ve Teknoloji Bakanlığı yürüttü. Strateji kapsamında Kamu YZ Ekosistemi Proje Çağrısı ve Kamu Veri Alanı gibi ulusal anlamda kıymetli pek çok çalışma başlatılmıştır [4].

3. Hukuki Düzenleme Çabaları ve Güncellenen Strateji, 2024

2024 yılı, Türkiye'nin YZ'ye dair ilk doğrudan yasa teklifinin gündeme geldiği yıl oldu. Aynı zamanda, hızlı küresel gelişmelere uyum sağlamak için ulusal stratejide güncelleme yapıldı.

3.1. Yapay Zekâ Kanun Teklifi

24 Haziran 2024'te TBMM'ye sunulan YZ Kanun Teklifi, Türkiye'de bu alana özgü ilk yasa girişimi olarak dikkat çekti. Tasarı, YZ'nin güvenli, etik ve adil şekilde kullanımını sağlamak için genel bir çerçeve çizmeyi hedefledi. İlk kez "yapay zekâ" terimi yasal olarak tanımlandı ve yüksek riskli uygulamaların özel denetime tabi tutulması öngörüldü. Ayrıca, etik ilkelere aykırı uygulamalar ve ihlaller için yüksek tutarlı idari para cezaları düzenlendi. Henüz yasalaşmamış olan teklif, 2025 itibarıyla komisyon aşamasına taşınmıştır [5].

3.2. Ulusal Yapay Zekâ Stratejisi 2024–2025 Eylem Planı

Cumhurbaşkanlığı Dijital Dönüşüm Ofisi, Ağustos 2024'te mevcut stratejiyi güncelleyerek yeni bir eylem planı yayımladı [6]. Bu plan, özellikle üretken YZ, Türkçe büyük dil modelleri, yüksek performanslı hesaplama altyapıları ve veri erişimi konularına odaklandı. Aynı zamanda, girişimcilik ekosisteminin desteklenmesi ve uzman insan kaynağı yetiştirilmesi önceliklendirildi. Böylece Türkiye, hızla değişen küresel teknoloji trendlerine uyum sağlamak için stratejik önceliklerini yeniden şekillendirmiş oldu.

4. Meclis Araştırmaları ve Eğitim Politikaları, 2025

2025 yılı, YZ'nin yalnızca teknolojik değil toplumsal ve sektörel boyutlarının da daha geniş ele alındığı bir dönem oldu.

4.1. TBMM Yapay Zekâ Araştırma Komisyonu

Ocak 2025'te kurulan TBMM Araştırma Komisyonu, dört ay süren çalışmalar sonucunda kapsamlı bir rapor hazırladı. Raporda, Türkiye'nin YZ'de rekabet gücünü artırabilmesi için veri egemenliği, insan kaynağı geliştirme, etik ilkelerin kurumsallaştırılması ve uluslararası işbirlikleri gibi başlıklarda somut öneriler sunuldu [7]. Bu rapor, önümüzdeki dönemde hazırlanacak yasal düzenlemeler ve strateji belgeleri için bir yol haritası niteliği taşımaktadır.

4.2. Eğitimde Yapay Zekâ Politika Belgesi (2025–2029)

Millî Eğitim Bakanlığı, 17 Haziran 2025'te eğitim alanında YZ'ye yönelik ilk kapsamlı politika belgesini yayımladı [8]. Bu belge, eğitimde YZ kültürünün oluşturulması, müfredata entegrasyonu, karar destek mekanizmalarının güçlendirilmesi ve dijital altyapının geliştirilmesi gibi dört stratejik hedef etrafında yapılandırıldı. Kısa, orta ve uzun vadeli olmak üzere 40 eylem adımı belirlendi. Öğretmenlere YZ eğitimi verilmesi, öğrenciler için YZ okuryazarlığı programları ve etik kurulların oluşturulması gibi adımlar, belgenin somut uygulamaları arasında yer aldı.

5. Geleceğe Dönük Planlar

Türkiye'nin YZ politikalarında 2025 sonrası dönemde üç ana yönelim öne çıkmaktadır. İlk olarak, TBMM'de görüşülen Yapay Zekâ Kanunu'nun yasalaşması, ülkenin bağlayıcı bir yasal çerçeveye kavuşmasını sağlayacaktır. İkinci olarak, 2021–2025 stratejisinin sona ermesiyle birlikte, 12. Kalkınma Planı doğrultusunda yeni bir ulusal YZ stratejisi hazırlanması beklenmektedir. Bu stratejinin üretken YZ, veri alanları, YZ2nin iş gücüne ve ekonomiye etkisi, YZ risk yönetimi ve etik yönetim gibi konulara daha fazla ağırlık vereceği öngörülmektedir. Üçüncü olarak, sağlık, adalet ve tarım gibi alanlarda sektörel yol haritaları hazırlanması gündemdedir. KVKK da üretken YZ'ye ilişkin yeni bir rehber çalışmasına başlamış, TÜBİTAK ise Türkçe dil modelleri ve otonom sistemler gibi alanlarda Ar-Ge desteklerini artırmıştır.

6. *Sonuç ve Değerlendirme*

Türkiye, 2020'den bu yana YZ alanında hem kurumsal hem de hukuki zemin oluşturmaya yönelik önemli adımlar atmıştır. Strateji belgeleri ve etik rehberlerle başlayan bu süreç, kanun teklifleri, Meclis araştırmaları ve sektörel eylem planlarıyla daha da derinleşmiş olsa da ülkenin kurumsal yapılanmasındaki hızlı değişimler ve belirsizlikler, çok hızlı gelişen YZ teknoloji alanında zaman kaybetmesine ve orta güçteki rekabet halinde olduğu diğer ülkeler arasında stratejik ve gücünü artıracak bir pozisyon tutmasını zorlaştırmaktadır. 2025 sonrasında Türkiye'nin önünde, bağlayıcı yasal düzenlemeler, ulusal strateji güncellemeleri ve sektörel uygulama planları ile küresel standartlara uyumlu, insan odaklı bir YZ ekosistemi oluşturma fırsatı bulunmaktadır.

7. *Kaynaklar*

[1] *Kronoloji - TÜBİTAK BİLGEM*. (2025, Ağustos 21). Tübitak Bilgem. https://bilgem.tubitak.gov.tr/kronoloji/

[2] *Yapay Zekâ Alanında Kişisel Verilerin Korunmasına Dair Tavsiyeler.* (2025, Nisan), KVKK, https://www.kvkk.gov.tr/Icerik/7048/Yapay-Zekâ-Alaninda-Kisisel-Verilerin-Korunmasina-Dair-Tavsiyeler

[3] Ulusal Yapay Zekâ Stratejisi. (2021 Ağustos 20), Resmî Gazete, https://www.resmigazete.gov.tr/eskiler/2021/08/20210820-22.pdf

[4] *Türkiye Cumhuriyeti Cumhurbaşkanlığı Dijital Dönüşüm Ofisi - Ulusal Yapay Zekâ Stratejisi 2021-2025*. (n.d.). https://cbddo.gov.tr/uyzs

[5] *TÜRKİYE BÜYÜK MİLLET MECLİSİ*. (n.d.). https://www.tbmm.gov.tr/Yasama/KanunTeklifi/e21539a0-888a-4500-81be-01904a918c53

[6] *Türkiye Cumhuriyeti Cumhurbaşkanlığı Dijital Dönüşüm Ofisi - Ulusal Yapay Zekâ Stratejisi 2021-2025*. (n.d.). https://cbddo.gov.tr/uyzs

[7] *TBMM Yapay Zekâ Araştırma Komisyonu Başkanı Dönmez: Amacımız teknolojiyi desteklerken toplumsal zararların önüne geçmek*. (n.d.). https://www.aa.com.tr/tr/politika/tbmm-yapay-Zekâ-arastirma-komisyonu-baskani-donmez-amacimiz-teknolojiyi-desteklerken-toplumsal-zararlarin-onune-gecmek/3626795

[8] Başkanlığı, T. M. E. B. B. İ. D. B. İ. H. D. (n.d.). *"EĞİTİMDE YAPAY ZEKÂ POLİTİKA BELGESİ VE EYLEM PLANI" YÜRÜRLÜĞE GİRDİ*. © Millî Eğitim Bakanlığı. https://www.meb.gov.tr/egitimde-yapay-zek-politika-belgesi-ve-eylem-plani-yururluge-girdi/haber/37531/tr

7.1.3. Türkiye'de Yapay Zekâ Yönetişimi için Uyarlanabilir Düzenleyici Çerçeve Önerileri

1. Amaç ve Yaklaşım

Türkiye'nin YZ ekosistemi; strateji belgeleri, meclis çalışmaları, veri koruma rehberleri ve sivil/toplumsal paydaşların katkılarıyla olgunlaşıyor. Ancak teknolojinin hızı, tek seferlik ve katı mevzuat yerine öğrenen, risk-temelli ve bağlama duyarlı bir düzenleyici mimariyi gerektiriyor. Raporun bu bölümü, Türkiye'deki mevcut gelişmeleri ve uluslararası örnekleri birlikte okuyarak her konu için öneri, bu önerinin nasıl gerçekleştirilebileceği ve gerçekleşmesi halinde beklenen etki bağlamında uygulanabilir adımlar sunmayı amaçlamaktadır.

1.1. Kapsam ve Mevcut Durum: UNDP Türkiye için Veri Yönetişimi Çerçevesi

YZ yönetişiminin inşa edilmesinde, sağlam bir veri yönetişimi altyapısı kritik önemdedir. UNDP'nin 2024'te yayımladığı "Türkiye için Veri Yönetişimi Çerçevesi Öneri Raporu", bu alanda ulusal yol haritası için temel bir referans noktası oluşturmuştur [1]. Raporda, veri yönetişiminin beş temel sütun üzerine inşa edilmesi gerektiği belirtilmiştir: politika, mevzuat ve düzenlemeler; kurumlar, mekanizmalar ve süreçler; insan kaynağı; teknoloji ve altyapı; işbirlikleri ve ortaklıklar.

Türkiye'de bugüne kadar atılan adımlar –Kişisel Verilerin Korunması Kanunu, Ulusal Veri Sözlüğü, Açık Veri Portalı, TÜİK öncülüğünde resmi istatistik ekosistemi, TUCBS gibi coğrafi bilgi sistemleri– parçalı fakat ilerletilebilir bir temel sunmaktadır. Bununla birlikte, UNDP raporunda da vurgulandığı gibi, verinin silo yapılarında tutulması, farklı kurumlarda farklı standartların uygulanması, veri paylaşımına dair açık düzenleme eksiklikleri ve kapasite farklılıkları mevcut durumun temel sınırlılıklarıdır.

Dolayısıyla Türkiye'de YZ yönetişimine dair düzenleyici çerçevenin inşasında, UNDP'nin ortaya koyduğu bu beş sütunlu yaklaşım kritik bir dayanak oluşturabilir. Veri güvenliği ve mahremiyetin korunması, insan hakları merkezli etik ilkeler, kurumlar arası

birlikte çalışabilirlik, teknik altyapı yatırımları ve kamu-özel-akademi ortaklıkları, YZ'nin güvenilir ve uyarlanabilir biçimde yönetişiminin ön şartlarıdır.

2. Ulusal Gelişmeler ve Türkiye'ye Özgü Tasarım

2.1. Ulusal Yapay Zekâ Stratejisi ve Eylem Planı

Ulusal YZ Stratejisi (2021–2025) ve güncel eylem planları; hukuki rehberlik, ilke ve değerler çerçevesi, güven damgası/belgelendirme ve birlikte geliştirme ortamlarını öngören iki katmanlı bir yönetişim altyapısı tarif ediyor [2].

Öneri: Stratejideki ilke-rehber ve belgelendirme araçlarını, risk sınıfları ile doğrudan ilişkilendirerek kurumsallaştırmak; yüksek riskli alanlar (sağlık, enerji, kamu hizmetleri, adalet) için zorunlu gereklilikleri açık hükümlere bağlamak; düşük riskli alanlarda ise yeniliği hızlandıran düzenleyici kum havuzları (regulatory sandbox) işletmeyi planlamalıdır.

Uygulama: İlgili bakanlıklar ve düzenleyiciler (KVKK dâhil) tarafından ortak bir "Ulusal YZ Uyum Rehberi" yayımlanabilir; rehber risk sınıflarını, belgelendirme ölçütlerini, YZ model/ürün yaşam döngüsüne göre denetimleri ve kum havuzu prosedürlerini tek çatı altında toplar. Kum havuzları için sektör-bazlı çağrılar, etik ve güvenlik protokolleri, bağımsız teknik gözlem ve kamu raporlaması zorunlu kılınır.

Beklenen etki: Düzenleme ve yenilikçilik dengesi korunur; yüksek riskte kamu güveni artar, düşük riskte ürünleşme hızlanır; rekabet olumlu yönde etkilenir; hukuki belirsizlik ve uyum maliyeti azalır.

2.2. TBMM Yapay Zekâ Araştırma Komisyonu

Komisyon 2025'te sunduğu yol haritasında; altyapı, yetkinlik, etik ve hukuk başlıklarını bütüncül ele alarak devlet liderliğinde çok paydaşlı koordinasyon vurgusu yapmıştır [3].

Öneri: Bakanlıklar, düzenleyiciler, sektör ve STK'ların yer aldığı "Ulusal YZ Koordinasyon Kurulu" kurulması; ortak takvim, ölçülebilir hedefler (KPI), yıllık öğrenme raporu ve kamuya açık gösterge tablosu ile ilerlemenin izlenmesine yönelik çalışmalar teşkil edilmelidir.

Uygulama: Merkezi koordinasyon altında sekretarya; sektör düzenleyicilerinin (finans, sağlık, ulaştırma vb.) uygulama rehberi çıkarma yükümlülüğü; üniversite-STK-kamu-özel danışma forumlarıyla periyodik geri bildirim döngüleri oluşturulur.

Beklenen etki: Parçalı girişimler eşgüdümlenir, politika tutarlılığı ve hesap verebilirlik artar; düzenleme çevikliği yükselir.

2.3. KVKK'nın "YZ'de Kişisel Verilerin Korunması" Tavsiyeleri

KVKK; tasarımla mahremiyet, asgari veri, ayrımcılığın önlenmesi, kullanıcı hakları ve itiraz-telafi kanallarını somutlaştıran tavsiyeler yayımladı [4].

Öneri: KVKK rehberini, "Yapay Zekâ Etki Değerlendirmesi" ve "Algoritmik Etki Değerlendirmesi" şablonlarıyla bağlayıp kamu ve kritik sektörlerde zorunlu hâle getirmek; otomatik kararlar için "insan-döngüde" ilkesi ve açıklanabilirlik raporlarını standartlaştırılması.

Uygulama: Bu iki değerlendirme formu e-devlet tabanlı bir portalda; model kartları, veri kartları ve risk tedbir planlarının (mitigation log) yüklenmesi; rastgele ve risk-odaklı denetimler; uyumsuzlukta kademeli yaptırım ve düzeltici-önleyici faaliyet (CAPA) zorunluluğu istenebilir.

Beklenen etki: Temel haklara uyum güçlenir; ayrımcılık/doğruluk sorunları erken tespit edilir; geliştirici-işletici tarafında kurumsal kapasite artar.

2.4. STK ve İş Dünyası Girişimleri

Etik ilkeler, yaşam döngüsü risk yönetimi ve iyi uygulama örnekleri paylaşılıyor; iş dünyası akıllı ve orantılı düzenleme talep ediyor.

Öneri: Eş-düzenleme modeliyle sektör kodları ve gönüllü-zorunlu karması; sertifikasyon-etiketleme (ör. "Güvenilir YZ Damgası") şeffaf kriterlere bağlanmalı.

Uygulama: Sektörel birliklerle birlikte "uyum profili" kılavuzları; bağımsız üçüncü taraf test/denetim kuruluşlarının akreditasyonu; sonuçların kamu şeffaflık panellerinde yayımlanması.

Beklenen etki: Uyum maliyeti öngörülebilir olur; rekabet ve güven pekişir; küresel pazarlara giriş kolaylaşır.

3. Uluslararası Örnekler ve Türkiye'ye Uyarlama

3.1. AB Yapay Zekâ Yasası (AI Act): Risk Temellilik ve Korumalı Alanlar

AB YZ Yasası, kabul edilemez, yüksek, sınırlı ve düşük risk kategorileri üzerinden kademeli yükümlülükler öngörür ve bu yaklaşımı düzenleyici korumalı alanlarla (sandbox) destekler [5].

Öneri: Türkiye'de işlevsel uyum için, söz konusu risk sınıflarının iç mevzuata açık biçimde tercüme edilmesi; kısa vadede sağlık, finans, enerji ve kamu hizmetlerinde sektör bazlı kum havuzlarının hayata geçirilmesi; yetkili otorite ile şikâyet/itiraz mekanizmalarının ise net olarak tanımlanması önerilir.

Uygulama: Bu amaçla ikincil mevzuatta risk sınıfı–gereklilik matrisi yayımlanır; teknik belgelendirme, veri yönetişimi, insan gözetimi ve kayıt tutmaya ilişkin asgari gereklilik listeleri belirlenir. Kum havuzlarında bağımsız etik gözetim ile sızma/test protokolleri işletilir ve düzenli kamu raporlaması zorunlu kılınır.

Beklenen etki: Bu yaklaşım, AB ile mevzuat ve ticari uyumluluğu güçlendirir; güven ile yenilikçilik arasında dengeli bir çerçeve kurar; ihracat kapasitesini ve tedarik zincirlerine entegrasyonu kolaylaştırır.

3.2. Birleşik Krallık: İlke Tabanlı ve Düzenleyici Ağları

Tek bir "büyük yasa" yerine, Birleşik Krallık modeli güvenlik, adillik, şeffaflık ve hesap verebilirlik gibi yatay ilkeleri öne çıkarır ve bu ilkelerin sektör düzenleyicileri tarafından bağlama göre uygulanmasını öngörür [6].

Öneri: Türkiye'de benzer bir yapı için "Düzenleyici Ağlar" oluşturulmalı; finans, sağlık, ulaştırma, enerji ve medya gibi alanlardaki düzenleyiciler ortak ilkeleri benimsemeli ve kendi sektörlerine özgü rehberler yayımlamalıdır. Merkezî koordinasyon ise ölçütleri ve denetim standartlarını belirleyerek bu rehberlerin uyumunu sağlamalıdır.

Uygulama: Bu modelde öncelikle ortak ilkeler karara bağlanır; her düzenleyici kurumda YZ'den sorumlu bir odak noktası birimi oluşturulur. Yıllık "uygulama ve öğrenme" raporları yayımlanır, çapraz denetimler yapılır ve düzenleyiciler arası tecrübe paylaşımı toplantıları düzenlenir.

Beklenen etki: Böylece düzenlemede orantılılık ve esneklik sağlanır, aşırı veya yetersiz düzenleme riski azaltılır ve sektörlere özgü yenilik süreçleri hızlanır.

3.3. Singapur: Model Çerçeve ve YZ Doğrulama (AI Verify)

Singapur örneğinde görüldüğü gibi, gönüllü Model YZ Governance Framework ve YZ Doğrulama (Verify) aracı sayesinde açıklanabilirlik, şeffaflık ve güvenlik ölçütleri pratik test süreçlerine bağlanmaktadır [7].

Öneri: Türkiye'de benzer bir mekanizma için "Türkiye YZ Doğrulama Kriterleri" geliştirilmelidir. Bu kriterler; açıklanabilirlik kartları (model/data card), kırılganlık–çeviklik–dayanıklılık–esneklik testleri, güvenlik ve mahremiyet kontrolleri ile kullanım bağlamı raporlarını içermelidir.

Uygulama: Bu doğrultuda TSE ve üniversiteler işbirliğiyle açık test paketleri hazırlanmalı; akredite laboratuvar ağı kurulmalı; kamu alımlarında "test-geçer" şartı getirilmelidir. Ayrıca geliştiricilere test aracı ve SDK desteği sağlanarak sürece katılımları kolaylaştırılmalıdır.

Beklenen etki: Böyle bir doğrulama sistemi, ürün kalitesi ve güvenilirliği şeffaf hale getirir; KOBİ'lerin uyum maliyetlerini erişilebilir düzeye çeker ve kamu alımları aracılığıyla kalite standartlarını yükseltir.

3.4. Güney Kore: İçerik Etiketleme ve Kamu Destekli Altyapı

Güney Kore örneğinde görüldüğü gibi, yüksek etkili YZ uygulamalarında ilave yükümlülükler getirilmiş; özellikle üretken YZ için içerik etiketleme mekanizmaları ve devlet destekli altyapı yatırımları öne çıkmıştır [8].

Öneri: Türkiye'de de benzer şekilde içerik menşei/etiketleme standardı oluşturulmalı ve kamuya ait bulut/model altyapıları geliştirilmelidir. Bu yaklaşım, telif hakları, dezenformasyon ve güvenlik risklerini azaltacak etiketleme sistemlerini içerirken; Ar-Ge ve test süreçleri için kamu destekli, güvenli model altyapılarının oluşturulmasını sağlamalıdır.

Uygulama: BTK ve TSE koordinasyonunda içerik menşei standartları belirlenmeli, medya ve dijital platformlarla protokoller imzalanmalı, ayrıca güvenli model-veri odaları kurularak araştırmacıların erişimine açılmalıdır.

Beklenen etki: Böyle bir sistem, güvenlik ve telif süreçlerini şeffaflaştırır; yerli inovasyonun sabit maliyetlerini azaltır ve sorumlu üretken YZ uygulamalarının yaygınlaşmasını teşvik eder.

3.5. Kanada: Algoritmik Etki Değerlendirmesi (AIA) ve Kamu Hesap Verebilirliği

Kanada'da kamuda uygulanan Algoritmik Etki Değerlendirmesi (AİD) zorunluluğu, otomatik karar sistemlerine özgü hesap verebilirlik araçlarını hayata geçirerek dikkat çekmektedir [9].

Öneri: Türkiye'de de kamu kurumları için zorunlu AİD uygulaması başlatılmalı ve sonuçlar özet şeffaflık raporlarıyla kamuoyuyla paylaşılmalıdır. Ayrıca otomatik kararların etkilediği bireyler için itiraz-telafi mekanizmaları oluşturulmalı ve tüm süreçlerde insan-döngüde ilkesi güvence altına alınmalıdır.

Uygulama: Kamu alımlarında AİD skorlaması zorunlu hale getirilmeli; yüksek skorların ek kontrol ve değerlendirme süreçlerini tetiklemesi sağlanmalıdır. Kurum içi etik komiteler kurulmalı ve vatandaşların doğrudan başvuru yapabileceği kullanıcı dostu bir itiraz portalı devreye alınmalıdır.

Beklenen etki: Böyle bir sistem, kamuya olan güveni artırır; siyasa hatalarının erken görünür hale gelmesini sağlar ve daha kapsayıcı, vatandaş odaklı hizmet sunumunu güçlendirir.

4. Veri Yönetişimi ve Altyapı "Taşıyıcı Kolon"

Veri alanları (data spaces), açık veri uygulamaları, standartlar ve güvenli paylaşım mekanizmaları, YZ'nin güvenilirliğini ve yeniden kullanılabilirliğini doğrudan şekillendiren temel unsurlardır [10].

Öneri: Türkiye'de Kamu Veri Alanı kurulmalı ve bunun yanı sıra sağlık, mobilite ve üretim gibi öncelikli sektörlerde sektörel veri alanları oluşturulmalıdır. Bu yapıların işleyişini desteklemek üzere bir veri aracılık ve kayıt sistemi, anonimleştirme laboratuvarları ve ortak veri sözlüğü hayata geçirilmelidir.

Uygulama: Ulusal veri sözlüğü ve şema kütüphaneleri genişletilmeli; güvenli veri bağlantısı için IDSA/GAIA-X uyumlu konektörler kullanılmalıdır. Ayrıca veri bağışçılığı/emanet mekanizmaları geliştirilmeli, açık lisans modelleri benimsenmeli ve denetimli erişim protokolleri uygulanmalıdır.

Beklenen etki: Bu yaklaşım veri kalitesini ve erişilebilirliğini artırır, modellerin performansını ve adaletini iyileştirir, özellikle KOBİ'lerin veri temelli inovasyona erişimini kolaylaştırarak giriş bariyerlerini düşürür.

5. Uygulama, Kapasite ve Kamu Alımları

Öneri: Türkiye'de YZ yönetişimini güçlendirmek için üç tamamlayıcı adım öne çıkmaktadır. İlk olarak, yetkinlik ve denetim kapasitesi artırılmalıdır: düzenleyici kurumlarda

YZ konusunda uzman ekipler oluşturulmalı, denetim protokolleri ve model inceleme kılavuzları hazırlanmalı, ayrıca etik ve güvenlik odaklı eğitimler yaygınlaştırılmalıdır. İkinci olarak, kamu alımları yoluyla yönlendirme yapılmalı; "Güvenilir YZ" şartnameleri geliştirilerek açıklanabilirlik, kayıt tutma, veri yönetişimi ve üçüncü taraf test raporları zorunlu hale getirilmelidir. Bu süreçte her pilot uygulamanın ardından Açık Test/Öğrenme raporu sunulması şart koşulmalıdır. Üçüncü olarak ise açık öğrenme döngüsü kurulmalı; yıllık "YZ Güvenlik ve Uyum Raporu" yayımlanmalı, "hata ve yakın ıskalar" bildirim sistemi devreye alınmalı ve kamuya açık gösterge tabloları aracılığıyla şeffaflık sağlanmalıdır.

Uygulama: Bu hedeflere ulaşmak için merkezi bütçe kapsamında özel bir kapasite programı oluşturulmalı; TSE ve üniversiteler aracılığıyla eğitim ve akreditasyon süreçleri yürütülmelidir. Hazine ve Maliye Bakanlığı ile işbirliği içinde sonuç odaklı alım (outcome-based procurement) şablonları geliştirilmeli, pilot projelerden elde edilen kanıt tabanı düzenli olarak yayımlanmalıdır.

Beklenen etki: Bu adımlar sayesinde kurumsal öğrenme hızlanır, piyasada kalite standartları yükselir ve kamu kaynakları daha güvenilir, etik ve sorumlu YZ çözümlerine yönelmiş olur.

6. İzleme-Değerlendirme ve Şeffaflık

Öneri: Türkiye'de YZ yönetişiminin sürdürülebilirliği için kapsamlı bir gösterge seti (KPI/KRI) oluşturulmalı ve düzenli değerlendirme süreçleri işletilmelidir. Denetim sıklığı risk sınıflarına göre farklılaştırılmalı; kamu, özel sektör ve STK'ların temsil edildiği ortak bir izleme komitesi kurulmalı ve vatandaşların doğrudan geri bildirim verebileceği kanallar açılmalıdır.

Uygulama: Her risk sınıfı için çekirdek metrikler belirlenmelidir. Bunlar arasında hata oranı, yanlılık göstergeleri, itiraz süresi, düzeltilen olay sayısı ve geri çağırma/sürüm notları gibi kriterler yer alabilir. Ayrıca bağımsız bir yıllık değerlendirme yapılmalı ve bulgular parlamento ya da ilgili üst kurullara sunulmalıdır.

Beklenen etki: Bu sistem sayesinde hesap verebilirlik güçlenir, politika çevikliği korunur ve toplumsal lisans ile kamu güveni pekiştirilir.

7. Sonuç ve Yol Haritası

Türkiye, risk-temelli, bağlama duyarlı ve öğrenen bir düzenleyici tasarımla hem güvenilirliği hem rekabetçiliği birlikte büyütebilir. Ancak bu sürecin sağlam bir temele oturabilmesi için YZ yönetişimi, UNDP'nin önerdiği beş sütunlu veri yönetişimi çerçevesi ile uyumlu hale getirilmelidir.

- Sektörel kum havuzlarının başlatılması, kamuda zorunlu Algoritmik Etki Değerlendirmesi uygulamalarının hayata geçirilmesi, Düzenleyiciler Ağı'nın kurulması ve "Türkiye YZ Doğrulama Kriterleri"nin yayımlanması; ayrıca UNDP çerçevesinin ilk iki sütunu olan "mevzuat" ve "kurumsal mekanizmalar"a yönelik düzenlemelerin başlatılması.
- Veri alanlarının, güvenli veri paylaşım mekanizmalarının ve kamu test laboratuvarlarının işletilmesi; ulusal veri sözlüğü ve şema kütüphanelerinin genişletilmesi; böylece UNDP çerçevesinin üçüncü ve dördüncü sütunları olan "insan kaynağı" ve "teknolojik-altyapı"nın güçlendirilmesi
- AB uyumluluğunun sağlanması, sertifikasyon pazarının gelişmesi ve ihracat avantajlarının yakalanması; uluslararası veri ekosistemlerine entegrasyon ve çok paydaşlı işbirlikleriyle UNDP çerçevesinin beşinci sütunu olan "ortaklıklar"ın kurumsallaşması.

Bu yol haritası, Türkiye'nin YZ yönetişimini yalnızca uyarlanabilir ve risk temelli değil, aynı zamanda veri yönetişimi ile entegre, bütüncül ve sürdürülebilir bir yapı haline getirecek; ülkeyi bölgesel bir referans noktası konumuna taşıyacaktır.

8. Kaynaklar

[1] *Veri Yönetişimi Çerçevesi - Türkiye için Tavsiye Raporu*. (n.d.). UNDP. https://www.undp.org/tr/turkiye/publications/veri-yonetisimi-cercevesi-turkiye-icin-tavsiye-raporu

[2] *Türkiye Cumhuriyeti Cumhurbaşkanlığı Dijital Dönüşüm Ofisi - Ulusal Yapay Zekâ Stratejisi 2021-2025*. (n.d.). https://cbddo.gov.tr/uyzs

[3] *TBMM Yapay Zekâ Araştırma Komisyonu Başkanı Dönmez: Amacımız teknolojiyi desteklerken toplumsal zararların önüne geçmek*. (n.d.). https://www.aa.com.tr/tr/politika/tbmm-yapay-Zekâ-arastirma-komisyonu-

baskani-donmez-amacimiz-teknolojiyi-desteklerken-toplumsal-zararlarin-onune-gecmek/3626795

[4] *Yapay Zekâ Alanında Kişisel Verilerin Korunmasına Dair Tavsiyeler.* (2025, Nisan), KVKK, https://www.kvkk.gov.tr/Icerik/7048/Yapay-Zekâ-Alaninda-Kisisel-Verilerin-Korunmasina-Dair-Tavsiyeler

[5] *Article 57: AI Regulatory Sandboxes | EU Artificial Intelligence Act*. (n.d.). https://artificialintelligenceact.eu/article/57/

[6] *Introducing the AI Safety Institute*. (2024, January 17). GOV.UK. https://www.gov.uk/government/publications/ai-safety-institute-overview/introducing-the-ai-safety-institute

[7] *PDPC | Singapore's approach to AI governance*. (n.d.). https://www.pdpc.gov.sg/help-and-resources/2020/01/model-ai-governance-framework

[8] *South Korea leverages open government data for AI development*. (2024, October 28). https://govinsider.asia/intl-en/article/south-korea-leverages-open-government-data-for-ai-development

[9] Secretariat, T. B. O. C. (2025, June 24). *Algorithmic Impact Assessment tool*. Canada.ca. https://www.canada.ca/en/government/system/digital-government/digital-government-innovations/responsible-use-ai/algorithmic-impact-assessment.html

[10] Kızrak, M. Ayyüce, *"Kamuda Veriye Dayalı Yenilikçiliği Destekleyecek Bir Platform Olarak Kamu Veri Alanı"*, Anahtar Dergisi, sayfa: 17-19, 2024 Mayıs, https://edergi.sanayi.gov.tr/File/Journal/2024/5/Flipbook/assets/SAVGM_ANAHTAR_MAYIS2024%20web.pdf

7.1.4. “Yapay Zekâ Yönetişim Düzenlemeleri Takipçisi: 2025” Makale Özeti

"*AI Governance Regulations Tracker: 2025 Edition - What Businesses Need to Know About Responsible and Trustworthy AI Development and Use*" başlıklı makalenin ana hatlarını ve işletmeler için çıkarımlarını içeren bir özet aşağıdadır.

Bu makale, 2025 yılı itibarıyla YZ yönetişiminin artık teorik bir tartışma olmaktan çıkıp, işletmeler için somut yasal yükümlülükler ve stratejik bir zorunluluk haline geldiğini vurgulamaktadır. Gönüllülük esasına dayalı etik ilkeler dönemi sona ererken, yerini yaptırımları olan küresel ve bölgesel düzenlemeler almaktadır.

Küresel Düzenlemelere Genel Bakış

Makale, 2025'te yürürlüğe giren veya etkisi artan üç ana düzenleyici yaklaşımı ele alıyor:

1. Avrupa Birliği- YZ Yasası: Bu yasa, küresel standartları belirleyen en kapsamlı düzenlemedir. Risk tabanlı bir yaklaşım benimser ve YZ sistemlerini dört kategoriye ayırıyor:

- **Kabul Edilemez Risk:** Sosyal puanlama gibi sistemler tamamen yasaklanmaktadır.
- **Yüksek Risk:** İşe alım, kredi notlaması, kritik altyapı gibi alanlarda kullanılan YZ'ler için şeffaflık, veri kalitesi, insan gözetimi ve siber güvenlik gibi katı kurallar getirilmektedir. İşletmelerin bu sistemleri piyasaya sürmeden önce uygunluk değerlendirmesinden geçirmesi zorunludur.
- **Sınırlı Risk:** Chatbot'lar gibi sistemlerin, kullanıcıya bir YZ ile etkileşimde olduğunu açıkça belirtmesi gerekir.
- **Minimal Risk:** Video oyunları veya spam filtreleri gibi sistemler için herhangi bir ek yükümlülük getirilmez.

2. Amerika Birleşik Devletleri- Sektörel ve Eyalet Bazlı Yaklaşım: ABD'de AB gibi tek bir federal yasa yoktur. Bunun yerine, sektörel düzenlemeler (finans, sağlık) ve eyalet yasaları (Kaliforniya, Colorado) ön plandadır. Odak noktası daha çok ayrımcılıkla mücadele, tüketici haklarının korunması ve şeffaflıktır. Başkanlık kararnameleri ve NIST (Ulusal Standartlar ve Teknoloji Enstitüsü) tarafından geliştirilen YZ Risk Yönetimi Çerçevesi, şirketler için yol gösterici olmaktadır.

3. Diğer Bölgeler (Çin, İngiltere, Kanada): Çin, YZ'yi daha çok sosyal istikrar ve içerik denetimi odağında düzenlerken; İngiltere ve Kanada gibi ülkeler, AB modelinden esinlenen ancak daha esnek ve inovasyon odaklı kendi çerçevelerini geliştirmektedir.

İşletmelerin Odaklanması Gereken Temel İlkeler

Makaleye göre, düzenlemelerin coğrafi farklılıklarına rağmen, tüm işletmelerin benimsemesi gereken evrensel ilkeler şunlardır:

- **Hesap Verebilirlik (Accountability):** Bir YZ sisteminin neden olduğu zarardan kimin sorumlu olduğu net bir şekilde tanımlanmalıdır. Şirket içinde YZ yönetişiminden sorumlu birimler veya kişiler belirlenmelidir.
- **Şeffaflık ve Açıklanabilirlik (Transparency & Explainability):** YZ sistemlerinin nasıl karar verdiğini ("kara kutu" olmaktan çıkarıp) hem denetçilere hem de son kullanıcılara anlayabilecekleri bir dilde açıklayabilmek kritik öneme sahiptir.
- **Adalet ve Önyargı Azaltma (Fairness & Bias Mitigation):** YZ modellerini eğitmek için kullanılan verilerdeki potansiyel önyargıları tespit etmek ve bu önyargıların ayrımcı sonuçlara yol açmasını engellemek yasal bir zorunluluktur. Algoritmaların düzenli olarak denetlenmesi gerekir.
- **Veri Yönetişimi ve Gizlilik (Data Governance & Privacy):** YZ sistemlerinin kullandığı verilerin kalitesi, güvenliği ve mahremiyeti esastır. GDPR gibi mevcut veri koruma yasaları, YZ sistemleri için de geçerliliğini korumaktadır.
- **İnsan Gözetimi (Human Oversight):** Özellikle yüksek riskli sistemlerde, kritik kararların nihai olarak bir insan tarafından onaylanması veya denetlenmesi mekanizmaları kurulmalıdır. YZ, bir araç olarak kalmalı, kontrolü tamamen ele almamalıdır.

İşletmeler İçin Stratejik Adımlar

Makale, uyum sağlamak ve riskleri yönetmek için işletmelere şu somut adımları atmalarını öneriyor:

- **YZ Envanteri ve Risk Sınıflandırması Yapın**: Şirketinizde mevcut ve planlanan tüm YZ uygulamalarını listeleyin ve bunları AB YZ Yasası'na benzer bir risk matrisine göre sınıflandırılması.

- **Bir Yönetişim Çerçevesi Oluşturun:** YZ projelerini denetleyecek, etik ilkeleri belirleyecek ve yasal uyumu takip edecek disiplinler arası bir "YZ Etik Kurulu" veya sorumlu bir ekip kurulması.
- **Teknik ve Operasyonel Kapasitenizi Geliştirin:** Modellerin açıklanabilirliğini sağlayan teknolojilere (XAI - Explainable AI) ve veri yönetişimi araçlarına yatırım yapılması.
- **Çalışanlarınızı Eğitin:** YZ'nin potansiyel riskleri ve sorumlu kullanım ilkeleri konusunda mühendislerden pazarlama ekiplerine kadar tüm çalışanlarınıza yönelik eğitim programları düzenlenmesi.
- **Dokümantasyon ve Raporlama Alışkanlığı Edinin:** Yüksek riskli YZ sistemlerinizin veri setleri, test sonuçları ve karar süreçleri hakkında detaylı kayıtlar tutulması. Bu, olası bir denetimde hayat kurtarıcı olacaktır.

2025 yılı, YZ'nin sorumlu bir şekilde geliştirilmesi ve kullanılmasının artık bir "tercih" değil, yasalara uyum, marka itibarını koruma ve sürdürülebilir bir rekabet avantajı elde etme açısından bir "zorunluluk" olduğu bir dönüm noktasıdır. Proaktif bir yaklaşım benimseyen şirketler, bu yeni düzenleyici ortamda başarılı olacaktır.


Sonuç ve Değerlendirme

8. SONUÇ VE DEĞERLENDİRME

Yapay Zekâ (YZ) teknolojileri, toplumsal ve ekonomik yaşamın ayrılmaz bir bileşeni haline gelmiş; kamu hizmetlerinden üretim süreçlerine, bilimsel araştırmalardan hukuk ve güvenlik mekanizmalarına kadar kritik alanlarda dönüştürücü bir güç yaratmıştır. Ancak bu dönüşüm, yalnızca teknik ilerleme ile sınırlı olmayıp, etik, hukuki, yönetişimsel ve sosyal boyutlarıyla birlikte yönetilmesi gereken kapsamlı bir ekosistemi beraberinde getirmiş ve getirmektedir. Türkiye Bilişim Derneği (TBD) tarafından gerçekleştirilen bu rapor çalışması, söz konusu ekosistemin sürdürülebilir, güvenilir ve uluslararası standartlarla uyumlu biçimde yapılandırılabilmesi için gerekli bütünsel çerçeveyi ortaya koymaktadır.

8.1. Yapay Zekâ İçin Gereken Çok Boyutlu Yaklaşım

YZ sistemlerinin özellikleri, bu teknolojiyi geleneksel yazılım mühendisliği yaklaşımlarından ayırmakta ve daha karmaşık, öngörülemeyen davranış örüntüleri geliştirmesine neden olmaktadır. Bu bağlamda raporda açıkça ortaya konduğu üzere, YZ'nin sadece teknik bir araç değil; **toplumsal karar süreçleri, insan hakları, veri güvenliği, mahremiyet, etik ve adalet gibi alanlara doğrudan etki eden sosyo-teknik bir sistem** olduğu benimsenmelidir.

YZ'nin yaşam döngüsünün bütün aşamalarının uluslararası kabul görmüş standartlar çerçevesinde tasarlanması, hem projelerin başarısı hem de toplumsal güven açısından bir zorunluluktur. Özellikle:

- Veri toplama süreçlerinin şeffaf ve etik olması,
- Modellerin doğrulanabilir ve açıklanabilir olması,
- YZ çıktılarının izlenebilir ve hesap verebilir olması,
- Sistemlerin düşmanca (adversaryal) saldırılara dayanıklı ve güvenli olması,
- YZ kullanımının sürekli denetime tabi tutulması

raporun pek çok bölümünde altı çizilen ana ilkeler arasında yer almaktadır.

Bu nedenle Türkiye'nin YZ politika ve uygulamalarında **çok paydaşlı, çok katmanlı ve uyarlanabilir** bir yaklaşımı kurumsallaştırması gereken bir döneme girildiği açıktır.

8.2. Standartların Stratejik Rolü: ISO/IEC 5338 ve Diğer Çerçeveler

Raporun en önemli katkılarından biri, **ISO/IEC 5338 YZ Yaşam Döngüsü Standardı** başta olmak üzere, uluslararası standartların kapsamlı bir analizini sunmasıdır. Bu standart, YZ geliştirme süreçlerinin tüm aşamalarını tanımlayarak kurumlara uygulanabilir bir yol haritası sağlamaktadır.

Bu sonuç çerçevesinde üç temel nokta öne çıkmaktadır:

8.2.1. Standartların Birleştirici Gücü

YZ ile ilgili standartlar, farklı sektörlerde ortaya çıkan gereksinimleri ortak bir dile kavuşturarak:

- Kurumlar arası uyumu artırmakta,
- Denetimi kolaylaştırmakta,
- Kamu ve özel sektör uygulamalarını karşılaştırılabilir hâle getirmekte,
- İhale süreçlerinde objektif kriterler oluşturmakta,
- Uluslararası piyasada rekabet gücünü yükseltmektedir.

Nitekim raporda yer verilen değerlendirmeler, Türkiye'de TSE, bakanlıklar ve ihale makamlarının YZ projelerinde uluslararası standartları referans göstermesi halinde hem kamu yararı hem de operasyonel verimlilik açısından önemli kazanımlar elde edileceğini ortaya koymaktadır.

8.2.2. Türkiye İçin Uygulanabilirlik

Uluslararası standartların Türkiye'nin kurum yapısına uyarlanmasının mümkün olduğu; hatta savunma, sağlık, eğitim, çevre yönetimi ve kamu hizmetleri alanlarında uygulanmasının zorunlu hâle geleceği değerlendirilmektedir.

8.2.3. Risk, Güvenlik ve Kalite Güvencesi

YZ sistemlerinin güvenli, etik ve sürdürülebilir biçimde çalışabilmesi için ISO/IEC 23894, 24028, 42001 gibi standartların entegrasyonunun gerekli olduğu belirtilmektedir. Bu standartlar:

- Risk yönetimi,

- Güvenilirlik,
- Güvenlik,
- Kalite güvencesi

alanlarında kurumsal kapasite oluşturmanın temel araçları olarak değerlendirilmektedir.

8.3. Etik ve Sosyal Boyut: Güvenilir YZ'nin Temeli

YZ'nin teknik gelişiminin etik ve sosyal etkilerden ayrı düşünülemeyeceğini, hatta teknik standartların bile etik ilkelere uygun biçimde tasarlanması gerektiği vurgulanmaktadır.

YZ'nin etik kullanımına yönelik temel ilkeler şunlardır:

- Adalet ve ayrımcılığın önlenmesi,
- Şeffaflık,
- Hesap verebilirlik,
- Açıklanabilirlik,
- İnsan gözetimi,
- Veri mahremiyeti,
- Güvenlik ve sağlamlık.

Bu ilkelerin yaşam döngüsünün tüm aşamalarına yerleştirilmesinin hem kamu güveni hem de hukuki uyumluluk açısından kritik önemde olduğu değerlendirilmektedir. Kamu kurumlarında açıklanabilirlik, katılımcı yönetişim ve etik denetim mekanizmalarının zorunlu kılınması da öneriler arasında yer almaktadır.

8.4. Hukuki ve Düzenleyici Gereklilikler: KVKK, GDPR ve AB YZ Yasası

Türkiye'nin YZ alanındaki hukuki altyapısı, mevcut KVKK çerçevesi ile sınırlı olup, YZ'ye özgü riskleri tam olarak karşılamamaktadır. Raporda Avrupa Birliği'nin **GDPR** ve **AB YZ Yasası (AI Act)** ile kıyaslamalı bir analizi sunulmakta ve Türkiye'nin aşağıdaki alanlarda uyumlaştırma yapması gerektiği önerilmektedir:

- Yüksek riskli YZ sistemlerinin tanımlanması,
- Geliştirici ve sağlayıcıların yükümlülüklerinin belirlenmesi,
- Eğitim verisi kullanımında telif ve mahremiyet koruması,

- Açıklanabilirlik ve risk değerlendirmesi zorunlulukları,
- Algoritmik etki değerlendirmesi mekanizmaları.

Bu bağlamda Türkiye için **kademeli geçiş modeli**, **rehber politikalar** ve **etki analizlerine dayalı düzenlemeler** önerilmektedir.

8.5. İnsan Faktörü ve Kullanıcı Merkezli Yaklaşımın Önemi

Yapay Zekâ sistemlerinin başarısı yalnızca algoritmaların doğruluğuna değil, **insan bazlı tasarım süreçlerinin** etkinliğine de bağlıdır. İnsan faktörünün yaşam döngüsüne entegrasyonunu şu noktalarda ele alınmaktadır:

- Kullanıcı deneyimi tasarımı,
- Ergonomi standartları,
- Kullanım bağlamı analizleri,
- İnsan odaklı risk senaryoları,
- İnsan-denetimli YZ mekanizmaları.

Bu yaklaşımın kamu hizmetlerinde güvenilir ve kapsayıcı sistemlerin oluşturulmasında özellikle kritik olduğu vurgulanmaktadır.

8.6. YZ Güvenliği ve Ulusal Kapasite: Stratejik Bir Zorunluluk

YZ sistemleri; düşmanca (adversaryal) saldırılar, veri zehirleme, model manipülasyonu, kavramsal kayma gibi risklere açıktır. Bu nedenle ulusal düzeyde **YZ güvenlik mimarisi**, **siber güvenlik protokolleri**, **açıklanabilirlik mekanizmaları** ve **doğrulanabilir model davranışı** standartlarının oluşturulması gerekmektedir.

Bu bağlamda;

- TR 24027 gibi teknik raporların[9] Türkiye'nin güvenlik ve adalet gereksinimlerini karşılayabilmek için kritik öneme sahip olduğu;
- Kamu kurumlarının güvenilirlik standartlarını benimsemesinin sosyal hizmetlerde, istihdam taramalarında ve diğer kritik konularda daha adil kararlar alınmasını sağlayacağı

değerlendirilmektedir.

[9] https://cdn.standards.iteh.ai/samples/77607/c0664994eace4bd597db80bb10c18dec/ISO-IEC-TR-24027-2021.pdf

8.7. Türkiye İçin Ulusal YZ Yönetişim Modeli

Türkiye'nin YZ alanında bütüncül bir yönetişim çerçevesi oluşturabilmesi için kapsamlı öneriler sunulmaktadır. Bu öneriler arasında şunlar öne çıkmaktadır:

- **Ulusal Yapay Zekâ Komitesi** kurulması,
- **YZ Politika ve Güvenlik Araştırma Merkezleri** oluşturulması,
- Kamu kuruluşlarında **YZ sorumluları** atanması,
- Açık veri politikalarının geliştirilmesi,
- Sayısal içeriklerde **kimlik doğrulama** (content provenance) mekanizmalarının yaygınlaştırılması,
- Derin aldatmaca (deepfake) ve istismara yönelik düzenlemelerin güçlendirilmesi.

Bu yapının Türkiye'nin küresel rekabet gücünü artıracağı; kamu hizmetlerinde güvenilir, şeffaf ve etik bir YZ ekosistemi oluşmasının sağlanacağı değerlendirilmektedir.

8.8. Genel Değerlendirme ve Stratejik Öncelikler

Çalışma sonucunda Türkiye için dört stratejik öncelik ortaya çıktığı görülmektedir:

8.8.1. Uluslararası Standartlara Uyum

YZ'nin tüm süreçlerinde ISO/IEC standartlarının benimsenmesi, hem kamuda hem de özel sektörde kurumsal olgunluğu artıracaktır.

8.8.2. Etik ve İnsan Hakları Odaklı Yaklaşım

Adil, güvenilir, şeffaf ve hesap verebilir YZ sistemlerinin oluşturulması toplumsal kabul açısından zorunludur.

8.8.3. Risk ve Güvenlik Yönetimi

Kritik altyapılarda kullanılan YZ sistemlerinde güvenlik açıklarının yönetilmesi ulusal güvenlik açısından stratejik önemdedir.

8.8.4. Kapsamlı YZ Yönetişim Ekosistemi

Kurumsal roller, denetim mekanizmaları ve politika yapım kapasitesinin güçlendirilmesi gerekmektedir.

8.9. Türkiye İçin Yol Haritası

TBD'nin hazırladığı bu rapor, Türkiye'nin yapay zekâ alanında güvenilir, etik, rekabetçi ve uluslararası standartlara uyumlu bir ekosistem inşa edebilmesi için kapsamlı bir yol haritası sunmaktadır. Bu yol haritasının başarıyla uygulanması için:

- Kamu–özel sektör işbirliği,
- Akademik katkı,
- Düzenleyici kurumların aktif rolü,
- Toplumsal farkındalık,
- Sivil toplumun katılımı

yaşamsal önem taşımaktadır.

Sonuç olarak, YZ teknolojilerinin sunduğu fırsatların adil, güvenli ve sürdürülebilir biçimde değerlendirilmesi, ancak standartlara dayalı yönetişim yapılarının güçlendirilmesiyle sağlanabilecektir.

YZ alanındaki gelişmelerle sürekli olarak gelişip güncellenecek bu çalışmanın Türkiye'nin YZ alanında ulusal stratejisinin şekillenmesine önemli katkılar sağlayacağı açıktır.

Türkiye Bilişim Derneği

www.tbd.org.tr