



T.C.
AMASYA ÜNİVERSİTESİ
FEN BİLİMLERİ ENSTİTÜSÜ

YAPAY ZEKÂ ETİĞİ VE OTONOM SİSTEMLERDE ETİK MODEL
ÖNERİSİ

YÜKSEK LİSANS TEZİ

DENİZ DEVRİM TAŞDEMİR

HAZİRAN

DENİZ DEVRİM TAŞDEMİR

**TEKNOLOJİ VE İNOVASYON YÖNETİMİ
ANABİLİM DALI**

HAZİRAN 2023

**YAPAY ZEKÂ ETİĞİ VE OTONOM SİSTEMLERDE ETİK MODEL
ÖNERİSİ**

Deniz Devrim TAŞDEMİR

**YÜKSEK LİSANS TEZİ
TEKNOLOJİ VE İNOVASYON ANABİLİM DALI**

Danışman

Prof. Dr. Mehmet KARA

İkinci Danışman

Doç. Dr. Leyla Fırze ARDA ÖZALP

**AMASYA ÜNİVERSİTESİ
FEN BİLİMLERİ ENSTİTÜSÜ**

HAZİRAN 2023

Yüksek Lisans Tezi Kabul ve Onay Sayfası

Deniz Devrim TAŞDEMİR tarafından hazırlanan “Yapay Zekâ Etiği ve Otonom Sistemlerde Etik Model Önerisi” adlı tez çalışması aşağıdaki jüri tarafından OY BİRLİĞİ ile Amasya Üniversitesi Teknoloji Ve İnovasyon Yönetimi Anabilim Dalında YÜKSEK LİSANS TEZİ olarak kabul edilmiştir.

Danışman: Prof.Dr. Mehmet KARA

Makine Mühendisliği Bölümü, Amasya Üniversitesi

Bu tezin, kapsam ve kalite olarak Yüksek Lisans Tezi olduğunu onaylıyorum

Başkan : Prof.Dr. Oğuz FINDIK

Bilgisayar Mühendisliği Bölümü, Karabük Üniversitesi

Bu tezin, kapsam ve kalite olarak Yüksek Lisans Tezi olduğunu onaylıyorum

Üye : Prof.Dr. Recep ÇAKIR

Eğitim Bilimleri Bölümü, Amasya Üniversitesi

Bu tezin, kapsam ve kalite olarak Yüksek Lisans Tezi olduğunu onaylıyorum.

Tez Savunma Tarihi: 05 Haziran 2023

Jüri tarafından kabul edilen bu tezin Yüksek Lisans Tezi olması için gerekli şartları yerine getirdiğini onaylıyorum.

.....
Doç. Dr. Ümit Yıldırım

Fen Bilimleri Enstitüsü Müdürü

ETİK BEYAN

Amasya Üniversitesi Fen Bilimleri Enstitüsü Tez Yazım Kurallarına uygun olarak hazırladığım bu tez çalışmada;

- Tez içinde sunduğum verileri, bilgileri ve dokümanları akademik ve etik kurallar çerçevesinde elde ettiğimi,
 - Tüm bilgi, belge, değerlendirme ve sonuçları bilimsel etik ve ahlak kurallarına uygun olarak sunduğumu,
 - Tez çalışmada yararlandığım eserlerin tümüne uygun atıfta bulunarak kaynak gösterdiğimi,
 - Kullanılan verilerde herhangi bir değişiklik yapmadığımı,
 - Bu tezde sunduğum çalışmanın özgün olduğunu,
- bildirir aksi bir durumda aleyhime doğabilecek tüm hak kayıplarını kabullendiğimi beyan ederim.

Deniz Devrim TAŞDEMİR

YAPAY ZEKÂ ETİĞİ VE OTONOM SİSTEMLERDE ETİK MODEL ÖNERİSİ

(Yüksek Lisans Tezi)

Deniz Devrim TAŞDEMİR

AMASYA ÜNİVERSİTESİ
FEN BİLİMLERİ ENSTİTÜSÜ

Haziran 2023

ÖZET

Sanayi devrimi ile birlikte, başlangıçta özellikle fabrikalarda kullanılmak üzere üretilen makineler, diğer endüstri kolları ve sektörler için de insan gücünü taklit ederek belirli görevleri icra edebilen tekniklerin geliştirilmesine öncülük etti. Bunun bir adım ötesinde ise, insan zekâsından esinlenerek tasarlanan ancak çok daha hızlı işlem yapabilen akıllı makinelerin üretimi ile yeni bir döneme girildi. Robotların tamamen insan benzeri olacak şekilde yapılması için belirli bir zaman dilimine ihtiyacımız olsa da, bu teknolojilerin insan zekâsını aşarak daha yüksek seviyelere erişeceğini öngörmek zor değil. Gelecekte bir robot, insan gibi düşünebilecek ve hareket edebilecek, insanın sahip olduğu eksikliklerden, duygusal boşluklardan etkilenmeyen ve yorulmayan, daha nesnel kararlar alabilen bir varlık olacaktır. Yapay zekâ uygulamalarının toplumda kullanılmasıyla birlikte, sistem içinde etik kuralların tekrar düzenlenmesi ve otonom sistemlerin etik ilkelere uyumlu hale getirilmesi önem kazanmıştır. Araştırmada otonom sistemlerin etik kurallara uygun şekilde davranmasını sağlamak bir dizi yöntem ve tekniklerle etik kontrol modeli belirlenmiştir. Modelin amacı, otonom sistemlerin etik değerleri göz önünde bulundurarak ve kişilik haklarını ihlal etmeden karar vermesini sağlamaktır. Altı adımdan oluşan etik kontrol modelinde; yapay zekânın neden olduğu olası riskleri en aza indirmek için, kişisel verilerin kullanımında sorumlulukların belirlenmesi ve olası ihlallerde verilmesi gereken tepkiler ortaya koyulmuştur. Model, sistemin özelliklerinin belirlenmesi, potansiyel etik sorunların belirlenmesi, risk analizinin yapılması, etik kuralların belirlenmesi, sistemin test edilmesi ve sistemin yeniden değerlendirilmesi aşamalarını içerir. Model yapay zekâyâ sahip sistemlerin etik kurallara uygun olarak çalışmasını sağlamak için tasarlanmıştır. Modelde yapay zekâyâ ilgili toplumsal endişelere yol açan, henüz tam olarak ortaya koyulmamış etik ilkeler belirlenmiş ve otonom sistem tasarımında nasıl ele alınması gerektiği analiz edilmiştir. Çok disiplinli bir çalışma olan model, otonom sistemlerin ortaya çıkarabileceği etik sorunları ayrıntılı bir şekilde tespit etmesi, gerekli düzenlemeler ile ilgili destek vermesi, etik ikilemler sonucu ortaya çıkan problemlerin ve eşitsizliklerin üstesinden gelinmesinde önemli bir adım olacaktır.

Sayfa Adedi	: 68
Anahtar Kelimeler	: Yapay zekâ, Etik, uygulamalı etik, Otonom sistemler
Danışman	: Prof. Dr. Mehmet KARA
İkinci Danışman	: Doç. Dr. Leyla Fırze ARDA ÖZALP

ARTIFICIAL INTELLIGENCE ETHICS AND AN ETHICAL MODEL PROPOSAL FOR AUTONOMOUS SYSTEMS

(M. Sc. Thesis)

Deniz Devrim TAŞDEMİR

AMASYA UNIVERSITY

GRADUATE SCHOOL OF NATURAL AND APPLIED SCIENCES

June 2023

ABSTRACT

The industrial revolution sparked the creation of methods that could execute specific jobs by imitating human power for various industries and sectors. This was made possible by the machines that were created expressly for use in factories. A step further than this, a new era has begun with the development of smart machines, which are motivated by human intelligence but have far faster processing speeds. While it will take some time for robots to perfectly resemble humans, it is easy to predict that these technologies will eventually outperform human intelligence and develop to higher levels. A robot of the future will be a thing that can think and act like a human, that is not constrained by human limitations and emotional gaps, that is not worn out, and that is capable of making more unbiased decisions. It is now crucial to restructure ethical norms inside the system and align autonomous systems with ethical principles as a result of the use of artificial intelligence applications in society. To ensure that autonomous systems act ethically, a model of ethical control was developed in the study using a number of methodologies and techniques. The model's goal is to make it possible for autonomous systems to make judgments that respect individual rights and ethical principles. The ethical control model, which consists of six steps, reveals obligations on the use of personal data and responses to potential violations in order to reduce the hazards that artificial intelligence could provide. The model consists of the steps of defining the system's characteristics, seeing potential ethical issues, running a risk analysis, deciding on ethical guidelines, testing the system, and reevaluating the system. The model's goal is to guarantee that artificial intelligence-based systems operate morally. The model has identified unrevealed ethical principles that raise social concerns regarding artificial intelligence and has examined how they should be handled in the creation of autonomous systems. By specifically addressing the ethical issues that may come from autonomous systems and giving support for the appropriate rules, the model, which is a multidisciplinary study, will be a significant step in overcoming the issues and disparities that result from ethical dilemmas.

Page : 68

Keywords : Artificial intelligence, Ethics, Applied Ethics, Autonomous Systems

Supervisor : Prof. Dr. Mehmet KARA

Co- Supervisor : Assoc.Prof. Dr. Leyla Fırze ARDA ÖZALP

ÖN SÖZ VE TEŞEKKÜR

Disiplinlerarası bir çalışmanın ürünü olan bu yüksek lisans tezinde emeği geçen başta kıymetli hocam ve danışmanım Prof.Dr. Mehmet KARA ve eş danışmanım Doç.Dr. Leyla Firuze ARDA ÖZALP'e, son güne kadar desteklerini ve fikirlerini esirgemeyen Prof.Dr. Metin ORBAY, Prof.Dr. Oğuz FINDIK ve Prof.Dr. Recep ÇAKIR'a saygılarımı sunarım.

Çalışma sürecimde yanımda olan sevgili karım Özden ve oğlum Mete'ye de sonsuz teşekkürler. İyi ki varsınız.



İÇİNDEKİLER

	Sayfa
ÖZET	ii
ABSTRACT.....	iii
ÖNSÖZ VE TEŞEKKÜR	iii
İÇİNDEKİLER	v
ÇİZELGELER DİZİNİ	vi
ŞEKİLLER DİZİNİ	vii
1.GİRİŞ	1
1.1.Problemin Durumu.....	1
1.2.Araştırmanın Amacı ve Önemi	2
1.3.Sınırlılıklar	4
1.4.Sayıltılar.....	4
2.KURAMSAL ÇERÇEVE.....	5
2.1. Etik ve Ahlak	5
2.1.1.Etik sistem ve etik davranış	5
2.1.2.Etik sistem ve toplum.....	7
2.2. Zekâdan Yapay Zekâya.....	9
2.2.1.Kavram olarak yapay zekâ ve tarihsel gelişimi	10
2.2.2.Yapay zekânın kullanım alanları	14
2.2.3.Yapay zekâya veri sağlama ve verilerin kişiselliği.....	16
3.YAPAY ZEKÂ ETİĞİ.....	19
3.1.Yapay Zekâ Modellemesinin Aktörleri	20
3.2.Yapay Zekâ ve Tartışmalı Etik Hususlar	22
3.3.Yapay Zekâ Etiğinin Kaynağı.....	23

3.4.Yapay Zekâda Etik Sorumluluk.....	24
3.5.Yapay Zekâda Etik Sorun Alanlarının Belirlenmesi	25
3.6.Yapay Zekâda Etik Kontrol	27
4.YÖNTEM	30
4.1.Modelde Kullanılacak Etik Değerlerin Belirlenmesi.....	30
4.2.Modelde Kullanılan Risk Yönteminin Belirlenmesi.....	33
4.3.Verilerin Toplanması	35
4.4.Model Basamaklarının Belirlenmesi.....	38
4.4.1. Kullanım senaryosu belirlenmesi ve sistem incelemesi.....	308
4.4.2. Tasarımın yapılması tehdit ve hassasiyetlerin belirlenmesi.....	40
4.4.3. Etik olasılıklar, kontrol analizi ve eğitim verilerinin tespiti	42
4.4.3.1. Açıklık.....	45
4.4.3.2. Zarar vermeme	46
4.4.3.3. Adalet.....	47
4.4.3.4. Fayda sağlama ve iyilik	48
4.4.4. Uygulama oluşturulma ve etki analizi	48
4.4.5. Test etme ve risk belirleme	50
4.4.6. Sonuçların raporlanması ve sistemin takibi	52
4.4.7. Etik kontrol modelinin tamamlanması.....	53
5.SONUÇ VE ÖNERİLER.....	56
KAYNAKLAR	59
ÖZGEÇMİŞ	69

ÇİZELGELER DİZİNİ

Çizelge	Sayfa
Çizelge 2.1. Yapay zekâ tarihsel gelişiminde tanımlar.....	12
Çizelge 4.1. Literatür taramasında kullanılan anahtar kelimeler	32



ŞEKİLLER DİZİNİ

Şekil	Sayfa
Şekil 2.1. Etik algı oluşum süreci	9
Şekil 2.2. Kişisel verilerin yaşam döngüsü.....	17
Şekil 3.1. Yapay zekâ teknolojisinde üç grup.....	20
Şekil 3.2. Yasama çalışmaları.....	24
Şekil 3.3. Etik sorun alanları.....	27
Şekil 3.4. Dijital bir ekosistemi teşvik için 5 tavsiye	30
Şekil 4.1. Etik yapay zekânın dört önemli ilkesi	33
Şekil 4.2. Kullanım senaryosu belirlenmesi ve sistem incelemesi	40
Şekil 4.3. Tasarımın yapılması tehdit ve hassasiyetlerin belirlenmesi	42
Şekil 4.4. Etik olasılıklar, kontrol analizi ve eğitim verilerinin tespiti.....	44
Şekil 4.5. Uygulama oluşturulma ve etki analizi	50
Şekil 4.6. Test etme ve risk belirleme.....	52
Şekil 4.7. Sonuçların raporlanması ve sistemin takibi.....	53
Şekil 4.8. Etik kontrol modeli.....	55

1. GİRİŞ

1.1. Problemin Durumu

Yakın geçmişte kullanım alanları kısıtlı iken, günümüzde artan etkinlikleri ile hayatın her alanında makineler tüm toplum hayatını önemli ölçüde etkilemektedir (Khan, 2018). Akıllı cihazların insanlar için kolay erişilebilir hale gelmiş olması, gazetecilik gibi birçok meslekte cep telefonu ve benzeri cihazların kullanımını yaygınlaştırmıştır. Gelişen teknoloji ile birlikte akıllı teknolojilerin kullanım yerleri, fonksiyon alanları ve ürün çeşitlilikleri üssel olarak artmıştır. İşlemcilerin ilk ortaya çıkışından bu yana bilim insanları; yol, yöntem ve buluşlar ile işlemci hızını artırırken boyutlarını küçültmek için donanım ve uygulamalarda yenilik arayışında olmuştur. Geçmişte bir odayı kaplayabilen büyüklükte bir makine olan işlemciler, şuan kolumuzda taşıyabileceğimiz boyutlara kadar küçülmüştür.

İşlemcilerin küçülmesi, taşınabilir cihazlarla bütünleştirilmesi, kablosuz ağlar vasıtasıyla birbirleri ile iletişim kurması ve haberleşme aygıtları ile uyumlu hale getirilmesi, beraberinde bilgi sistem cihazlarının iş ve özel hayatımızın hemen hemen her yerine dâhil olduğu bir duruma yol açmıştır. Sistemlere kazandırılan bu ilave zekâ ise, geline sürecin doğal bir parçası olarak görülebilir (McCarthy, 2007). Sanayi devrimi ile beraber, başlangıçta özellikle fabrikalarda kullanılmak üzere üretilen makineler, devamında diğer endüstri kolları ve sektörler için de üretilmiştir. İnsan marifetiyle yapılan teknik ve tekrarlayan belirli görevleri başarılı şekilde icra etmek üzere geliştirilen makineleşme eğilimi, günümüzde; hataları en aza indirdiği, yorgunluk ve diğer insana özgü olgulardan etkilenmemesi gibi gerekçelerle artarak devam etmektedir (Ford, 2016). Bu gerçeğin bir yansıması olarak, birçok endüstride fabrikalardaki insan emeğine olan ihtiyaç azalmış fakat buna rağmen sanayinin yükselişi beraberinde sayısız yeni iş alanı da açmıştır.

Başlangıçta işsizliğe sebep olacağı düşünülen sanayileşme, kısa sürede birçok yeni meslek yaratmış ve insanlar tarafından daha da geliştirilmesi gereken bir boyut kazanmıştır (Köklü, 2018). Makineleşmenin bir adım ötesinin ise; insan aklından yola çıkılarak modellenen fakat çok daha yüksek hızda işlem yapabilen akıllı makineler olduğu aşikârdır. İnsan ırkının sayısız becerisine baktığımızda; insan imkân ve kabiliyetlerini tamamen taşıyan robotlar yapmak için önümüzde belirli bir zaman olduğu sonucuna varsak da, bu teknolojilerin

zamanlaması dışında insan zekâsını yansıtmakla kalmayıp, daha ötesinde bir noktaya ulaşacağını tahmin etmek zor değildir. Kana enjekte edilebilen ve insan fonksiyonlarının kontrolünde kullanılan nanorobotlar, yapay zekâ ve makinelerin gelebileceği bir üst noktanın olmadığına açık bir örneğidir. Gelecekte bir robot; insan gibi düşünüp davranırken, aynı zamanda insana özgü birtakım eksikliklerden, duygusal boşluklardan etkilenmeden ve yorulmadan, daha fazla veri işleyerek daha nesnel kararlar alabilecektir (Simborg, 2017).

Barrat (2015)'e göre yapay zekâ ile donatılmış makineler; insan aklını yansıtan ve onun gibi davranan veri işleme kabiliyetine sahip ama buna ek öz farkındalık ve belki de kişilik kazanmak suretiyle belleğinde bulunan yazılımı değiştirme kabiliyetine sahip olacaktır. Bu tanımlamadan yapay zekânın bizden bağımsız öğrenen makineler haline geleceği sonucuna varmak mümkündür. Düşüncemizi genişlettiğimizde insan aklını, imkân ve kabiliyetlerini taklit edebilen, insan gibi düşünen ve onun duyarlılıklarını da veri olarak işleyebilen makinelerin, kendi imkân ve kabiliyetlerini tıpkı insanlar gibi ama insanlardan çok daha büyük bir hızla, çok yeni yönlerle genişletmeleri kaçınılmazdır.

1.2. Araştırmanın Amacı ve Önemi

Teknoloji ile yeniden şekillenen yeni dünyada bilimsel gelişmeler ve yapay zekâ uygulamaları, toplumda hesap verebilirliği ve eşitliği sağlamada kişisel verilerin kullanımı ve özellikle korunması konularında dikkatli bir çalışmayı gerekli kılmıştır (Taddeo ve Floridi, 2018). Teknolojinin kullanımı ile istenildiğinde güçlüler tarafından, kullanıcı iletişimlerinin engellenebilir olması, belirli anahtar kelimeleri içeren mesajların yayınlanmasının engellenmesi mümkündür. Akademik araştırmalar son dönemde; etik ile yapay zekâ arasındaki ilişkiye, gizliliğin korumasına, problem çözümünde kullanılan algoritmalara erişim eşitliğine ve sistemlerin kararlılığı gibi birçok ahlaki açıdan önemli konunun temellerine odaklanmıştır (Kizza, 2016). Günümüzde yaygın şekilde kullanılan otonom sistemlerin belirli bir etik kurallar olmadan kullanılmasının toplumsal alanda birtakım endişelere yol açacağı açıktır. Bu çalışmanın temel amacı; yapay zekâyâ sahip sistemlerin ortak olarak belirlenmiş kurallara sahip olmadan kullanılmasının toplumda yol açabileceği endişelere karşı etik ilkelerin korunabilmesi için kontrol modeli ortaya koymaktır.

Ayrıca bu çalışma ile literatürde sıklıkla yer bulmasına rağmen henüz tam olarak çözülmeyen, otonom sistemlerde etik ilkelerin korunması konusunda alınması gereken tedbirlerin belirlenmesi amaçlamaktadır. Çalışmada otonom sistemlerin ortaya çıkarabileceği etik sorunlar ayrıntılı bir şekilde tespit edilmiş ve gerekli düzeltmeler ortaya koyularak, etik ikilemler sonucu ortaya çıkabilecek problemlerin ve eşitsizliklerin üstesinden gelebilmek için bir etik kontrol modeli ortaya koyulmuştur.

Otonom sistemlerin gelişmesiyle beraber kişisel verilerin toplanması, işlenmesi, korunması ve kullanımı konusunda ortaya çıkan sorunların kişisel verilerin korunması açısından etik yönlerini ele alan bu çalışma beş bölümden oluşmaktadır. Birinci bölümde, etik ve ahlak kavramları ve toplum içindeki önemi incelenmiştir. İkinci bölümde, zekâ kavramı ve makinelerden otonom sistemlere geçiş için gerekli olan yapay zekâ ele alınmıştır. Üçüncü bölümde, yapay zekâ etiği kavramına yer verilmiştir. Dördüncü bölümde ise yapay zekâ etiğinde karşılaşılan sorunlar, tespiti ve engellenmesine yönelik çalışmada kullanılan veriler ve yöntemler tanıtılarak, bir etik kontrol modeli ortaya koyulmuştur. Çalışmanın sonuç bölümünde ise genel bir değerlendirme yapılmış ve çalışmanın ana bulguları vurgulanmıştır.

Bu çalışma, otonom sistemlerin yol açabileceği etik sorunlara odaklanması ve yapay zekâ etiği konusunda ayrıntılı bir analiz sunarak model önermesi nedeniyle önemlidir. Ayrıca, bu çalışma, yapay zekâ uygulamalarında kullanılması gereken etik kuralların ve ilkelerin belirlenmesine katkı sağlamaktadır. Çalışma aynı zamanda, yapay zekâ uygulamalarının geliştirilmesi ve kullanımında açıklık, hesap verebilirlik ve eşitliği sağlamak için alınacak tedbirlerin önemini vurgulamaktadır. Son olarak araştırmada, otonom sistemlerde kişisel verilerin gizliliğinin sağlanması için bilginin işlenmesi ve kullanımı sırasında kullanılabilecek etik yaklaşımların neler olabileceği, kullanıcılara nasıl daha fazla kontrol sağlanabileceği de ortaya koyulmuştur.

Araştırmanın bir diğer amacı ise; giderek yaygınlaşan yapay zekâ teknolojilerinde kullanılan kişisel verilerin korunması ve kullanıcı mahremiyetinin önemi konusunda farkındalık yaratmaktır. Araştırmanın sonunda ortaya koyulan model; yapay zekâ geliştiricileri, hukukçular ve karar alıcılar gibi çeşitli paydaşlar için sağlık sektörü ve hukuk gibi hassas alanlarda ortaya çıkabilecek etik sorunlara çözüm bulmak maksadıyla kılavuz olarak da kullanılabilir.

1.3. Sınırlılıklar

Araştırmada dört adet sınırlılık tespit edilmiştir. Bu sınırlılıklar aşağıda verilmiştir.

1. Bu araştırmada önerilen etik kontrol modeli 2021-2022 yılları arasında yapay zekâ ve etik alanında toplanan veriler ile oluşturulmuştur. Araştırmada kullanılan veriler bu zamanlar ile kısıtlıdır.
2. Araştırmanın temelini oluşturan yapay zekâ teknolojilerinin gelişimindeki hız, araştırmada kullanılan modelin zaman zaman güncellenmesine neden olabilecek teknik sınırlamalara yol açmaktadır.
3. Araştırmanın önemli kavramlarından olan etiğin farklı kültürler ve toplumlar arasındaki değişkenliği nedeniyle, ortaya koyulan model her toplum için kabul edilebilir nihai durumu bilmeyecektir. Bu nedenle, model kullanırken içinde bulunulan toplumun farklılık gösteren etik standartları dikkate alınmalıdır.
4. Bu araştırmada literatürde fikir birliğine varılmış ve modele katkısı olacağı değerlendirilen dört etik konu değerlendirilmiştir. Bu nedenle, model dört etik sorunu ele alır ve diğer etik problemleri kapsamaz.

1.4. Sayıtlar

Araştırmada modeli oluşturulurken kullanılan sayıtlar aşağıda verilmiştir.

1. Yapay zekâyâ sahip sistemler, kişisel bilgileri gizli tutacak ve kötü yazılım, istenmeyen erişimlere karşı güvenle saklayacaktır.
2. Yapay zekâ insan odaklı ve insan ihtiyaçlarını karşılamaya yönelik davranacak, toplumsal yarar sağlamaya katkı sağlayacaktır.
3. Yapay zekâ tüm kullanıcı ve paydaşlara karşı eşit davranacak herhangi imtiyazlı bir grup meydana getirmeyecektir.
4. Yapay zekâ sistemleri, karar alma süreçleri açıklık taşıyacak ve sistemler kararlı, güvenli sonuçlar üretecektir.

2. KURAMSAL ÇERÇEVE VE İLGİLİ ARAŞTIRMALAR

2.1. Etik ve Ahlak

Etik kavramını anlamak, insanın bir toplum içerisindeki tüm davranışlarını kavrayabilmek kadar zordur. Toplumsal yaşam içerisindeki değişim ve gelişimin bir sonucu olarak, etik kavramı için tek ve sabit bir tanım yoktur ve kavram toplumsal düzenin gelişimi ile şekillenir. Etik sistemin bütününe oluşturan prensipler, içinde bulunulan döneme has ilkelerden, sosyal kurallardan ve yaşam biçimlerinden etkilenir (Örselli, 2010). Toplum gerektiğinde yeni kurallar koyarak etik sistemini geliştirirken, dönemin gereklerine uymayan etik kurallar da kaldırılır.

Ahlak kavramının kökeni ise Aristoteles'e kadar uzanır ve sonraki filozoflar tarafından da yaygın bir şekilde kullanılır. Ahlak kavramı, etik ilkelerin tespiti için önemli bir yer tutar. Ahlak kavramının toplum içinde açık veya dolaylı olarak yerleşmesi ve kullanılması, etik ilkeler ve ahlak arasındaki ilişkinin başlangıcını oluşturur. Başka bir deyişle, etik, toplum içinde bireylerin birbirleriyle iletişim kurarken ahlaki davranışlarda bulunmaları ve karşılıklarına saygı duymaları için bir arayıştır. Etik, tarafsız olamayan ama ahlaklı kişilerin birbirleriyle iletişim kurma ve anlaşma yolu olarak görülebilir (Capurro, 2008:21).

2.1.1. Etik sistem ve etik davranış

Temelleri yaklaşık olarak beş bin yıl geriye dayanan ve tarihte ilk uygulamalarına Sümerlerin yazdığı “İştar'ın Davranış Kanununda” rastlanan etik, ilk ortaya çıktığında toplum içerisindeki ticari ilişkileri düzenlemekle başlamış, devamında insanların rahat ve refahı için uyulması gereken kurallar bütününe kadar geliştirilmiştir. Çağımızdaki etik ilkelerin temeli olarak da adlandırabileceğimiz bu ilk kurallar; kişiler arası borç ilişkilerini, borç alma-verme kurallarını, borcun ödenmemesi durumunda yapılması gerekenleri ve borcunu ödemeyenlerin katlanması gereken cezaları düzenlemiştir (Singer, 2011). Milattan önce iki bin yedi yüz yıllarında ise farklı coğrafi bir bölgede bulunan Mısır'da; yine etik kuralların varlığı tespit edilmiş ama kuralların daha çok ölümden sonra iyi bir yaşam için hayattayken yapılması gerekenler olduğu görülmüştür.

Hâlihazırda kullandığımız etik sistemin oluşumunda, tarihi toplumların etik düzenlerinin etkilerini gösteren farklı örnekler görmek de mümkündür (Polanyi, 1944). Günümüzde kullandığımıza en yakın ve kapsamlı; toplumsal düzeni sağlamaya yönelik, devlet işleyişine yön veren, ekonomik ve günlük olayları düzenleyen ilk kurallar ise üç yüz yıllarında Roma’da görülmüştür. Bu dönemde; çağımızdaki sistemin temeli sayılabilecek ilkeler, Roma vatandaşlarının günlük hayatlarını düzenlemek adına benimsenmiş ve kullanılmıştır (Tonybee, 1946). Tarihsel olarak etik sistemin toplum üzerindeki işlevine bakıldığında; kişilerin uyum içerisinde bir arada yaşamını sürdürmesine olanak sağlayacak şekilde düzenlendiği ve refah, adalet, kişilik hakları, güvenli bir ortam oluşturma, ticari ve kişisel alanda çıkarların korunması gibi ilkeler barındırdığı görülür (Rodrigues, 2009). Etik kuralların zaman ve coğrafi alana göre farklılaşmaları, Aristoteles’ten bu yana üzerinde düşünülmesine ve yorumlanmasına karşın, etik kavramının tüm toplumlar için aynı ortak anlamı ifade edecek şekilde kavramsallaştırılamamasına neden olmuştur.

Toplumsal alışkanlıkların dönemsel farklılığına bağlı olarak tarih içerisinde güzellik ve doğruluğun, kötülük ve çirkinliğin tanımları dahi değişiklik göstermiştir. Bunun getirisi olarak temel anlamı aynı kalsa da zaman zaman etik prensiplerin tanımlarında birtakım değişiklikler meydana gelmiş ve kavram karmaşık kabul edilen konular arasında görülmüştür (Santiago, 2019). Etik kavramı için farklı bakış açıları ve kullanım alanlarına göre birbirini tamamlayan ve destekleyen çeşitli tanımları mevcuttur. Etik kimi zaman başkalarına dürüstçe davranmak, kişisel hırslardan arınmak ve iyi karakterli olmak olarak değerlendirilirken bazı durumlarda sonuçları önemseyen, davranışın sonucuna bakan ve düşünceyi iyi, kötü, çirkin diye niteleyen hali ile de kullanılabilir.

Tüm bu etik türlerinin yanı sıra günümüzde çalışanlar için mesleki etik, akademik kurumlardaki personel için akademik etik, öğretmenleri ve öğretmenlik mesleğini yöneten kurallar bütünü olarak öğretmen etiği, bilgi çalışanları için bilgi yönetimi etiği, devlet çalışanları için kamu hizmet etiği, doktorlar ve hemşireler için sağlık hizmetleri etiği, çalışma ve anket verilerinin kullanımına ilişkin veri ve pazar araştırmaları etiği, bilgi teknolojileri uzmanları için sistem yönetimi etiği, gazeteciler için gazetecilik etiği gibi tanımlamaları bulunur. İlgili mesleğe ve alana dair kuralların bir arada sunulması temel olarak meslek etiği olarak adlandırılabilir (Arslan ve Berkman, 2009).

2.1.2. Etik sistem ve toplum

Ticari ilişkileri düzenlemekle başlayan kurallar; çağımızda insan hakları, haksızlıklarla mücadele, sağlık, güvenlik ve çevre gibi konulara ve en önemlisi toplumun ahenk içinde yaşamasına odaklanmıştır. Bu haliyle etik; yazılı ve yazısız hukuk kurallarını içinde barındıran, hem teorik hem de uygulamalı bir ilkeler bütünü olarak karşımıza çıkmaktadır. Başka bir deyişle etik, kararlı bir şekilde toplum içindeki bireylerin yaşamını düzenleyen ilkeleri ortaya koymaya yarayan kurallardan oluşmaktadır.

Etik'in asıl amacı, toplumsal davranışlara sebep olan temel olguları araştırmak, toplumun bu şekilde hareket etmesine neden olan değerleri araştırmaktır (Ülgen ve Mirze, 2004). Toplumsal hayatta düzen sağlamak amacıyla kullanılabilen, toplumsal yaşamı düzenleme ve yönetme biçim olan etik, kişilerin kurallara uymasıyla sosyal düzen ve antlaşmaya katkı sağlamaktadır. Kurallara uyulması; toplumsal bütünleşme, refah artışı, suç oranlarında azalmaya katkı sağlarken, etik olmayan davranışlar toplum içerisinde çözülme ve istenmeyen olayların artışına neden olmaktadır.

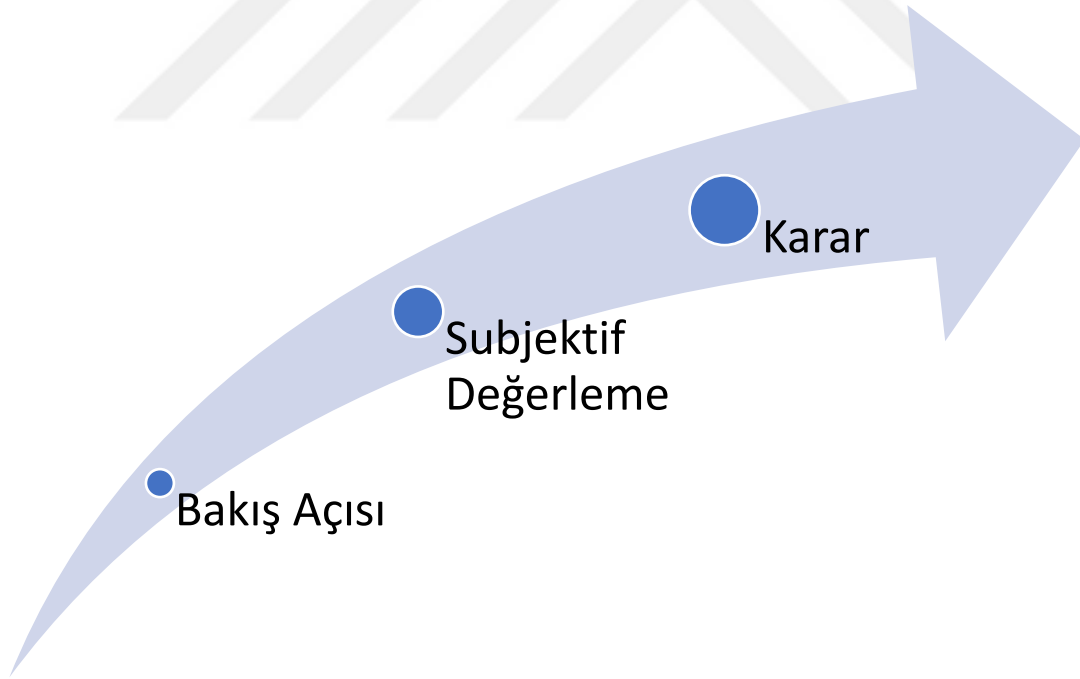
Toplum içerisinde yaşayan bireylerin özellikle ailesi ve kendi yakın çevresinin koyduğu kurallara uyma ve bunları kabullenme hızının yüksek olduğu görülür (Velasquez, 2017:139). Toplumsal yaşamı düzenlemek için koyulan kurallar, uyum sağlama sürecinden sonra bireyler tarafından otorite olarak görülür ve tersi durumda artık uyulması gerekmeyen kurallar terk edilir. Etik kurallar, şartlara uyan ve uymayan kuralların değerlendirilmesi ile devamlı olarak güncellenir. Uygulamaya gerek görülmeyen etik kurallar terk edilir, gerekiyorsa yerine yeni etik ilkeler belirlenir. Çevresel şartların günümüzdeki hızlı değişimi ile beraber teknolojik gelişmeler sonucu meydana çıkan ürünler toplumsal yaşayışı etkilemiş bunun devamında da etik kurallarda birtakım değişikliklere neden olmuştur.

Tarihsel süreç içerisinde ateşin icadından, cep telefonuna kadar her yenilik; aslında yeni kural koyma ihtiyacını da beraberinde getirmiştir. Tüm bu kuralların koyulmasının temelinde kişisel faydadan uzaklaşılması ve genelin refah düzeyinin ve faydasının artırılmak istenmesi yatmaktadır (Winner, 1980). Etik, ahlaki normların toplumsal düzende ne şekilde kullanılacağına veya kullanılmaya devam edilip edilmeyeceğinin cevabını ortaya koyar. Bu ahlaki normların toplumdan topluma farklı uygulanması ise genel bir etik ilkeler bütününden, birbirinden farklı alt sistemlerin ortaya çıkmasına neden olur. Temel olarak

kavramın tanımını yapmak mümkünken, uygulandığı topluma veya duruma göre şekillenmiş, temel tanıma uyan ama birbirinden farklı sistemler meydana getiren tanımlar türetmek mümkün hale gelmektedir.

Lawati, Sayed ve Caldwell (2019)'a göre; sadece iş etiği ile ilgili en az 16 farklı yaklaşım ve farklı etik sistem bulunmaktadır. Farklı etik alt sistemlerin nedeni esasta insanların etik konusundaki farklı algılamaları, bakış açıları ve subjektif değerlendirmeleridir. Bireyler, diğerlerinin davranışlarını ve etik davranıp davranmadığını kendi kriterlerine göre değerlendirmektedir. Bu kriterlerin kişiler arasında farklılaşması, etik kuralların farklı yorumlanmasına neden olmaktadır. Etik algının oluşum süreci bu subjektif karar ile neticelenir ve sonucunda birbirine benzeyen fakat farklı sistemleri meydana getirir.

Etik ortamın yapısına bakıldığında kişilerin bir diğerinin eylemlerini farklı bakış açılarıyla değerlendirdiği ve yorumladığı subjektif bir çerçeve olduğu görülür (Kızıl, 2021).



Şekil 2.1. Etik Algı Oluşum Süreci (Caldwell ve Clapham, 2003'den uyarlanmıştır)

Örnek olarak yerleşik olmayan gezici toplumlarda özel mülke saygı göstermek daha az önemsenirken, karşısındakine hediye vermek daha erdemli bir davranış olarak kabul edilebilir. Tam tersi durumda, gezici olmayan ve tarım ile geçinen insanların oluşturduğu

topluluklarda, yoğun fedakârlık ve emek ile üretilen ürünler, zor zamanlarda kullanılmak üzere saklanır. Bu toplumlarda özel mülkiyete izinsiz girmek, birinin malını rızası dışında almak veya kullanmak çok daha dikkat edilmesi gereken bir kural olarak karşımıza çıkar (Callicott, 2014:56).

2.2. Zekâdan Yapay Zekâya

Zekâ; insanların hayatlarını devam ettirebilmek için doğuştan var olduğu kabul edilen, çevresel değişimlere ayak uydurabilme yeteneği olarak adlandırılmaktadır (Taşdemir, 2020). Genetik yatkınlık, beyin ve vücudun ihtiyacı olan besinlerin karşılanma düzeyi, fiziksel ve ruhsal duygu durumu, yaş ve diğer çevresel faktörlerden etkilenmektedir. Farklı disiplinlerde sıklıkla kullanılan zekâ kavramı için ortak bir tanıma ulaşmak veya mutlak bir doğru iddia etmek son derece güçtür. Sadece insanlara özgü bir yetenek olduğu düşünülen zekâ, bilgi sistemlerinde meydana gelen değişimlerle salt insanda var olmadığı, yapay bir şekilde de modelleneceği kabul edilmiştir (Cerka, Jurgita ve Gintarė, 2015). Literatürde kullanıldığı alana dair önemli hususlar dikkate alınarak birbirinden farklı tanımlamalar bulunmaktadır. Biyoloji biliminde zekâ, beyindeki sinir bağlarının sayısı olarak ifade edilir.

Sinir bağlarının sayısı ve dolayısıyla zekânın, yeni öğrenmeler ile artırılabilmesi veya kullanılmadığında azalacağı düşünülür. Süreç yoluyla varlıkların daha yetenekli, bilgili, zeki ve yaratıcı olabilecekleri düşünülür. (Şimşek, 2007). Sosyal bilimlerde çevreyle uyum yeteneği, toplum bilimlerinde zekâ bir kapasite iken eğitim bilimlerinde canlılara özgü ayrı bir yetenektir. İnsanların doğumdan itibaren zaman içinde geliştirildikleri düşünülen zekâ, insanların gelişim şekline ve durumuna göre insan yeteneklerinde farklılaşmaya da sebep olmaktadır (Schmidt, 2012). Zekâ sahibi olarak eylemlerden doğan sonuçların sorumluluğunu almak ise temelde bilinç ile ilgilidir. Bilinçsiz varlıkların yaptıkları davranışlar ne kadar şaşırtıcı olursa olsun, bir eylemin sonuçlarının tartışılabilmesi ve önem arz etmesi için bilinçli varlıklar tarafından ve istemli şekilde yapılması gerekir.

Korienek ve Uzgalis (2002)'e göre canlılara ait bir özellik olarak görülen bilinç ve zekânın, insan yapımı makinelerle aktarılması yapay zekâ kavramının temel açıklamasını oluşturur. İnsandan farklı temel ilkeler ile hareket eden sistemlere, zekâ ve bilinç olgularının teknolojik imkânlarla aktarılması, makinelerle öğrenme kabiliyeti kazandırılması ve dolayısıyla zekâyı sahip otonom sistemlerin ortaya çıkması yapay zekâ süreci olarak

adlandırılmaktadır (Ereken, 2021). Günümüzde yapay zekâ sıklıkla farklı teknolojilerle bütünleşmiş olarak ve çok sayıda bilim ile bağlantılı bir kavram olarak kullanılır. Yapay zekâyı diğer bilimlerle ilişkilendiren sayısız araştırmanın yanı sıra, özellikle nanorobotlardan sinirbilimine ve beynin çalışma yöntemlerinin keşfine kadar bir süreçte yapay zekâ kavramının günümüzdeki önemi görülmektedir.

Yapay zekâyı anlamaya giden yolda, zekâ kavramını daha iyi ortaya koymak için insan zekâsının merkezi olan beyin üzerine kapsamlı bilimsel araştırmaların ilerlemeye yardımcı olacağı açıktır (Maalej ve Ayachi, 2021). Yine nöronlar üzerine yapılan bilimsel araştırma ve çıkarımların; taklit eden yapay zekâdan "bizim gibi düşünen ve öğrenen " sistemlere dönüşümdeki önemi büyük olacaktır (Lake, Ullman, Tenenbaum ve Gershman, 2017). Son gelinen aşamada, sinir bilimi, yapay sinir ağları ve makine öğrenmesinin birbirinden nasıl fayda sağlayabileceği ve birbirlerini ne şekilde destekleyebileceği ile ilgili önemli çalışmalar bulunmaktadır.

Brooks (2002:165)'e göre insan zekâsı manevi duyguların dâhil olduğu, algıların zihinsel ve fiziksel yeteneklere dönüştüğü bir alan iken; insandan farklı ilkeler ile işlem yapabilen makineleri, otonom yapay zekâyâ sahip sistemlere çevirme sürecinde sorulması gereken en önemli sorulardan biri, yapay zekâlı otonom sistemlerden bizim gibi davranmalarını mı istemeliyiz? Yoksa bunu istemek bir anlamda akıllı sistemlerin insanı taklit etme oyununa mı dönüşür? İnsan beyni ile insanın dahi yapamayacağı işlemleri saniyenin binde birinde tamamlayabilen akıllı sistemler modellemek mümkündür. Makine öğrenmesi, veri madenciliği, nesnelerin interneti gibi bu alana veri sağlayan teknolojiler, insan yaşamının hemen hemen diğer her alanında kısaca "akıllı bilgi sistemleri" adı altında kullanılmaktadır (Stahl ve Wright, 2018). Gelişen teknoloji ile hayatımızın her alanında kullanılan ve küçülen akıllı bilgi sistemlerinin ileriki yıllarda çok daha az kaynak kullanılarak üretililebilecek olması ve ürettiği verinin her yıl üstel olarak artması beklenen bir durumdur.

2.2.1. Kavram olarak yapay zekâ ve tarihsel gelişimi

Terim olarak "Yapay Zekâ" ilk olarak McCarthy (1955) tarafından kullanılmasına rağmen bilim alanının doğuşu 1956 yılındaki Dartmouth konferansına atfedilir (Crevier, 1993; McCorduck, 2004). Yapay zekâ bilimi; insan zekâsını anlamak için bilgisayarlar, akıllı sistemler ve programlar yapma bilimidir. Yapay zekânın diğer işletim sistemleri

tasarlamaktan farkı ise, modellenen sistemlerin veri girişinde sistemin kendini sınırlandırmasına gerek olmamasıdır (McCarthy, 2007). Diğer sistemlerde belirli algoritmalar ile süreç işletilirken, yapay zekâda otonom olarak toplanan bilgilerin veri olarak kullanılması mümkündür.

Yapay zekâ, insana zekâsına özgü bazı yönleri gerçekleştirmeyi hedefleyen bir bilim dalıdır. Bu alanda araştırma yapan bilim insanları; insan kavrama ve anlama becerilerinin, sorun çözme tekniklerinin ve bilişsel süreçlerinin makinelere kazandırılmasını amaçlar. Bellman (1978) yapay zekâyı insan düşünme becerilerinin otonom hale getirilmesi olarak tanımlamıştır. Haugeland (1985), bilgisayarların tamamen düşünme yeteneğine sahip makinelere dönüştürülmesi için çalışmalar yürütmüştür. Kurzweil (1990), insanlar tarafından zekâ ile yapılan işlemleri yapabilen makineler yaratmanın bir sanat olduğunu belirtmiştir. Luger ve Stubblefield (1993), yapay zekânın zekâ göstergesi davranışları inceleyen bilim olduğunu ifade etmiştir. Nilsson (1998), insanlara özgü bilişsel işlemleri yapan ve hatta araba kullanabilen sistemler oluşturmayı amaçlayan bir yaklaşımı savunmuştur. Yapay zekâ çalışmalarının tarihsel gelişiminde tanımlar Çizelge 2.1’de özetlenmiştir.

Çizelge 2.1. Yapay Zekâ Tarihsel Gelişiminde Tanımlar

Bellman (1978)	Karar verme, problem çözme ve öğrenme gibi insan düşünme becerilerinin otonom hale getirilmesi.
Haugeland (1985)	Bilgisayarları tamamen düşünme yeteneğine sahip makinelere dönüştürmek
Kurzweil (1990)	İnsanlar tarafından zekâ ile yapılan işlemleri yapabilen makineler yaratma sanatı
Luger and Stubblefield (1993)	Zekâ göstergesi davranışları inceleyen bilgisayar bilim alanı
Nilsson (1998)	İnsanlara özgü bilişsel işlemleri yapabilen ve hatta araba kullanabilen sistemler oluşturma

Güncel anlamıyla yapay zekâyı ulaşmadan önce “anlamsal ağ, karar destek sistemleri, uzman sistemler” gibi kavramlar da sıklıkla kullanılmıştır. Berners-Lee (1998) anlamsal ağın

makinelerin belgeleri anlaması olduğunu ve yapay zekâ ile bu anlamda farklılaştığını savunmuştur. Anlamsal ağ olarak adlandırılan bu alanda; makineler, insanların tanımladığı sorunları, eksiksiz girilmiş verileri kullanarak çözmektedir. Kısaca anlamsal ağda insanların çaba sarf etmesi gerekirken, yapay zekâda aynı şey makinelerden beklenmektedir. Yapay zekâ sistemleri problemlerin üstesinden gelmek adına insan müdahalesi olmadan otonom olarak karar verebilmektedir (Bartfield ve Datta, 2018). Yapay zekâyâ ulaşmadan önce modellenen diğer sistemlerden olan karar destek sistemleri ve uzman sistemler ise özellikle bilgi toplamak, önerilerde bulunmak, soruları cevaplamak ve karar süreçlerine yardımcı olmak için kullanılmıştır. Bu sistemler; belirli bir alanda otomatik olarak toplanan verileri kullanarak, analiz yapar ve bu yolla sorunlara çözüm yeteneği elde eder. Bu sistemlerde kararlar ve sonuçlar, istatistiksel hesaplamalara dayanmaktadır.

Neden sonuç ilişkileri olasılık hesaplamaları üzerinden yapılır ve yöneticilere verilecek karar konusunda destek sağlanır. Kısaca problem çözmek için bir uzman bilgisi üreten ve yeniden değerlendiren bir dizi kurallar üzerine kuruludur (Sucu, 2020). Makine öğrenmesi ise 21. yüzyılda oldukça gelişen; yapay zekânın daha dar bir alanıdır. İlk olarak Samuel (1959) tarafından kullanılmıştır. Bir sistemin açık programlama yerine, sürece sokulan veriler ile öğrenmesini sağlama yöntemidir. Model eğitilmeyi müteakip, bu yeni verileri gerçek ve tam zamanlı olarak öğrenmeye devam etme amacıyla kullanılabilir.

Bu noktada makinelerden verileri analiz etme, kategorilere ayırma, sınıflama ve etiketleme konusunda faydalanılmaktadır. Açıkça programlama yapılmayan bir alanda sistemin kendi kendine öğrenmesine imkân veren bir alt bilimdir. Makine öğrenmesi sayesinde, sistem önceden karşılaşılmamış problemlere birtakım algoritmalar kullanılarak çözümler sunabilmektedir (Mockenhaupt, 2021). Yukarıda anlatılan bu sistemler; insan emeğinden tasarruf sağlaması, bunu yaparken daha fazla veri işleyebilme kapasite ve hızına sahip olması ve objektif sonuçlar sunması nedeniyle toplum yaşamının birçok alanını etkilemiştir. Bilginin yönetilmesi ve isimlendirilmesi, belgelerin dijital bir ortamda sınıflandırılması, tanımlanması ve indekslenmesi, bilgilerin analizi için teknik destek sağlamasının yanı sıra; toplum için bilgi toplamadaki verimliliği ile daha çok, daha alakalı ve daha faydalı bilgi sağlama konusunda avantaj sağlamıştır.

Günümüzde yapay zekâyâ sahip sistemlerin bir araç olmaktan çok, bir zorunluluk haline dönüşmesi ve zekâ yarışında organizasyonların insanlara yapay zekâ sistemleri ile tecrübe

kazandırmak, ürünlerin özelliklerini istenilen şekilde optimize etmek veya otonom sistemler vasıtasıyla topluma yeni bakış ve perspektif kazandırmak amacıyla kullanacakları düşünülmektedir (Alexandre, 2017). Yapay zekâ, işlenmesi gereken veriler arttıkça bilgi gereksinimlerinin karşılanması için kullanılabilecek faydalı bir araç olarak kullanılmaktadır.

Yapay zekânın ortaya atılışının ilk yıllarında, insan zekâsını taklit etmenin karmaşıklığından kaynaklanan sorunları aşmak için çaba sarf edilmiştir. Yine bu dönemde ABD'deki akademik araştırmalarda yapılan kesintiler ve İngiltere'de makine pazarının çöküşü, alana olan ilginin azalmasına neden olmuştur. Sonraki yıllarda yatırımın artması ve teknolojik büyüme, yapay zekâ gelişim sürecine olumlu katkı sağlamış, dünya satranç şampiyonunu yenen yapay zekâ bilgisayar ve umut vadeden otonom araçlar; daha da iyi sonuçlar almak ve daha iddialı hedeflere ulaşmak için toplumu cesaretlendirmiştir (Russell ve Norvig, 2009). Özellikle organizasyonel yönetim ve bilgi toplama alanlarında, yapay zekâ yaygın olarak kullanılmıştır. Bu kapsamda; dünya çapında internetin yaygın kullanımı, üretilen verilerin üssel şekilde büyümesi ve depolanabilmesine, bilginin sınırsız hale gelmesine neden olmuştur (Silva, Xavier ve Nathansohn, 2018). Yapay zekâ, bir bilgisayarın veya bilgisayar kontrollü bir robotun, genellikle akıllı varlıklarla ilişkili görevleri insan zekâsının beş temel yönü olan öğrenme, akıl yürütme, problem çözme, algı ve dili kullanarak yerine getirme yeteneğidir. Terim sıklıkla akıl yürütme, anlam keşfetme, genelleme yapma veya geçmiş deneyimlerden öğrenme gibi insanlara özgü entelektüel süreçlere sahip sistemler için kullanılır. Yapay zekâyâ sahip otonom sistemler; görüntüleri algılayabilir, sesleri tanıyabilir ve bunları makine algoritmalarına çevirebilir.

Günümüzde yapay zekânın; kendi işletim sisteminde değişik yapma, yazılımsal geri bildirimlerde bulunma, kendi sistemini iyileştirme yeteneklerini kazanması tartışılmaktadır (Kurzweil, 2006). Yapay zekâ kavramı kısaca; insanların basit bir şekilde işlem yapamayacağı kadar büyük verilerin ve bilgi kaynaklarının analiz edilerek harekete dönüştürülmesi ve gerekli noktalarda yapay zekâ işlemcilerinin kendi yazılımını değiştirme yeteneği olarak tanımlanabilir. Yakın gelecekte bu yeteneklerin zekâ ve biliş anlamında büyük bir ilerleme meydana getirmesi ve süper yapay zekâyâ evrilmesi beklenmektedir (Bostrom, 2016).

2.2.2. Yapay zekânın kullanım alanları

Yapay zekânın toplumsal hayatta daha sık yer bulması genel tahminlerin aksine, yok olacağı düşünülen iş kollarını tamamen ortadan kaldırmamış, aksine değişiklik gösterip devamlı bir hal almasını sağlamıştır. Bazı durumlarda ise, kaynak ve ekonomik eşitsizlik yaratarak da olsa, yeni birtakım iş kollarının yaratılmasına sebep olmuştur (AI Now Report, 2017). Yapay zekâyâ iletişim alanında sıklıkla başvurulmaktadır. Teknolojinin insan iletişiminde yoğun şekilde kullanılması sonucu iş hayatında gerekli yer ve zamanlarda çokça kullanılan e-postalar, yapay zekâ kullanılarak süzgeçten geçirilmekte, geliştirilen algoritmalar sayesinde içerikler kontrol edilerek tasniflenmekte ve kullanıcılara sadece gerekli olduğu düşünülenler ulaştırılmaktadır (Kaplan A. ve Kaplan N., 2020). Yapay zekâ hava ve kara ulaştırma faaliyetlerini düzenlemede de kullanılmaktadır. Pilot adaylarının yetiştirilmesi maksadıyla kullanılan ve gerçek uçuş deneyimine yakın sonuçlar elde edilebilen simülasyonların yanı sıra, seyrüsefer yardımcı sistemleri sayesinde uçağın kalkışından itibaren otonom sistemler tüm kontrolleri devralabilmekte veya gelişen teknoloji yardımı ile kalkış da dâhil olmak üzere tamamen otonom sistemler kullanılabilmektedir.

Karayollarında ise trafik yoğunluğunun ölçülmesi, denetlenmesi, hız limitlerinin belirlenmesi, uygunsuz durumların tespiti, plaka tanıma sistemleri, şerit denetim mekanizmaları, trafik ışıklarının süresinin ayarlanması gibi etkinlikler vasıtasıyla yapay zekâ, insan hayatını düzenlemektedir (Erdal, 2018). Bakım ve onarım faaliyetlerinde de yapay zekâdan faydalanılmaktadır. Yüksek maliyete neden olan bakım onarım giderleri ile etkin olarak başa çıkabilmek adına; otomatik olarak hesaplanan iş saatleri, parça değişim süreleri vb. faktörler yapay zekâ temelli sistemler ile belirlenmekte ve bu sayede üretkenlik artışı sağlanırken, maliyetler düşürülmektedir (Smith ve Johnson, 2018). İş hayatında faaliyetlerin koordine edilmesinden, toplantı planlamasına, gelecek dönem kar zarar tahminlerinin modellenmesinden hisse senedi fiyatlarının olası yönünün belirlenmesine kadar ki süreçte yapay zekâyâ sahip sistemlerden faydalanılmaktadır. Hayatımıza giren internet bankacılığı işlemlerinde dahi finansal durumun belirlenmesi, kredi notunun hesaplanması gibi işlemler yapay zekâ ile yapılabilmektedir.

Kişiyeye uygun bankacılık ürününün seçimi ve teklif edilmesi sonucu, bankalar tarafından ürün satışları çok daha zahmetsiz ve kişiyeye özel olarak yapılabilmektedir (Calayır ve Kabak, 2021). Yapay zekânın kullanıldığı bir diğer alan elektronik ticaret siteleridir. Kullanıcıların

demografik bilgilerine göre sunulan film önerileri, ürün renk ve modelleri, daha önce satın alınan ürünlere göre ürün teklifleri yapay zekânın elektronik ticarete sıklıkla kullanıldığı durumlardır. Son dönemde otonom işlem yapan iş makinelerinden, daha hassas makinelere kadar geniş bir alanda önemli gelişmeler kaydedilmiştir. Günümüzde tamamen otonom şekilde tıbbi bir işlem gerçekleştirilemese de; hastalık tanılama, kanserli hücrelerin ayrıştırılması gibi karar destek sistemlerinde yapay zekâdan önemli ölçüde faydalanılmaktadır.

Yakın zamanda otonom sistemlerin insan kontrolünde ve gözetiminde ameliyat yapma olasılığı bulunmaktadır (Güven ve Ayvaz Güven, 2023). Taşıma ve nakliye alanlarında da yapay zekâ kullanılmaktadır. Robotik alanında kullanılan yapay zekâyâ sahip sistemler yörünge kontrolü yapabilmekte, forklift robotları otonom olarak çalışabilmekte ve rota belirlemeler akıllı bilgi sistemleri ile icra edilebilmektedir (Güney ve Özkaya, 2020). Evin krokisini hafızasına kaydederek temizlik yapan ev robotlarından, gökyüzünde belirli rotalarda istenilen video kaydı, fotoğraf çekimi gibi işlem yapan dronlara kadar ki süreçte yapay zekâ kullanımı yoğun olarak görülmektedir. Daha az malzeme ve iş gücü ile üretilen bu tip ürünler günlük hayat için olmazsa olmaz hale geldikçe gelişimleri de hızlanmaktadır. Robot üreten yapay zekâlı otonom sistemlerin gelecekte mümkün hale gelmesi ile insan nesli gibi bir robot neslinin ortaya çıkabileceği günümüzde değerlendirilmektedir.

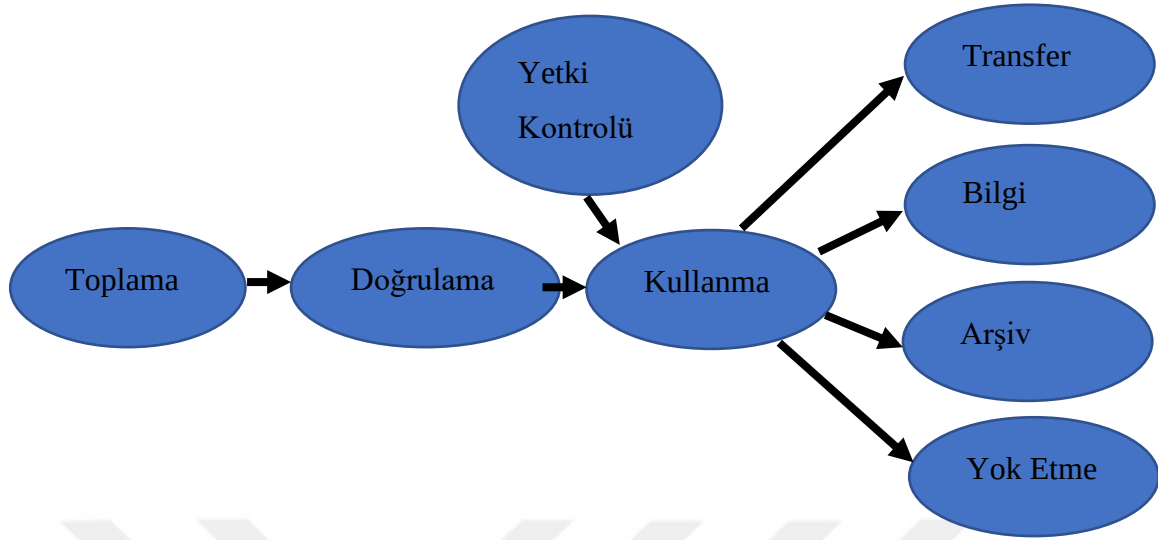
Gruplara ayırma ve hiyerarşilerin belirlenmesi insan düşünce sisteminin bir parçasıdır. Bilgi sistemleri öğrenme teknikleri ise, öğeleri kümelemek veya birbirine bağlamak için istatistiksel yaklaşımlara dayanır ve büyük veri setlerinde, ilişkileri modellemek ve keşfetmek için olasılık ve oran hesaplamaları yapar (Kurbanoğlu, 1992: 189). Bilgi sistemlerinin olasılık hesaplamalar konusunda insanlardan daha başarılı olduğu aşikârdır. Bu anlamda toplumsal alanda otonom sistemlerden, olasılık hesaplayabilmesi ve kararlara dayanak oluşturması açısından fayda sağlanmaktadır. Karar almaya yardımcı otonom sistemler vasıtasıyla verilerin işlem hızının artması, aynı zamanda işletmeler için bilgiyi daha erişilebilir hale getirmekte ve ekonomik getiri artışı sağlamaktadır (Dönerçark ve Tecim, 2020). Özellikle son yirmi yılda, işlenecek verilerin artması ve bilginin önemli oranda sayısal hale gelmesi ile otonom sistemlerin kullanımı, yığın haldeki bilginin yönetimi açısından bir zorunluluk olarak ortaya çıkmaktadır. Yapay zekâ sistemleri ile toplumun yönetilmesi, işleyişin düzenlenmesi, verimliliğin artması ve ekonomik çıkar sağlanmasının yanı sıra, doğru şekilde kullanıldığında, önemli ölçüde güvenlik kazanımı da

sağlanmaktadır. Güvenlik alanında otonom yüz tanımlama sistemleri toplumsal düzeni sağlamak için çeşitli ülkelerde kullanılmaya başlamıştır. Geldiğimiz noktada çok yoğun ve çeşitli şekillerde üretilen veriler çok yüksek hacimlere ulaşmıştır (Atalay ve Çelik, 2017).

2.2.3. Yapay zekâya veri sağlama ve verilerin kişiselliği

Neredeyse her değişimde olduğu gibi, toplumun kullandığı ve toplumsal hayatı etkileyen teknolojik gelişmeler de bünyesinde fayda ve zarar barındırmaktadır. Bu nedenle olası zararları önlemek adına, yapay zekâya sahip otonom sistemlerin modellemesinde; disiplinler arası, özverili ve dikkatli bir çalışma yapılması gerekir. Yapay zekâya sahip otonom sistemlerden istenen faydanın sağlanabilmesi için; devletlerin, uluslararası kuruluşların ve ulusal etik toplulukların ortak beklentisi; insanın zekâ ve düşünme becerilerine sahip ve soruna neden olmayacak bir sistemin ortaya koyulabilmesidir (Ma ve Liu, 2019). Son dönemde bu çabalar ilk basamakta; verilerin toplanması, algoritmaların sıkı denetlenmesi ve önyargı besleyecek ya da diğer bir taraf lehine haksız fayda sağlayacak işlemlerin engellenmesine yoğunlaşmıştır (CAHAI, 2020). Yapay zekâya sahip sistemlerin kullanacağı verilerin toplanması aşaması ise birtakım hassasiyetler barındırmaktadır. Verilere erişim sağlarken, hassas kişisel bilgilerin gizli kalması ve kişi mahremiyetinin muhafaza edilmesi gerekmektedir.

Mahremiyet konusunda en hassas alanların adli durum, mesleki bilgiler, tüzel kişilik bilgileri ve ailesel bilgiler olduğu görülmektedir. İnsanın yaratılış ve yaşayışına özgü verileri içeren; genetik ve sağlık bilgileri, ırk ve etnik köken, siyasi veya dini inançlar, cinsel tercihler gibi hususlar hakkında bilgi toplanması ve bu bilgilerin kullanılabilmesinin de belirli esaslara göre yürütülmesi gerekmektedir (Canada, 2000). Bu esasları ortaya koymak adına temel bir veri toplama algoritması ile kişisel verilerin elde edilmesinden, yok edilmesine kadar geçen kişisel verilerin yaşam süreci şekil 2.2.de gösterilmiştir.



Şekil 2.2. Kişisel Verilerin Yaşam Döngüsü (Romansky, 2015’den uyarlanmıştır.)

Hassas kişisel bilgilerin gizli kalması ve kişi mahremiyetinin muhafaza edilmesi sürecinde gerekli olan diğer bazı gereksinimler aşağıda vurgulanmaktadır (Romansky, 2015).

- Kişinin rızasıyla ve meşru bir sebebe dayalı olarak kişisel verilerin toplanması
- Önceden tanımlanmış kriterlere göre toplanan verilerin korunması
- Meşru kişiler tarafından ve kimlik doğrulama, yetkilendirme ve hesap verebilir şekilde kullanım
- Kişisel verilerin doğruluk, kapsamlılık ve güncellik açısından bütünlüğü ve içerik yönetimi
- Katı ve kurallara dayalı aktarım
- Yasaya uygun olarak arşivleme
- Mahremiyet kavramının, kültürel gruplar veya bilgi alanları içinde ülkeden ülkeye değişkenliğinin göz önüne alınması

Yukarıdaki maddelerin ortak özelliği kişisel verilerin toplanması ve kullanımında; yasalara ve diğer yazılı-yazısız kurallara uygun davranılmasını sağlamasıdır. Neredeyse her değişimde olduğu gibi, toplumun kullandığı ve toplumsal hayatı etkileyen teknolojik gelişmeler de bünyesinde fayda ve zarar barındırmaktadır. Bu nedenle olası zararları önlemek adına, yapay zekâya sahip otonom sistemlerin modellemesinde; disiplinler arası, özverili ve dikkatli bir çalışma yapılması gerekir. Bu bilgilerden hareketle kişisel verilerin

mahremiyetinin sağlanması, yapay zekâ etiğine giden yolda öncelikli konuların başında gelmektedir.



3. YAPAY ZEKÂ ETİĞİ

İlk olarak Anderson ve diğerleri (2005) tarafından kullanılan yapay zekâ etiği, insanların teknoloji ile ilişkili davranışlarıyla ilgilenen bilgisayar etiği ve akıllı makinelerin tasarımı, yapımı ve insan-makine etkileşiminde insanların etik davranışlarını ifade eden robot etikten farklı bir kavramdır. Yapay zekâ etiği; otonom sistemleri tasarlama veya kullanma prosedürlerinin, toplumsal düzen içerisinde önemli olan ahlaki kavramlar, ilkeleler, mevcut yasal ve sosyal düzenlemeler ile uyumlu hale getirilmesi ile ilgilenmektedir (Anderson ve Anderson, 2006). Yapay zekâ etiğinde, sistemlerin eylemlerinin ahlaki sonuçlarına özellikle vurgu yapılmaktadır. Fakat bu noktada, bizden çok üstün birtakım mantıksal önermeler ve sayısallaştırılmış veriler ile hesaplamalar yapabilen işletim sistemlerinin, toplumun etik düşünme yapısına uzak olduğu görülmektedir (Floridi, 2021). Bu problemin çözümü için otonom sistemlerin öncelikli olarak; kimin için, ne için ve hangi etik ilkelerle üretileceğinin belirlenmesi gerekir.

Otonom sistemlerin nihai bir ürün olarak ortaya koyulabilmesi için; yasal gereklilikleri karşılamış, etik kurallara uygun, hesap verebilir ve gelişime katkı sağlayacak şekilde tasarlanması gerekmektedir (European Commission, 2019). Bawden ve Robinson (2012)'a göre; sistemlerde kullanılacak gerekli ayrıntıların belirlenmesinde güçlük yaşanmakta ve bu güçlük aslında günlük hayatta kullandığımız faydalı bilginin sayısallaştırılması ve internet teknolojilerinin her alana yayılması ile kontrolün zorlaşmasından kaynaklanmaktadır. Bilginin insan yaşamından sayısal ortama aktarılmasında ve depolanma sürecinde aşağıdaki problemlerle karşılaşmaktadır (Bawden ve Robinson, 2012).

- Gizlilik ve Veri Koruma
- Bilgilerin Mülkiyeti ve Bilgilerin Ticari Kullanımı
- Sansür ve Entelektüel Özgürlük Arasındaki Çelişkiler
- Bilgi Eşitsizliği ve Bilgi Sahipliğindeki Uçurum
- Adil Kullanım
- Fikri Mülkiyete Saygı
- Evrensel Erişim
- Denge ve Yanlılık Sorunları

Teknolojinin yaşamımıza girmesi ile sıklıkla karşılaşılan bu problemler tek başına süreklilik arz etmeyeceği göz önüne alınmalı ve birbiri ile paralel olarak değerlendirilmelidir (Floridi, 2013).

3.1. Yapay Zekâ Modellemesinin Aktörleri

Yapay zekâya sahip otonom sistemlerin hayatımıza girmesi ile beraber süreç içerisinde temel olarak üç ayrı grup yer almaktadır. Bu gruplar; yapay zekâ geliştirenler, yapay zekâyı dağıtanlar yani hizmet sunanlar ve bu otonom sistemin kullanıcılarıdır (Avrupa Parlamentosu, 2020). Önümüzdeki yıllarda tamamen otonom sistemlerin ortaya çıkması ile kendi davranışlarından sorumlu tutulabilecek dördüncü bir aktörün tanımlanması olasıdır (Cerca ve diğerleri, 2015: 383).



Şekil 3.1. Yapay Zekâ Teknolojisinde Üç Grup (Avrupa Parlamentosu, 2020’den uyarlanmıştır.)

➤ Yapay zekâ geliştirenleri, otonom sistemi modelleyen teknik ekiplerdir. Genellikle firma, fabrika veya kuruluşları kapsamaktadır.

➤ Yapay zekâya sahip otonom sistem hizmeti sunan dağıtıcılar ise, bu sistemlerin kullanımından doğan fayda ve riskler üzerinde bilgisi olan kişilerdir. Sunulan hizmetin hassasiyetlerini, toplumsal olarak ne şekilde yarar getireceğini veya insana olan zararlarını ön görmesi gereken topluluktur.

➤ Yapay zekâ kullanıcıları ise sunulan hizmetlerden belirli maddi, manevi esaslar karşılığında yararlanma hakkına sahip kişilerdir.

Bu rollerin her durumda bu kadar ayrı olarak ifade edilmesi mümkün olmayabilir. Bazı durumlarda geliştirici aynı zamanda hizmet sunucusu olabilir. Etik kuralların uygulanması ya da uygulanmamasından doğan sorumlulukların belirlenmesinde, geliştiricinin otonom sistemin altındaki teknoloji hakkında kapsamlı bir bilince sahip olması doğal karşılanabilirken, hizmet sunucusunun daha az teknik bilgiye sahip olması ama bunun yanında toplumsal kullanımındaki problemlerden daha çok sorumlu tutulması makul karşılanabilir.

Yapay zekâ etik ikilemleri, örneğin ölümlü bir otonom araç kazasındaki riskleri yanlış hesaplayan geliştiricinin kazada ne kadar kusurlu olacağı, ya da otonom araç satan dağıtıcı firmanın ne oranda suçlamaya maruz kalacağı konularında belirsizlik yaratmaktadır. Belirsiz olan bir diğer konu ise; otonom araç kazalarında çarpışmanın önlenmesi veya olası en az zarar kavramının ortaya koyulmasında etiğe nereden bakılması gerektiği halen belirlenememiştir (Wolkenstein, 2018). Etik düzlemde maksimum faydayı sağlamak maksadıyla; otonom sisteme kapsamlı bir veri seti yükleyerek, ahlaki hesaplamaları sayısallaştırabilir ve matematiksel işlemler ile en az yaş toplamı insan sayısının kazada kaybı sağlanabilir. Fakat cevaplanması gereken asıl soru, toplumun bu yöndeki kararının gerçekte ne olduğudur (Velasquez, 2017). Algoritma hangi etik kriterlere göre karar vermelidir? Kurtarılan hayatların toplamına göre mi yoksa ilgili kişilerin “niteliğine” göre mi? Karar vermeden önce, böyle bir seçimi modellemenin insan sürücülerde mümkün olmadığını da göz önüne almak gerekir. Algoritma tarafından verilen herhangi bir karar, bu nedenle bir insan sürücününkinden daha rasyonel programlanması daha kolay ve tahmin edilebilir olacaktır. Etik olarak tanımlamanın zor olduğu bu durumda cevap toplumsal değerlere bağlı olarak değişecektir.

Örnek olay olarak literatürde sıklıkla dile getirilen durumda; hızla seyreden bir aracın karşısına çıkan başka bir araç, yolun bir tarafında çocuk, diğer tarafında yaşlı bir kadın ve

verilmesi gereken hızlı ama doğru bir karar yer almaktadır. Seçenekler arasında tartışmasız her toplumda korunması gereken yaşlı bir kadın ve bir çocuğun farklı taraflarda yer aldığı bir durum söz konusudur. Otonom araç; çarpışmadan kaçınmak için direksiyonu, yaşının zaten yeterince fazla olduğunu hesapladığı kadına mı yöneltecek, yoksa çocuğun ölmesinde bir sakınca olmadığına mı karar verecektir. Ya da bu alternatiflerden farklı olarak her iki tarafa da direksiyonu çevirmeyip içerisindeki kullanıcıyı ya da karşıdaki aracın sürücüsünü mü ölüme götürecektir. Sıklıkla dile getirilen bu problemin, bilim adamlarının ve toplumda etik konusunda çalışanların dikkatini çekmesi bir tesadüf değildir. Çünkü problem muhtemel gerçek hayat senaryolarıyla aynı çoklu alternatifleri barındırmaktadır. Etik kararlar yaşlılara saygının çok olduğu bir kültürde diğerlerinden farklıdır. Awad, Dsouza ve Kim (2018) tarafından yapılan çalışmada Fransızlar kadınların, Japonlar ise yaşlıların kurtarılması gerektiğini düşünmektedir. Çoğunluk ise; kaza sırasında uygulanacak davranış tarzının, belirli sınırlamalar içeren algoritmaya değil, şansa bırakılmasını desteklemektedir.

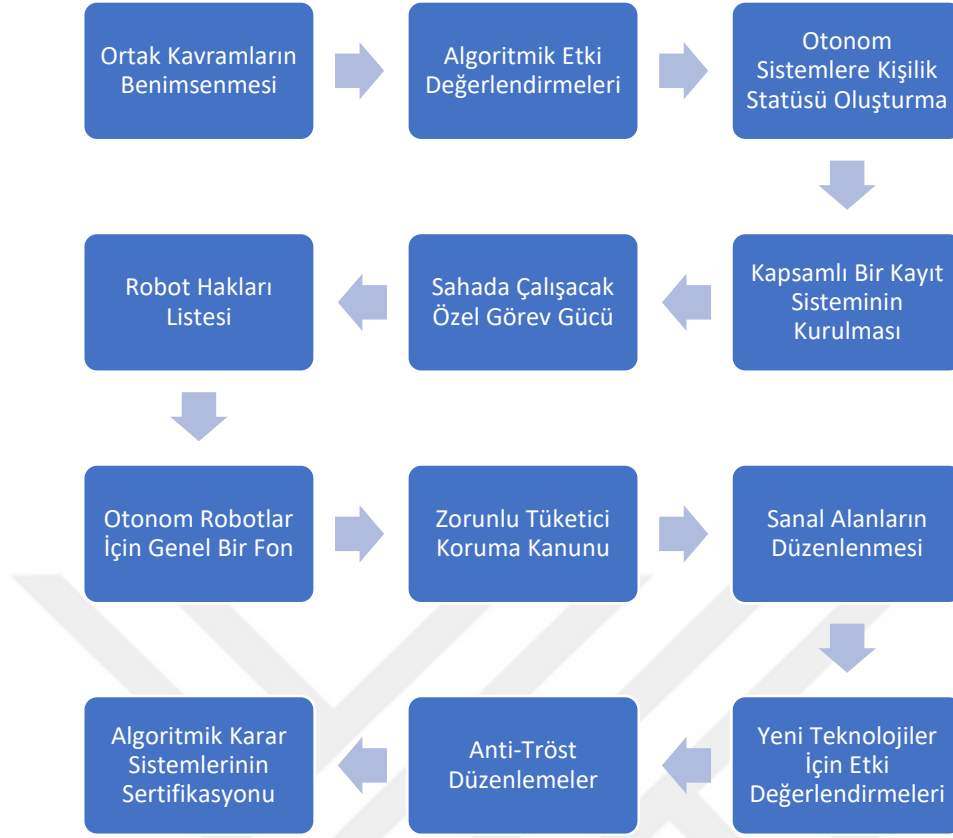
3.2. Yapay Zekâ ve Tartışmalı Etik Hususlar

Makinelerin bilgi işleme ve karar verme süreçleri insanlardan bağımsız ve farklıdır. Yapay zekâ günümüzde topladığı verileri, algoritmalar vasıtasıyla analiz ederek karar vermektedir. Otonom sistemlerde büyük miktarlarda hassas veri analiz edilmekte ve karar vermek için gerekli olacağı düşünülen bilgiler kullanılarak sonuçlar üretilmektedir. Analizin hata riskini en aza indirmek maksadıyla yine algoritmalar vasıtasıyla etki değerlendirmeleri yapılarak; yanlış karar verme riskleri azaltılır. Karar verme mekanizmaları açısından değerlendirildiğinde, birey ile otonom sistem birbirinden ayrılır. İnsanın içinde yaşamış olduğu toplumsal yapı, tarihsel bir gelişimin bir sonucudur ve kişiler için uyulması gereken davranış kalıpları belirlenmiştir. Yapay zekâyâ sahip sistemlerin karar verme mekanizmaları ise algoritmalarından oluşmaktadır. Her yeni harekette süreç yeniden işletilerek değerlendirme yapılır ve dolaylı öğrenme gerçekleştirilir (İnce, İmamoğlu, S.E. ve İmamoğlu, S.Z., 2021). Bu noktada makineler insanlar tarafından otonom hareket edebilme yeteneği verilmesine rağmen, karar alma sistemlerindeki değerlendirme kriterlerinin insanlar tarafından belirlenmesi, karar verme süreçlerindeki insanın etkisinin tamamen ortadan kalkmadığını gösterir. Makineler için yapılan bağımsız yapay zekâ vurgusuna rağmen, karar verme ve öğrenme süreçlerinin insanlar tarafından sağlanan algoritma ve modeller tarafından yapılıyor olması, makinelerin öğrenme sürecinde kendisine sunulan veri seti kadar tarafsız olabileceğini ortaya koyar.

Otonom sistemlerde belirli ölçüdeki insan katılımı ve sistemin hangi bilgiyi kullanacağını seçememesi, sürecin tarafsız olmasını önemli ölçüde etkilemektedir. Özellikle eksik veriler ve güvenilir olmayan sebep-sonuç ilişkilerinin olduğu önermelerde bazı sistemler yetersiz kalmakta ve gerekli etkinliği gösterememektedirler (Başoğlu ve Bulut, 2017: 577). Tüm bunlara karşı yapay zekâya sahip otonom sistemler vasıtasıyla; yüksek verimlilik ve üretkenlik düzeylerinden kaynaklanan ekonomik kazanımların, yeni yeteneklerin, indirgenmiş hata oranlarının ve toplumun daha iyi yaşamasına katkı sağlayacak yüksek refah vaat eden uygulamaların topluma kazandırılması mümkündür. Bu amacın gerçekleştirilebilmesi için, toplumlar ve insanlar arasında adaletli ve uygulanabilir etik kurallar oluşturulması yapay zekâ kullanımı ile mümkün hale gelebileceği ve yapay zekânın etik amaçlarla kullanılması konusundaki fikir birliği artmaktadır (Berendt, 2019).

3.3. Yapay Zekâ Etiğinin Kaynağı

Yapay zekânın toplum içerisinde neden olabileceği etik sorunlar; büyük oranda kişisel verilerin korunması ve eşitliğin ihlali gibi insan hakları sorunlarından oluşmaktadır (Rodrigues, Panagiotopoulos ve Lundgren vd. 2019). Olumsuz bir durumda etkilenecek en büyük topluluğun insan olması, insanı korumak için başvurulacak düzenlemenin insan hakları mevzuatı temelli olmasını gerekli kılmaktadır. Yeni otonom sistemlerin kullanımı sonucu meydana gelecek risklerin en aza indirilmesi ve potansiyel zararların önlenmesi için insan hakları düzenlemeleri yakından takip edilmelidir. Rodrigues ve diğerlerine (2019) göre; günümüzde tartışılmakta olan mevcut yasama çalışmaları aşağıdaki başlıklar altında toplanmaktadır.



Şekil 3.2. Yasama Çalışmaları (Rodrigues ve diğerleri, 2019'dan uyarlanmıştır.)

Yapay zekâyı toplumsal gelişim için kullanmak üzere modellerken, gerekli yapay zekâ mevzuatının oluşturulması ve küresel ölçekte tartışılıp karar verilmesi gerekir. Yapay zekâ teknolojisinin yaygınlaşmasıyla beraber otonom sistemlerin sebep olduğu zararların sorumluluğun belirlenmesinde yapılacak yanlış uygulamaların, gelecekte fayda sağlayabilecek birçok potansiyel otonom sistemin geliştirilmesini de engelleyeceği göz önüne alınmalıdır (Doshi-Velez ve Mason, 2017:2).

3.4. Yapay Zekâda Etik Sorumluluk

Teknolojik gelişmelerin hızı, toplumun algılama hızının üzerinde artış göstermiş ve bunun sonucu olarak gelişimin doğru şekilde yönetilebilmesi için otonom sistemlerin kullanımında ahlaki ve etik kuralların belirlenmesine ihtiyaç artmıştır (Beranger, 2021:53). Etik kurallar başlangıçta büyük ölçüde teknolojiyi kullanan insanların davranışlarına odaklanmıştır. Yeni dönemde artık makineler açısından da etik davranışların tartışılması ve makinelerin otonom olarak verdikleri kararların ve eylemlerin incelenmesi gerekmektedir (Müller, 2020).

Otonom sistemlerin günlük hayatta daha sık kullanım alanı bulmasıyla, özellikle kişisel verilerle ilgili işlevlerine yönelik sorular tartışılmaya başlamıştır. Cevap bulunması gereken temel nokta, zeki ve etik olması beklenen robotların eylemlerinin sorumluluğunu üstlenip üstlenemeyeceğidir (Baum 2020). Etik sorumluluğa sahip olması gereken bir makinenin, toplum içerisinde özerk olarak çalışabilmesini sağlamak için etik ilkeler belirlenmesi, bu etik ilkelerin bir sistemin karar sürecine dâhil edilmesi gerekir. Bu anlamda makinelerin otonom davranışlarıyla ilgili eksik bilgilerin tamamlanması ve hangi etik ilkelerin kullanılacağıının tam olarak belirlenmesi gerekmektedir (Anderson ve diğerleri, 2005).

3.5. Yapay Zekâda Etik Sorun Alanlarının Belirlenmesi

Verilerin korunması yapay zekâ sistemlerinde etik denetimin temel amaçlarından biridir. Bir ileri basamak olarak yapay zekâ etiği, aslında bir mahremiyet koruma mekanizması olarak da kullanılabilir. Var olan durumlar ile ilgili kural ve ilkeler bütünü ve pozitif bir bilim olan hukukun; toplumda güvenlik endişesi yaratma riski taşıyan ve etik kaygılara neden olan hususlarda gerekli alt yapıyı oluşturması gerekmektedir. Günümüzde etik problemler arasında en sık dile getirilene gizlilik ve kişisel verilerin gereken hassasiyetle saklanabilmesi ve depolanmasıdır (Buttarelli, 2018). Yapay zekânın günümüzde kullanılmasından meydana gelen etik sorunlar en genel anlamda kültürel, sözel ve coğrafi farklılıklara göre şekillenebilir. Teknolojinin hayatımıza girmesi ile beraber; büyük şirketlerin herhangi bir bedel ödmeden, kullanıcılarının yapmış olduğu işlemler yoluyla kişisel bilgilere sahip olabilmeleri, dijital verilerin mülkiyet ve yararlanma haklarının belirlenememesi gibi konular bilgi sistemlerinde etik tartışmaların odağını oluşturmaktadır (Spinello ve Tavani, 2004). Bir diğer genel etik kaygı ise; yapay zekâ ve makine öğrenmesi yardımı ile verilen kararların doğruluğunun, sisteme girilen verilerin doğruluk ve kalitesine bağımlı olmasıdır.

Toplum içerisinde hassas olarak kabul edilen bilgilerin kullanıldığı sistemlerde yanıltıcı sonuçlar veya mahremiyete neden olacak durumlardan kaçınabilmek adına gerekli tedbirlerin öncelikli olarak alınması gerekir (Canbay ve Sağıroğlu, 2020). Bazı durumlarda verilerin yanlış olma ve yanıltma potansiyeli birtakım güvenlik ihlallerinden ya da hatalardan kaynaklanırken, toplumsal dinamiklerden veya birtakım teknolojik yetersizliklerden de kaynaklanabilmektedir. Bu durumda, otonom sistemlerin vereceği kararların güvenilirliğinin ve etik kurallara uyumunun tespit edilmesi neredeyse olanaksız hale

gelmektedir (Canbay ve Demircioğlu, 2021). Bu problemin üstesinden gelmek adına, birden fazla kontrol mekanizması kurulmalı ve aynı durumu ölçen farklı veri setleri sürece dâhil edilerek hata olasılığı minimum seviyede tutulmalıdır. Hali hazırda otonom hava ve kara araçları muharebe ortamında kullanılmaya başlamıştır.

Fakat sistemin kimi düşman olarak belirleyip imha edeceği, ya da makine kararı ile insan hayatının sonlandırılıp sonlandırılmayacağı tartışılması gereken konular arasındadır (Sharkey, N ve Sharkey, A, 2010). İnsan yaşamını sonlandırmak bir yana, insan yaşamını kurtarma noktasında da etik ikilemler bulunmaktadır. Teknolojik gelişmelerin insan ömrünü uzattığı varsayımını kabul etsek dahi, robot cerrahlar tarafından yapılacak operasyonlarda sağlık için birtakım riskler ve belirsizlikler bulunmaktadır. Yapay zekâ sistemleri kullanmaya başlamadan önce; daha aşılmaz toplumsal eşitsizlikler, sosyal sorunlar, güç asimetrisi, ekonomik adaletsizlikler, bireysel özgürlüklerin azaltılması ve hareket özgürlüklerinin kısıtlanması gibi önemli tehlikelere hazırlıklı olmak ve tüm bunlar için engelleyici tedbirler geliştirmek insan türünün devamı için önemli bir görevdir (Sharkey, N ve Sharkey, A, 2010). Burada genel etik sorun aslında; ayrımcılık, kişisel bilgilerin mahremiyeti, önyargılar, verinin korunması veya bir şeyin yapılıp yapılmamasına ek olarak, insan ile teknolojik sistemlerin yer değiştirmesinin konusunda yaşanan endişelerdir. Yapay zekâ etiğinin daha iyi anlaşılması için yapay zekâ sistemlerinin neden olabileceği sorunları sıralamak ve bunların ne şekilde algılandığına da değinmek gerekir. Stahl (2021)'e göre günümüzde yapay zekâ alanında tespit edilen etik sorunlar sıralanmaktadır.



Şekil 3.3. Etik Sorun Alanları (Stahl, 2021’den uyarlanmıştır.)

Görüldüğü gibi etik problemler, birbirinden çok farklı alanlarda ve çeşitlidir. Bazı sorunların diğerlerine göre öncelikli olması da muhtemeldir. Gizliliğin korunması, veri güvenliği veya doğruluğu ile ilgili sorunların, bir bakıma etik problemin çekirdeğini teşkil ettiği söylenebilir. Bunun yanında verilerin kötü maksatlarla kullanımı veya kamusal alanda oluşabilecek güvensizlik durumları günümüzde sıklıkla yaşanan önemli sorunlardır. Her bir problem kaynağının çözümü için farklı mekanizmaların devreye girme gerekliliği ve çözüm alanlarının farklı disiplinlerde olması aslında yapay zekâda etik problemlerin ne kadar karmaşık ve çözümü zor olduğunun bir göstergesidir (Stahl, 2021).

3.6. Yapay Zekâda Etik Kontrol

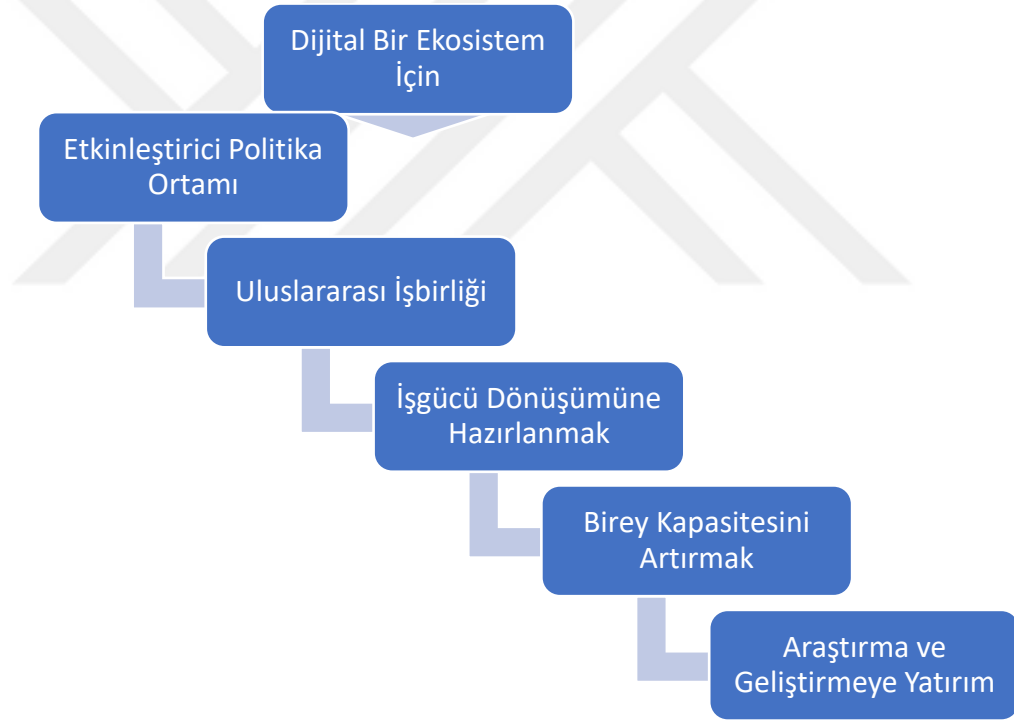
Etik kısmen de olsa, zaman içerisinde değişime uğraması muhtemel bir olgudur (Yatkin, 2009). Toplumun ahlak algısı ve etik ilkelere bakış açısı değişirken; yapay zekânın etik kuralları tespit etmesi ve uygulaması güçtür. Hâlihazırda aynı dönem içerisinde dahi; farklı coğrafi bölgelerde, farklı kişisel veri koruma kuralları, farklı toplumsal kabuller mevcuttur.

Bu da neyin etik olarak kabul edileceği konusunda tartışmalara yol açar. Diğer ülkelerden farklı olarak Çin'de, kişisel verilerin korunma kanunu “devlet organları hariç” vatandaşların bilgilerini firma ve üçüncü şahıslara karşı korur (Creemers ve Webster, 2021). Bu noktada çoğunluğa ait fikirlerin uygulamaya geçirilebilmesi için uluslararası organizasyonların tavsiyeleri etik yapay zekâ kavramının önümüzdeki yıllarda nasıl düzenleneceğini önemli oranda etkileyecektir. Toplumsal gelişim aşamalarındaki evrensel iddiaların başında; “insanın mutlu olmak için çaba gösterdiği ve buna bağlı olarak teknolojilerin haklı bir amaç uğruna kullanılması ve etik durumu desteklemesi gerektiği” varsayımı vardır (Bolay, 2013:262). Fakat tarihe bakıldığında başarılı olup olamadığımız halen belirsizliğini korumaktadır.

Gerek teknik, gerek sosyal veya etik; toplumun kabul etme hızının üzerinde olan yapay zekâ teknolojisindeki ilerlemeler, yapılacak çalışmaları fazlasıyla zorlu hale getirir. Son günlerde toplumsal yapımızda önemli değişime yol aşan COVID-19 salgını; evden çalışma, kısmi zamanlı çalışma, tokalaşmama, yakınlaşmama, sanal konserler gibi yaygın değişimlere neden olmuş ancak aynı anda işverene, yeni gözetim ve işçi haklarını sömürme biçimleri de sunmuştur (Harwell, 2020). Bu tip sosyal değişime neden olacak olayları tahmin etmenin imkânsızlığı bir yana, toplum içerisindeki etkileşimini dahi ön görmek olanak dışı bir durum olarak karşımıza çıkmaktadır. Yapay zekâ modellemesinde, etik ilkelerin toplumsal yaşamda halen var olup olmadığı ve var olanların nereye kadar sorunlara yol açabileceği etraflıca irdelenmelidir. Hızlı değişen teknolojik koşullarda, teknik gelişmelerin toplumsal düzeni sağlayan mevzuat maddelerinde bir değişikliğe yol açıp açmadığı veya mevzuat hükümlerinin halen ihtiyacı karşılama durumu göz önüne alınmalıdır.

Etik meselelerde uygulanan hukuk genel anlamda üç bölüme ayrılır. Bunlardan ilki üretilen eserin veya ürünün fikri mülkiyet hakkını koruyan hukuk, bir diğeri kişisel verilerin korunmasını sağlamakla amir olan hukuk ve sonuncusu kişiler, firmalar gibi müteşebbislerin birbiri arasındaki rekabeti düzenleyen hukuktur. Temel anlamda bu üç hukuk alanı kullanılarak insanların teknolojik gelişmeler ya da çevresel değişimler ne olursa olsun, mevzuata güven duyması sağlanır (Hukuk Çalıştayı, 2019). Güvenmediğimiz yeni bir teknolojiyi üretmek ve toplum hayatında sıklıkla kullanacağımızı varsaydığımız otonom sistemleri güvenilir hale getirmeden modellemek anlamsız bir uğraş olacaktır (Liao, 2020). Buradan hareketle, yeni otonom sistem üretme konusunda ön koşullardan biri güvenilirlik boyutudur. Bu durumu açıklamak üzere, 2019 yılında AB'de yeni bir kavram olarak ortaya

ıkan ‘‘Gvenilir Yapay Zek ’’ kullanılmaktadır. Kavramın kullanım yerlerindeki ortak nokta; insan hakları, toplumsal dzen ve etik ilkeleri gz nnde bulundurarak yapay zek nın geliřiminin saėlanması ve kullanım alanlarının bu esaslar ile dzenlenmesine olan vurgudur (European Commission, 2019). Gvenilir yapay zek ya ulařmada; kiřisel verilerin iřlenmesi ve korunması kapsamında en baėlayıcı ve kapsayıcı metin Avrupa Parlamentosunun 679/2016 sayılı Tzė olarak karřımıza ıkmaktadır. Bu tzk ile Avrupa Parlamentosu; etik kuralların yeterliliėi, alaka dzeyi, doėruluėu, uyumluluėu, rıza sreci, kiřisel veriye eriřim standartları, řeffaflık, silme ve iřlem gvenliėi ile ilgili hususlarda ayrıntılı dzenlemeler yapmıřtır. Yine bu erevede Yapay Zek  Tavsiye Komisyonu tarafından gvenilir yapay zek  oluřturmak ve dijital bir ekosistemi teřvik iin, řekil 3.2 de gsterilen beř politika konusunda hassas olunması gerekir.



řekil 3.4 Dijital Bir Ekosistemi Teřvik İin 5 Tavsiye (European Commission, 2019’dan uyarlanmıřtır.)

Yapay zek ya sahip sistemlerin toplumsal fayda saėlayabilmesi, tutarlı, adil ve hesap verilebilir olması iin; etkinleřtirici politika ortamı, uluslararası iř birliėi, iřgc dnřmne hazırlanmak, birey kapasitesini artırmak, arařtırma ve geliřtirmeye yatırım yaparak, dijital bir ekosisteme hazırlanmak gerekmektedir.

4. YÖNTEM

4.1. Modelde Kullanılacak Etik Değerlerin Belirlenmesi

Araştırmada kullanılacak etik kontrol modeli, yapay zekâyâ sahip sistemlerin etik kurallara uygun şekilde davranmasını sağlamak için kullanılan bir dizi yöntem ve tekniklerden oluşturulmuş bir algoritmayı ifade eder. Etik kontrol modeli ile otonom sistemlerde etik sorunların ortaya çıkmasını önlemeye ve otonom sistemin topluma zarar vermesine engel olmaya yönelik bir yaklaşım ortaya koyulmuştur. Modelin amacı; yapay zekâyâ sahip sistemlerin topluma ait etik değerleri göz önünde bulundurarak, bireylerin kişisel haklarını ihlal etmeden karar vermesini sağlamaktır. Bu nedenle; araştırmada kullanılan etik kontrol modelinde, yapay zekânın neden olabileceği olası riskleri en aza indirmek için, otonom sistemin ve verilerin kullanımında sorumlulukların kimde olduğu ve bunun sistem modelleme üzerinde yaratacağı etki ele alınmıştır. Kısaca model yapay zekâyâ sahip sistemlerin etik kurallara uygun olarak çalışmasını sağlamak için kontrol modeli olarak tasarlanmıştır. Otonom sistemleri denetlemek için kullanılabilecek kontrol modeline sahip olmak, yapay zekâ teknolojisinin insanlar ve toplumlar üzerindeki olası olumsuz etkilerini en aza indirmek için oldukça önemli bir basamaktır.

Modele başlamadan önce ilk karşılaşılan gerçek; otonom sistemlerin yol açabileceği etik problemlerin azaltılması maksadıyla gösterilen çabalara rağmen, henüz tam anlamıyla yapılması gerekenlerin belirlenememiş olmasıdır. Gerekliliklerin belirlenememesi beraberinde toplum içerisinde otonom sistemlerin kullanımı ile ilgili karmaşa ve tereddüte neden olmuş, bunun neticesinde toplumun faydalanabileceği muhtelif sistemlerin yaygınlaşması yavaşlamış, yer yer tamamen durmuştur (Cellan-Jones, 2018). Günümüzde artık daha çok evde olan otonom robot makinelerin, ev içerisinde bir “Truva Atına” dönüşebileceği, sistemin arka kapısı olarak adlandırılan yasal olmayan yöntemlerle ev içerisinde ses ve görüntü kayıt edebileceği düşüncesi ulusal basında yer almıştır (Postaktuel, 2021). Etik olarak yeterli mekanizmaların oluşturulamaması nedeniyle toplum nazarında, otonom sistemlerin başarısından çok başarısız olduğu düşüncesi yaygınlaşabilir. Bu da toplumun otonom sistemlerini kabulünü zorlaştırabileceği gibi, yeni teknolojinin benimsenmesini de azaltır.

En nihayetinde otonom sistemler tarafından sunulacak faydaların tamamen geri itilmesi ile insan yaşamını önemli ölçüde geliştirebilecek ve destekleyebilecek potansiyel yeni teknolojik sistemlerden de vazgeçilmiş olur (Cookson, 2018). Bu problemlerin üstesinden gelmek maksadıyla; toplum yaşayışında kullanılan soyut etik ilkelerden uygulanabilir, pratik kurallara ulaşmak önemli bir adımdır. Bu somutlaştırma işleminden sonra etik yapay zekâya ulaşmada bir sonraki basamak ise hangi kuralın ne şekilde sistemde yer alacağının belirlenmesidir.

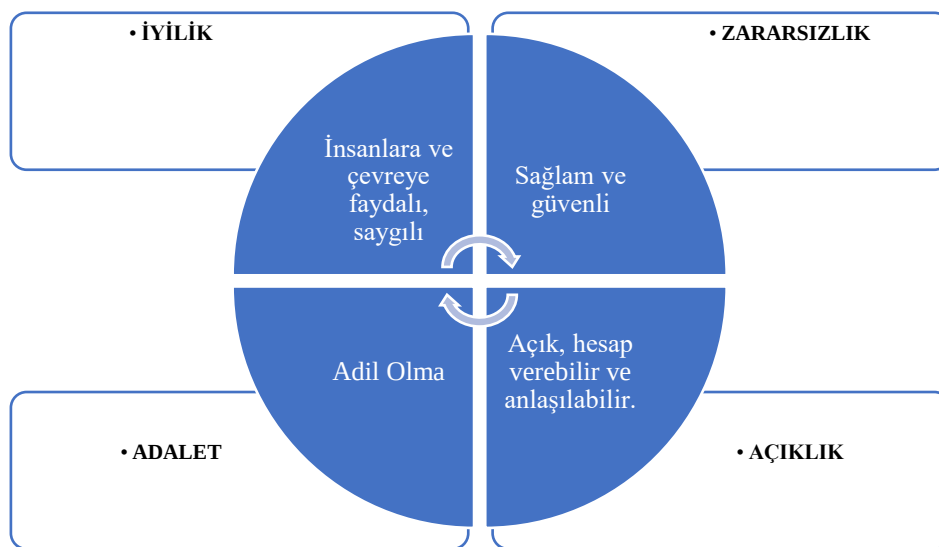
Soyut etik kavramlar ile tamamen uygulamalı olmak zorunda olan yapay zekâ modelleme basamakları arasındaki boşluk, konu ile ilgili belirli standart kuralları koymayı zor hale getirir de, uygulanması düşünülen her bir etik ilkenin bileşenlerinin iyi analiz edilmesi, alt parçalara bölünmesi, cevaplara ulaşmak için doğru soruların sorulması ve modelleme sürecinin her anında yakından denetimi ile mümkün hale getirilebilir (Floridi ve diğerleri, 2018). Endüstri firmalarının etik yapay zekâ tasarımı ve geliştirmesi için kullandıkları yöntemleri araştıran Vakkuri ve diğerleri (2019), farklı alanlarda yapay zekâ geliştiren şirketlerin etik sorunlar ile ilgili bilgi sahibi olduklarını tespit etmişlerdir. Yine aynı çalışmada bu şirketlerin potansiyel tehlikelere hâkim oldukları fakat analiz ve modelleyici olarak çalışan personele etik uygulamalar hakkında araç veya yöntem sağlamadıkları ve herhangi bir eğitim vermedikleri tespit edilmiştir.

Bu boşluğu doldurmak, endüstri profesyonellerinin etik yapay zekâ tasarımı için hangi uygulamaları, araçları veya yöntemleri kullanabileceğini tespit etmek üzere yazılan literatür belirli kilit kelimeler ile taranarak değerlendirmeye alınmıştır. Aramalar 2022 yılı Ocak ve Temmuz ayları arasında gerçekleştirilmiştir. Bulunan tüm sonuçlar teori ve uygulanabilirlik açısından değerlendirilmiştir. Aramalar esnasında kullanılan anahtar kelimeler tablo 4.1’de verilmiştir. Kilit kelimeler seçilirken yukarıda tarif edilen kapsayıcılık, yol ve yöntem sağlama, alan profesyonellerine hitap etme, etik literatüre uygunluk gibi değerlendirmeler göz önünde bulundurulmuştur. Kilit kelimeler Türkçe ve İngilizce olarak iki farklı dilde kullanılmıştır.

Çizelge 4.1. Literatür taramasında kullanılan anahtar kelimeler

ANAHTAR KELİMELEER		
Etik	Sistem Modelleme	İş/Sistem Değerleme
Bilgi Etiği	Otonom Sistem Tasarımı	Otonom Sistemler
Teknoloji Etiği	Uygulamalı Etik ve Zarar Vermeme	Teknoloji ve Şeffaflık
Büyük Veri	Uygulamalı Etik ve İyilik	Yapay Zekâ Hukuğu
Etik Teori	Uygulamalı Etik ve Adalet	Yapay Zekâ Modellemesi
Ön Yargı	Uygulamalı Etik ve Toplum	Kişisel Verilerin Korunması
Ahlak	Uygulamalı Etik ve Mahremiyet	Uygulamalı Etik ve Sorumluluk

Teknoloji ve etik ile ilgili literatürün sistematik bir incelemesi sonucunda Şekil 4.1. de görülen ve sıklıklarına göre diğerlerinden ayrılan; açıklık, adalet ve hakkaniyet, zararsız olma, fayda sağlama temalarının çalışmaların önemli kısmında öne çıkan ilkeler olduğu görülmektedir (Jobin ve diğerleri, 2019). Bunlara ek olarak sorumluluk taşıyabilme, güvenlik, özerk çalışabilme yeteneği ve kontrole açıklığın da sıklıkla tekrar ettiği ortaya koyulmuştur (Royackers ve diğerleri, 2018). Literatür incelendiğinde ve ortaya çıkan fikir birliği göz önüne alındığında, aynı kavramlar uluslararası standartlar belirlenirken istenen ve tavsiye edilen ilkeler olarak karşımıza çıkmaktadır (OECD, 2019).



Şekil 4.1 Etik Yapay Zekânın Dört Önemli İlkesi (OECD, 2019'dan uyarlanmıştır)

Etik yapay zekânın taşınması gereken asgari değerleri içeren Etik Yapay Zekânın dört önemli ilkesinin değindiği konular aşağıda verilmiştir (European Commission, 2019).

- Fayda sağlama ve iyilik: Toplum değerlerini kapsayan ve büyümeye destek olan, sürdürülebilir şekilde toplum içerisinde kalkınma ve refah artışına neden olan otonom sistemler tasarlama prensibidir.
- Zarar vermeme: Öncelikli olarak insanlara ve daha sonra diğer hiçbir canlıya zarar vermeyecek derecede güvenlik ve emniyet hususlarında yapılan testleri geçmiş sağlam ve güvenilir otonom sistemler tasarlanmasını içerir.
- Adalet: Güç dengelerini bir diğeri aleyhine değiştirmeyecek, kullanana haksız üstünlük sağlamayacak, genel toplumsal adaletin sağlanması hususunda etik ilkelerden ayrılmayacak otonom sistemlere olan ihtiyaçtan bahseder.
- Açıklık: Modelleme basamağından kullanıma kadar her safhada şeffaflık sağlayan, yapılan işlemin sebebini makul ve mantıklı bir şekilde açıklayabilen ve dolayısıyla hesap verebilir otonom sistemlerin tasarlanmasıdır.

4.2. Modelde Kullanılan Risk Yönteminin Belirlenmesi

Modelde kullanılan risk yönetim yöntemi belirlenirken literatürde bulunan çalışmalar ve bilgi kaynakları; ortaya koydukları teoriler, alana katkı sağlamak üzere sundukları yol ve yöntemler, etik yapay zekâ modellemesi için kullanılabilir olduğu değerlendirilen araçlar ve sonraki araştırmalar için verilen tavsiyeler gözden geçirilmiştir. Yapay zekâ etiği ile ilgili literatür taramasından elde edilen önemli bir sonuç, bugüne kadar yapılan araştırmaların büyük bir çoğunluğunda konunun teknik bir çözümünün var olduğunu ortaya koyması fakat teknik açıdan nasıl yapılacağı ile ilgili açıklama içermemesidir. Miller ve Coldicott (2019)'a göre; teknoloji alanında çalışan kişilerin %79'u etik hususlarda kendilerine yardımcı olacak teknik olarak açıklayıcı, kullanılabilir ve pratik kaynaklara ulaşmakta problem yaşamaktadır. Yapay zekâ modelleme söz konusu olduğunda asıl ihtiyaç, soyut ilkelere ziyade teknik açıklamalar ve model önerileri olduğu bir gerçektir.

Önerilen modelde ortak temel etik ilkelerin ortaya koyulmasından sonra, “soyut ilke ortaya koyma” probleminin çözümü için sıradaki basamak etik yapay zekânın çerçevesinin ve model basamaklarının belirlenmesidir (Maas, 2019). Toplum tarafından kabul görebilecek etik bir çerçeve ortaya koymadan, hayatımızın her alanında kullanılması için otonom

sistemlerin üretimine onay vermek olanaksızdır. Bu eksikliğin yukarıda sayılan ve literatürde sıklıkla yer bulan temel ilkeleri de içeren bir model ile giderilmesi, ortak bir algoritma belirlenerek toplumun beklentilerinin karşılanması son derece anlamlıdır.

Bu aşamada araştırma modeline ilk olarak Siber Güvenlik Çerçevesi dahil edilmiştir. Amerika Birleşik Devletleri Ulusal Standartlar ve Teknoloji Enstitüsü (NIST) tarafından siber güvenlik alanındaki; ölçüm, ağırlık, zaman referansları ile metot ve standartları içeren, siber güvenlik açıklarının analiz edilmesi, problem sahalarının yönetilmesi ve azaltılması ile sahip olunan büyük verilerin güvenliğinin geliştirilmesi için oluşturulmuş bir rehberdir. NIST Siber Güvenlik Çerçevesi temel olarak tanımlama, koruma, algılama, müdahale etme ve kurtarma adımlarını içeren beş basamaktan oluşur ve güvenilir bir araç olarak kabul edilir (FTC, 2023). Özel sektör ve diğer ülkeler tarafından da sıklıkla kullanılmaktadır. Siber Güvenlik Çerçevesi 'ne göre ilk basamak olan tanımlama aşamasında; kullanılan bilgisayar, tablet, cep telefonu ve diğer tüm veri taşıyan, aktaran ve depolayan malzemeler belirlenir. Bu ilk aşamada verinin korunması gereken fiziksel ve sanal alanlar tespit edilir. Daha sonra bir güvenlik açığına karşı verilere erişimi olan tüm roller ve sorumluluklar tespit edilerek kapsayıcı güvenlik politikası oluşturulur. Koruma aşaması ağ ve verilerin kontrol edilmesi için şifreleme ve güvenlik yazılımlarının kullanılmasını içerir. Personele güvenlik eğitimi verilmesi, güvenlik yazılımlarının güncel tutulması, veri yedeklemelerinin planlanması ile risklerin azaltılması sağlanır. Algılama basamağında yetkisiz ve olağan dışı girişlerin tespit edilmesi hususuna yoğunlaşılır. Bu aşamada fiziksel de dâhil olmak üzere tüm araçların yetkisiz erişimleri kontrol altına alınır. Dördüncü aşama “yanıtlama” aşamasıdır. Siber güvenlik açığının meydana geldiği durumlarda hareket tarzlarını içerir. Saldırı anında ve sonrasında ilgili kişilere bildirmek, gerektiğinde sistemin işleyişini durdurmak ve saldırıyı kontrol altına almak adımları uygulanmalıdır. Son basamak olan “iyileştirme” adımı ise saldırı sonrası etkilenen verilerin geri yüklenmesi ve siber güvenlik politika ve planlarının güncellenmesi hususları incelenir (FTC, 2023).

Modelde kullanılan ikinci kılavuz, Nieves, Dempsey ve Pillitteri (2017) tarafından oluşturulan "Bilgisayar Güvenliği için Bir Giriş: NIST El Kitabı" (NIST SP 800-12) adlı el kitabıdır. Yine Ulusal Standartlar ve Teknoloji Enstitüsü tarafından onaylanmıştır. Bilgi sistemleri ile ilgili temel ifadeler, saldırı çeşitleri ve bunlara karşı alınabilecek önlemleri içerir. 23 alt bölümden oluşan kılavuzda temel 5 bölüm bulunur. Tanımlama bölümünde; varlıklar, veriler, idari yetkiler, sorumluluklar tanımlanır. İkinci bölüm olan korumada ise

temel olarak erişim ve veri gizliliği esas alınmaktadır. Algılama bölümünde saldırıların tanımlanması ve analizi ele alınır. Sonraki bölümde ise verilecek yanıtlar ayrıntılı şekilde anlatılmıştır. Son bölüm sistemin kurtarılması ve iyileştirme adımlarını ayrıntılı bir şekilde ortaya koyar. Siber Güvenlik Çerçeveleri kapsamlı bir şekilde sonuca odaklı olarak çalışır. Yapay zekâya sahip sistemlerde kullanılacak olan etik kontrol modelleri de; sistemin kullanılacağı ortam, veri ve uygulamalar ile kullanıcıları kapsar. Ayrıntılı bir model; kullanıcı uç noktalarını, bulut ve diğer fiziksel depolama birimlerini, veri koruma ve kurtarma protokollerini ve ağ şemalarını içermesinin yanı sıra çeviklik ve esnekliğe imkân sağlayacak ölçüde olmalıdır (TechT, 2023). Oluşturulan modelin yukarıda sayılan hususları karşılaması ve etkinliğinin artırılması hedeflenmiştir.

4.3. Verilerin Toplanması

Etik yapay zekâ oluşturmak için ilk adım, modelleme sürecinin her aşamasında etik sorunların göz önünde bulundurulmasına yarayacak bir çerçeve oluşturmaktır. Bu maksatla dört önemli etik ilke (Fayda ağlama ve iyilik, Zarar vermeme, Adalet ve Açıklık) ile Stoneburner, Goguen ve Feringa (2002) tarafından oluşturulan bilgi teknolojilerinde “Risk Yönetim Kılavuzu” ve Nieles ve diğerleri (2017) tarafından oluşturulan Bilgi Güvenliği El Kitabı basamakları bütüncül bir şekilde ele alınmış ve tüm kontrollerin modelleme sürecine tam olarak yansıtılması amaçlanmıştır. Modelin birleştirilmesindeki nihai hedef, yapay zekâyla ilgili etik risklerin daha iyi yönetilmesini sağlamaktır. Bu sayede yapay zekâ etiği konusunda daha kapsamlı bir yaklaşım ortaya koyulması sağlanmıştır.

Bu amaçla ilk olarak etik yapay zekâ tasarımı için hangi uygulamaların, araçların veya yöntemlerin kullanabileceğini tespit etmek üzere literatür belirli kilit kelimeler ile taranarak analiz edilmiş ve kapsayıcılık, yol ve yöntem sağlama ve etik literatüre uygunluk gibi kriterlere göre değerlendirilmiştir. Etik sorunlarla ilgili yapılan çalışmalar ve bilgi kaynakları, ortaya koydukları teoriler, alana katkı sağlamak üzere sundukları yol ve yöntemler, etik yapay zekâ modellemesi için kullanılabilir olduğu değerlendirilen araçlar ve sonraki araştırmalar için verilen tavsiyeler gözden geçirilmiştir.

Bu değerlendirme neticesinde günümüzde yapay zekâ alanında tespit edilen etik sorunların birbirinden farklı alanlarda ve birbirleri arasında önceliklendirilmiş şekilde sıklıkla var olduğu görülmektedir. Etik sorunların temel olarak gizliliğin korunması, veri güvenliği ve

doğruluğu ile ilgili olduğu bunun yanı sıra verilerin kötü maksatlarla kullanımı veya kamusal alanda oluşabilecek güvensizlik durumlarının da yaşanan önemli sorunlar olarak karşımıza çıktığı görülmektedir (Stahl, 2021). Etik sorunların üstesinden gelmek için yapılan sistematik inceleme sonucunda, açıklık, adalet ve hakkaniyet, zararsızlık ve fayda sağlama temalarının diğerlerine göre daha sık ele alındığı ortaya çıkmıştır. Ayrıca, sorumluluk taşıma, güvenlik, özerk çalışabilme yeteneği ve kontrole açıklık gibi ilkelerin de önemli ölçüde ele alındığı görülmüştür.

Meydana gelme olasılığı ve etkileri dikkate alındığında, otonom sistemlerdeki etik ihlallerin toplumsal bir güvenlik açığına neden olması olası bir durum olarak karşımıza çıkmaktadır. İlk analiz neticesinde tespit edilen etik konuların Royakkers ve diğerleri (2018), Jobin ve diğerleri (2019), Treviño ve Weaver (2001), IEEE (2017), Montreal Declaration (2017), Asilomar (2017) ve Floridi ve diğerleri (2018) tarafından yapılan çalışmalarındaki sonuçlarla uyumlu olduğu tespit edilmiştir.

Etik değerler belirlendikten sonra etik yönetim modeli oluşturma aşamasına geçilmiştir. İhlallerin engellenmesi noktasında kullanılabilecek araçlardan olan etik yönetim; güvenlik açıklarına karşı etik standartlar belirleme, değerlendirme ve etik dışı davranışları asgari düzeye indirmek için bir süreç ortaya koyma süreci olarak ifade edilir. Etik yönetim modelinin tasarlanmasındaki amaç; otonom sistemlerin gerek modelleme gerekse kullanım aşamasında karşılaşılabilen etik riskleri tespit etmek ve etik dışı davranışların azaltılması için yapılması gerekli işlem basamaklarının tespit edilmesine imkân sağlamaktır. Etik problemlerin belirlenmesinde yapay zekânın taşınması gereken önemli asgari değerlerin tespit edilmesinde, Otonom ve Akıllı Sistemlerle İnsan Refahını Önceliklendirme Vizyonu tarafından geliştirilen Prensipler (IEEE, 2017), Yapay Zekânın Sosyal Olarak Sorumlu Gelişimi Projesi kapsamında Montreal Üniversitesi tarafından geliştirilen “Sorumlu Yapay Zekâ için Montreal Deklarasyonu” (Montreal Declaration, 2017), Yaşamın Geleceği Enstitüsü tarafından öne sürülen “Asilomar Yapay Zekâ Prensipleri” (Asilomar, 2017) ve Floridi ve diğerleri (2018) tarafından yapılan kapsamlı çalışmalar değerlendirilmiş ve sonuç olarak tespit edilen dört temel ilke: fayda sağlama, zarar vermeme, adalet ve açıklanabilirlik ilkeleri bu çalışmada kullanılan modelin çekirdeğini oluşturmuştur.

Çalışmada kapsamlı bir literatür taraması yapılmış, elde edilen bilgiler model oluşturmak için kullanılmıştır. Çalışmada ayrıca belirli bir araştırma sorusu veya konu hakkında var olan

tüm bilimsel literatürün taranması, derlenmesi, analizi ve sentezi süreci olan Sistematik inceleme de kullanılmıştır (Maşa ve Karadağ, 2019). Yapılan sistematik inceleme ile yapay zekâda etik kontrol modeli oluşturmak için en uygun ve güvenilir model belirlenmeye çalışılmıştır.

Çalışmada ele alınan bir diğer durum sistemlerdeki güvenlik açıklarını tespit edebilmek ve risk faktörlerini belirleyebilmek için kullanılması gereken metodolojinin tespit edilmesidir. Modelin risk yönetim sürecinde potansiyel riskleri belirlemek, anlamak ve gerekli önlemlerin alınmasını sağlama üzerine odaklanılmıştır. Bu yaklaşım için uygun bir yöntem olan NIST Çerçevesi araştırma modeline dâhil edilmiştir. NIST Güvenlik Çerçevesi; tehditler, güvenlik açıkları ve etkileri gibi güvenlik risklerinin tespit edilmesi ve azaltılması konusunda ABD Siber Güvenliği Geliştirme Yasasında (NIST Law, 2018) onaylanan ve özel sektör firmaları ile kritik altyapı sunan işletmeler için kullanılması gerekli görülen endüstri standartlarını ve en iyi uygulamaların bütününe ifade etmektedir. Temel olarak NIST Çerçevesi hesap verebilirlik ve diğer siber güvenlik risklerini değerlendirme ve yönetme fonksiyonlarında rehber ve düzenleyici bir yaklaşım olarak kullanılmaktadır. Stoneburner ve diğerleri (2002) tarafından oluşturulan risk yönetim kılavuzu ve Nieves ve diğerleri (2017) tarafından oluşturulan bilgi güvenliği el kitabının temel amacı sistemin görevlerini gerçekleştirmesini sağlamak için bilgi depolayan ve işleyen süreçlerde güvenlik risklerini en aza indirmektir.

Stoneburner ve diğerleri (2002) tarafından oluşturulan “Risk Yönetim Kılavuzu” Sistem Karakterizasyonu, Tehdit Belirleme, Güvenlik Açığı Belirleme, Kontrol Analizi, Olasılığın Belirlenmesi, Etki Analizi, Risk Belirleme, Kontrol Önerileri ve Sonuçların Belgelenmesi adımlarından oluşan bir süreç içermektedir. Modelde kullanılan kılavuz deneyim ya da teknik bilgisine bakılmaksızın uygulayıcılara yol gösterir. Süreçlerin hedefi riski en aza indirmek veya sonlandırmak için önlemlerin belirlenmesine yardımcı olmaktır. Kılavuza göre olası olumsuz durumları belirlemek için tehdit ve güvenlik açıkları zararın büyüklüğüne göre analiz edilmelidir.

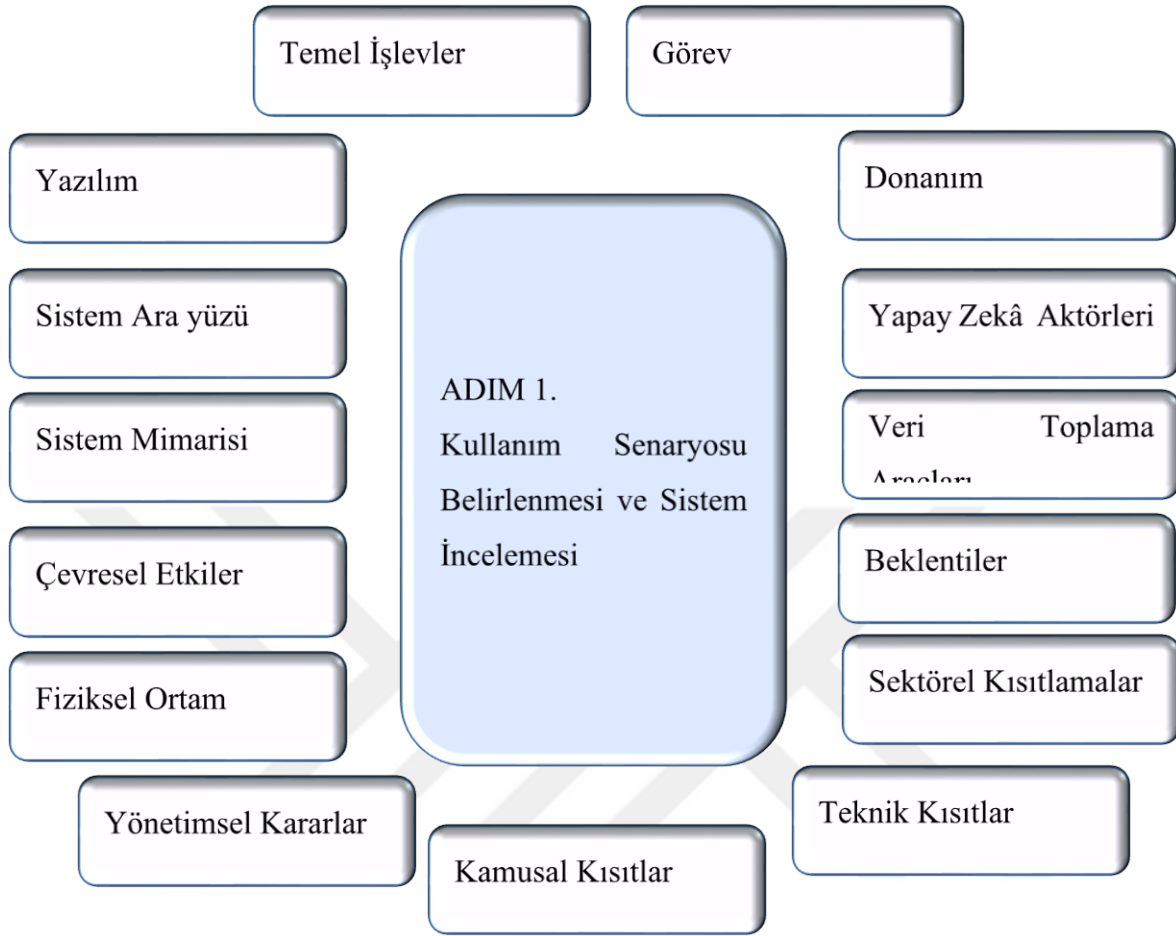
Aşağıda basamakları sunulan modelde; Stoneburner ve diğerleri (2002) tarafından oluşturulan bilgi teknolojilerinde “Risk Yönetim Kılavuzu” , Nieves ve diğerleri (2017) tarafından oluşturulan Bilgi Güvenliği El Kitabı basamaklarına, Otonom ve Akıllı Sistemlerle İnsan Refahını Önceliklendirme Vizyonu (IEEE, 2017) ve Yaşamın Geleceği

Enstitüsü tarafından geliştirilen Asilomar Yapay Zekâ Prensiplerine (Asilomar, 2017), Floridi ve diğerleri (2018) tarafından sunulan Yapay Zekâda Etik Kontrol politikalarına ve Montreal Üniversitesi (Montreal Declaration, 2017) tarafından geliştirilen Sorumlu Yapay Zekâ Deklarasyonuna uygun olarak hazırlanmıştır.

4.4. Model Basamaklarının Belirlenmesi

4.4.1. Kullanım senaryosu belirlenmesi ve sistem incelemesi

Önerdiğimiz modelde yapay zekâ sisteminde etik dışı davranışlar değerlendirilirken ilk adım, Stoneburner ve diğerleri (2002) tarafından oluşturulan Risk Yönetim Kılavuzu ve Nieves ve diğerleri (2017) tarafından oluşturulan Bilgi Güvenliği El Kitabında da olduğu gibi kullanım senaryosunun belirlenmesi ve kapsamın tanımlanmasıdır. Bu adımda, otonom sistem oluşumuna dâhil olacak donanım, yazılım, kullanılacak ara yüz gibi sistemsel birimler ile toplanması gereken veri, sisteme dâhil olması planlanan insanlar ve en önemlisi yapılması gereken görev birlikte değerlendirilerek yapay zekâ sisteminin kullanım sınırları belirlenir. Kullanım senaryosunu oluşturmak, olası etik dışı davranışların daraltılması ve alınabilecek tedbirlerin belirlenmesi için gerekli olan birtakım önemli bilgileri sağlar. Bu bilgilerden biri olan ve sistemin kullanım senaryosu gereği bağlantılı olabileceği ortamları belirlemek; etik ilkelerin uygulanması için gerekli olan metodolojinin de ortaya koyulmasını sağlar.



Şekil 4.2. Kullanım Senaryosu Belirlenmesi ve Sistem İncelemesi (Stoneburner vd, 2002 ve Nieves vd. 2017'den uyarlanmıştır)

Modelde yapay zekâ sisteminde kullanılabilecek etik kuralları belirleyebilmek için öncelikli olarak otonom sisteminin detaylı incelemesi yapılmalıdır. Etik değerlendirmeler yapabilmek adına otonom sistemin temel özellikleri anlaşılmalı ve ortaya koyulmalıdır. Otonom sistem incelenirken tablo 4.1'de gösterilen asgari bilgiler toplanmalıdır (Stoneburner ve diğerleri, 2002). Yapay zekânın çalışacağı fiziksel ve sosyal ortam, sistem tarafından gerçekleştirilmesi istenen görevler ve sistemin içinde bulunacağı süreçler belirlenmelidir. Sistemin ulaşacağı verinin içeriği, hassasiyeti ve bu hassasiyeti karşılamak üzere sahip olması gereken donanım, otonom sistemin yapısal bileşenlerini belirlemede kullanılacaktır. Yapay zekâ aktörleri ve bu aktörlere ek olarak teknik destek sağlayan kişi veya kuruluşlar, sistem tasarımında çalışan personelin güvenilirliği, sisteminin içinde bulunduğu nem, sıcaklık, toz gibi fiziksel ve çevresel faktörler, güvenlik ortamı belirlenirken incelenmelidir.

Araştırma safhasında herhangi bir faaliyete başlamadan önce kullanıcı ihtiyaçları öncelikli olarak değerlendirilmelidir (Higgins ve diğerleri, 2014).

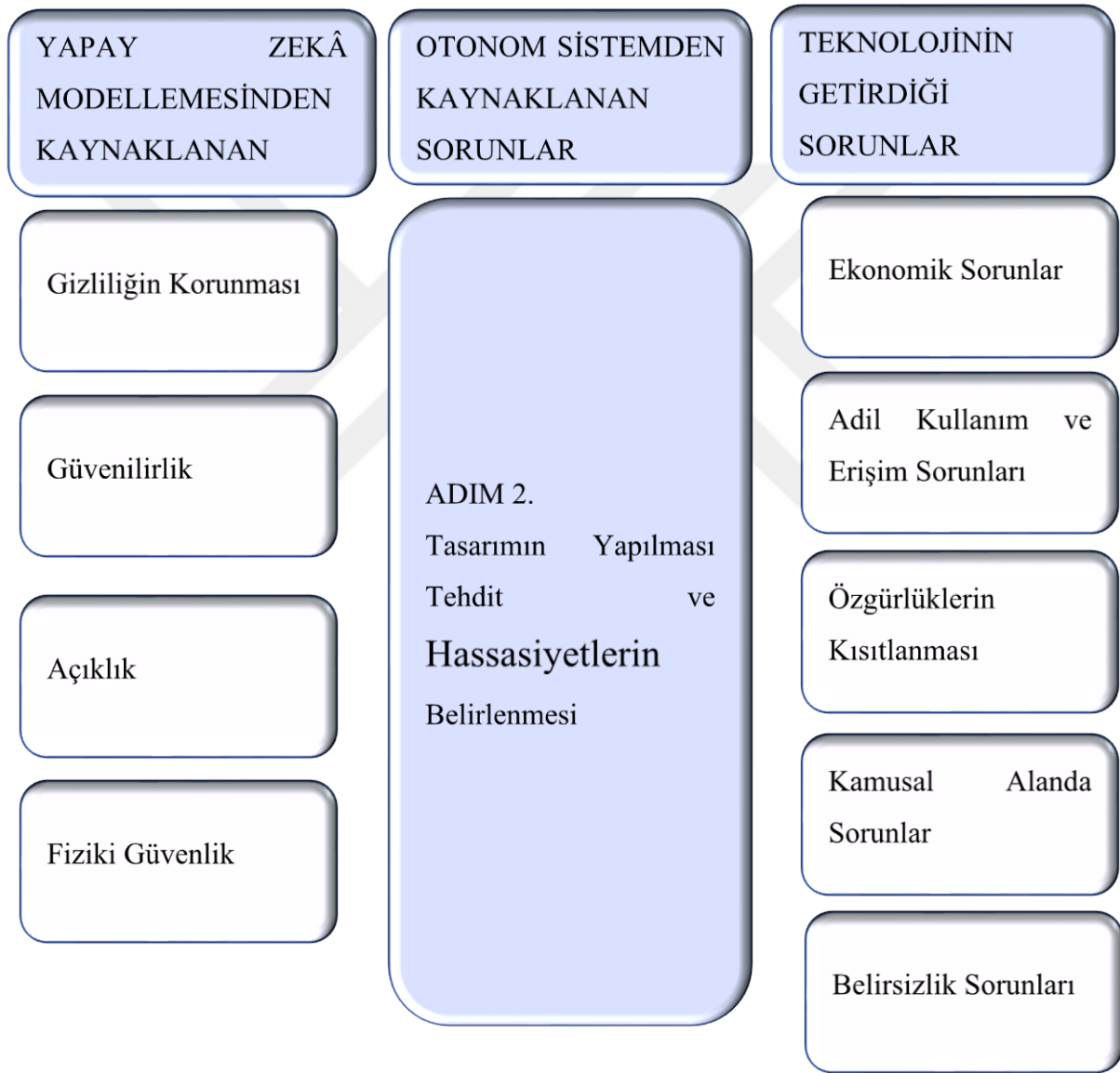
Ülkesel güvenlik politikaları ve gereksinimler, yönetsel karar ve kurallar ile sektörel güvenlik algıları otonom sistemin dış ortama uyumlu hale getirilmesine katkı sağlayacaktır. Sistemin mimarisi, sistem akış şeması ve topolojisi, veri depolama/koruma mekanizmaları, kimlik doğrulama ve şifreleme gibi teknik kontrol birimleri, erişim sağlama/engelleme ve gerektiğinde erişim kısıtlama prosedürleri; yönetici/üretici/kullanıcı şifresi belirleme, yeniden başlatma ve kurtarma seçenekleri, beklenmedik durumlara karşı yedekleme ve sürdürülebilir bir bakım yönetimi kullanmak sistemin güvenliğini artıracak gibi etik ihlallerin de önüne geçilmesine imkân tanıyacaktır. Engellenemeyen saldırılardan korunan bir sistem sağlamak için ağ kaynaklarını ve ağdaki iletişim ve cihaz bağlantıları gibi verileri güvence altına alması gerekir. Güvenliğin sağlanması için sisteme getirilen değişiklikler de ayrıca değerlendirilmelidir (Yick, Sharma, Lopez ve Park, 2016).

4.4.2. Tasarımın yapılması tehdit ve hassasiyetlerin belirlenmesi

Etik ihlalleri önlemek adına önerilen modelde, tasarımın yapılması adımı dikkat edilmesi gereken iki önemli nokta bulunmaktadır. Bunlar tehditler ve hassasiyetlerdir. Tehdit, otonom sistem için iç veya dış faktörler tarafından, belirli bir görev icra edilirken karşılaştığı sistemde açık meydana getiren tehlikelerdir. Hassasiyetler ise tehditler tarafından istismar edilme olasılığı olan zayıflık potansiyelidir. Tehdidin var olması, tek başına etik dışı uygulamalara yol açmayacaktır. Tehditlerin kaynağı, otonom sistemin görevi ve mevcut tedbirlerin bütünsel şekilde değerlendirilmesi esastır (Stoneburner ve diğerleri, 2002). Bu adımda amaç; tasarımda göz önüne alınması gereken potansiyel tehdit kaynakları ve hassasiyetleri belirlemek ve tasarım sürecinde dikkate alınması için potansiyel tehdit ve hassasiyetleri belirlemektir.

Yukarıda da bahsedildiği gibi tehdit kaynağı, bir yapay zekâ sisteminin işleyişinde, topluma veya kişilere zarar verme potansiyeli olan herhangi bir durum veya ihlaldir. Tehdit kaynağı kişisel, çevresel veya toplumsal olabilir. Değerlendirme yapılırken, tüm potansiyel tehditlerin ve hassasiyetlerin dikkate alınması; bilgilerinin ne şekilde ve ne maksatla saklandığının bilinmesi gerekir (Güngör, 2015). İhmal ve hatalar gibi kasıtsız ve beklenmedik durumlar dışında, kasıtlı etik ihlaller de tehdit kaynağı olabilir. Yapay zekâ

sistemlerinin ve var olan verilerin etik olarak kullanılabilirliğini veya gizliliğini bozmak için sistem güvenliğini kasıtlı olarak aşma girişimleri de göz önüne alınmalıdır. Burada önemli husus bütünlüğün ve kullanılabilirliğin sistemdeki verilerin gizliliğini sürdürecektir şekilde sağlanmasıdır (İren ve Can, 2017). Tüm bu tehdit ve hassasiyetler bir arada değerlendirilerek; Tasarımın Yapılması, Tehdit ve Hassasiyetlerin Belirlenmesi adımı tablo 5.2. de görüldüğü şekilde bütünsel olarak değerlendirilebilir.



Şekil 4.3. Tasarımın Yapılması Tehdit ve Hassasiyetlerin Belirlenmesi (Stoneburner vd, 2002 ve Nieleş vd. 2017’den uyarlanmıştır)

Tasarımın yapılması sırasında dikkat edilmesi gereken sorunları temelde dört ana başlıkta inceleyebiliriz. Yapay zekâ modellemesinden kaynaklanan sorunlar kapsamında; gizlilikten

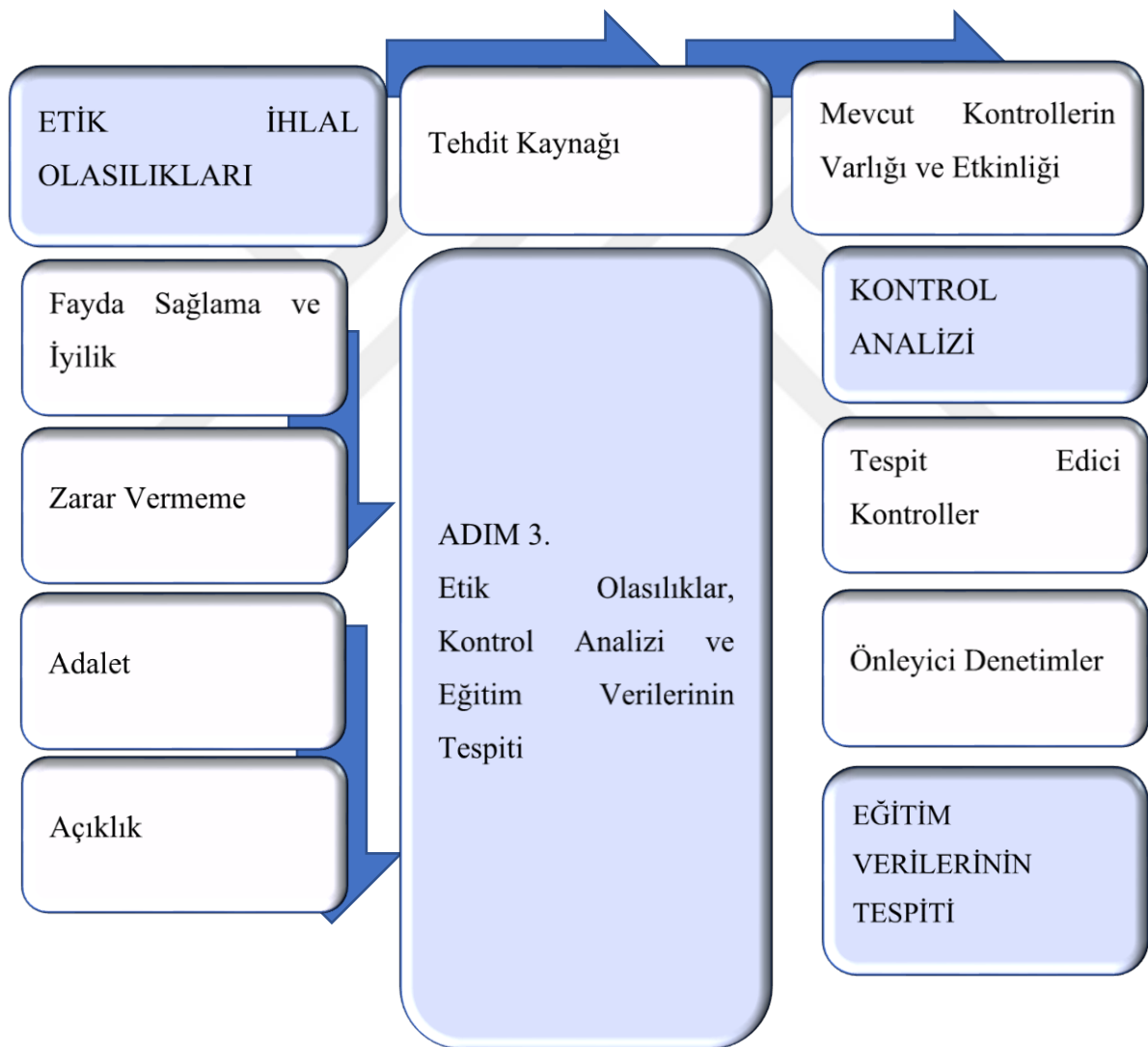
yoksun işlemlere izin verilmesi ve kişisel verilerin kötüye kullanılması gibi “gizlilik sorunları”, kaliteli veri eksikliği, veri doğruluğu eksikliği ve bütünlük gibi “veri güvenilirlik sorunları”, sorumluluk ve şeffaflık eksikliği, önyargı ve ayrımcılık gibi “açıklık sorunları” ve fiziki sorunlar olarak değerlendirilebilir (Stahl, 2021). Yapay zekâ sistemleri kendisinde önceki teknolojinin getirdiği sorunları da toplumsal olarak derinleştirme ve geri dönülemez duruma getirme potansiyeline sahiptir. Var olan bazı işlerin ortadan kalkması, ekonomik gücün tek aile ya da dar alanlarda yoğunlaşması, inovasyon maliyetlerinin artması teknolojinin getirdiği “ekonomik sorunlardır”. Adil kullanım ve erişim, gelir dağılımında bozulma ekonomik sorunları daha da ileriye taşıma potansiyeline sahiptir (Yereli ve Şahin, 2019). Tartışmalı veri sahipliğine imkân sağlaması, son kullanıcıların temel insan haklarını ihlal edebilme potansiyeli, adalet sistemi üzerinde olumsuz etki, kamu hizmetlerine erişim eksikliği ve hassas gruplar üzerinde yaratabileceği olumsuz etkiler iş ve kullanım senaryosuna göre belirlenmelidir (Stahl, 2021). Eşitsiz güç ilişkileri ve güç asimetrisi yaratma potansiyeli, demokrasi üzerindeki olumsuz etkileri, askeri kullanım potansiyeli ve insan teması ve kararının azalması kamusal alanda karşılaşılabilecek sorunlardandır (Isaak ve Hanna 2018; Coeckelbergh, 2020). İstenmeyen ve öngörülemeyen etkileri de yapay zekânın “belirsizlik sorunları” olarak değerlendirilir.

Henüz çağımızda karşılaşmadığımız ama hazırlıklı olmamız gereken sorunların da değerlendirmeye alınması gerekir. Bu kapsamda makine bilincinin oluşması, süper zekânın orantısız ve yapay zekânın uyanışı gibi hususları “otonom sistemden kaynaklanan sorunlar” olarak değerlendirilebilir (Stahl, 2021). Etik ihlal geçmişinin kayıt altına alınması, belirli aralıklar gözden geçirilmesi, ihlal ve olay sonuç raporlarının analizi yapay zekâ sistemini potansiyel tehlikeler karşısındaki kararlı hale getirir. Yapay zekâ kullanıcı ve üreticilerinin de dâhil olduğu deneyim paylaşım toplulukları kurmak; bilgi alışverişini artırır ve farkındalık sağlar.

4.4.3. Etik olasılıklar, kontrol analizi ve eğitim verilerinin tespiti

Bu adıma kadar otonom sistemin iş ve kullanım senaryoları belirlenmiş ve sistem detaylı şekilde incelenmiştir. Otonom sistemin tasarımına etki eden tehditler ve hassasiyetler bir önceki adımda belirlenmiştir. Etik olasılıklar, kontrol analizi ve eğitim verilerinin tespiti adımında, yapay zekâ sisteminde ihtiyaç duyulan etik ifadeler yeniden değerlendirilir ve sistemin amacı ile karşılaştırılır. Belirlenen etik riskler ve bu risklerin üstesinden gelmek

için gerekli olan güvenlik gereksinimleri ortaya koyulur. Yapay zekâ sistemleri yorum yapma, bilgi işleme, tahmin etme gibi belli amaçlara göre tasarlanır, programlanır veya geliştirilir (Nabiyev, 2005). Etik olasılıkların belirlenmesinde, yapay zekâ sistemin tasarımı ve mimarisi etik ihlallere yol açabilme ihtimallerine göre analiz edilmelidir. Sistemin güvenlik gereksinimlerini karşılama potansiyeli değerlendirilmeli, gerekirse yeniden yapılandırılmalı ve doğrulanmalıdır. Otonom sistemin belirlenen durumlar karşısındaki tedbirlerini içeren etik ihlalleri önleme süreci, sistem üretilmeden veya kullanıma başlanmadan önce tamamlaması hedeflenmelidir.



Şekil 4.4. Etik Olasılıklar, Kontrol Analizi ve Eğitim Verilerinin Tespiti (Stoneburner vd, 2002 ve Nieves vd. 2017'den uyarlanmıştır)

Yapay zekâ sistemlerinin yapısında potansiyel bir ihlal olasılığını belirlemek için tehdit kaynağı, etik ihlalin nitelik ve niceliği, ihlal durumunda önceden belirlenen önleme mekanizmaları ve mevcut kontrollerin varlığı ile etkinliği bütünsel olarak değerlendirilmelidir. Potansiyel tehdit kaynağının bir ihlale dönüşme durumunu sayısal veya kategori olarak belirtmek mümkündür (Stoneburner ve diğerleri, 2002). Bu modelde ortaya çıkış ihtimali için sayısal bir değer yerine kategorilere ayırmak daha kullanışlı olabilir. Tehdit kaynağı ile karşılaşma olasılığı; yüksek, orta ve düşük olarak tanımlanabilir. Tehdit kaynağı; toplumsal olarak oldukça yaygın ve hassasiyet gösterilen durumlar, yüksek tehdit kaynağı olarak değerlendirilir. Bu etik ihlalde, istenmeyen etkiler diğer durumlara göre daha fazladır. Orta tehdit kaynağı toplumsal olarak karşılaşılan ve zaman zaman görülen durumları kapsar. Etik ihlallerin önlenmesi için birtakım önlemler sistemsel olarak başarılı bir şekilde uygulanmıştır. Düşük tehdit durumlarında ise; etik ihlal toplumu ve kişileri rahatsız etmeyecek veya sistemsel olarak zarar veremeyecek durumdadır. İhlali önlemek veya zararını engellemek için sistemsel kontrollerin mevcut olduğu ya da kısa sürede sağlanabildiği durumlardır.

Etik olasılıklar ve kontrol analizi adımıyla amaç, bir ihlalin yapay zekâ sistemlerinde toplumsal olarak istenmeyen güvenlik açıkları meydana getirme olasılığını en aza indirmek veya ortadan kaldırmaktır. Otonom sistemin kullanım amacı ve senaryosu ile ilişkili olan ortamının yapısına göre, etik dışı durumların potansiyel bir tehdit meydana getirme olasılığı belirlenirken mevcut veya planlanan güvenlik kontrolleri dikkate alınmalıdır. Yapılan kontrollerde kapsamı geniş tutmak, etik ihlalleri önleme başarısını artırabilir. Kontrol tedbirleri genel ve teknik olarak iki alt başlıkta incelenebilir. İş yapma prosedürlerinin ayrıntılı yazılması, görev tanım formlarında etik ihlallerden kaçınılmasının açıkça ifade edilmesi, işe alımlarda kritik personel için ayrıntılı inceleme yapılması gibi hususlar genel kontroller iken; erişim kontrol için tanımlama ve kimlik doğrulama algoritmaları, şifreleme mekanizmaları, izinsiz giriş tespit yazılımları teknik kontrol tedbirlerindendir. (NIST SP 800-18, 1998). Kontrol listesi oluşturma faaliyeti, bir kere yapılan ve devamlı uygulanan bir statik bir liste olarak değerlendirmemek gerekir. Bunun yerine; otonom sistemlerdeki ve toplumsal dinamiklerdeki değişimler analiz edilerek, kontrol listesinin geçerliliğini sağlamak gerekir.

Güvenilir yapay zekâ sağlama konusunda Floridi ve diğerleri (2018) tarafından beş ilke ortaya koyulmuştur. Bu ilkelerden iyilik, zarar vermeme, özerklik, adalet ve açıklanabilirlik

prensipleri bu modelde benimsenmesine rağmen; özerklik kavramı etik ilkeler ile doğrudan bağlantılı olmadığı gerekçesi ile çalışma kapsamına alınmamıştır. Modelde kullanılmayan özerklik kavramı temel olarak yapay zekânın karar verme erkine kendi başına sahip olması ve insanların bu gücü bilerek ve isteyerek bir denge içerisinde otonom sisteme devretmesini kapsamaktadır. Ortaya konulan model insan özerkliği ve yapay zekânın karar alma mekanizmasını nasıl yönetmesi gerektiğini ortaya koymayı amaçlamıştır. Bu nedenle özerklik ilkesi hem etik değerlerle direk ilgili olmadığından hem de insan kararlarının otonom sistemlere entegrasyonunu ele aldığından araştırmaya dâhil edilmemiştir. Modelde etik kontrol olarak uygulanması gereken dört prensip aşağıda sıra ile açıklanmıştır.

4.4.3.1. Açıklık

Açıklık ilkesi yapay zekâ sistemlerinin, tasarım modellerinin ve sistemlerde kullanılan verilerin izlenebilir ve şeffaf olmasını kapsar. Veri toplama ve verilerin tasniflenmesine ilave olarak, algoritmalar da dâhil olmak üzere tüm süreçlerin; açıklık, kontrol edilebilirlik ve şeffaflığını artırmak ilkenin en önemli amaçlarından (Floridi ve diğerleri, 2018). Yapay zekâyâ sahip sistemlerde açıklık sağlamak üzere, kontrol mekanizmaları kurmak ve denetimi kolaylaştırmak gerekir. Kontrol mekanizmaları hatalı kararların tespit edilmesine ve önlenmesine destek olur. Karar noktalarında kullanılan algoritmaların sonuçlara nasıl vardıklarının insan kontrolüne açık olması bu ilke kapsamındadır (Robbins ve Henschke, 2017). Açıklık ilkesinin uygulanmasında, toplumsal kurallar kadar teknik uygulamaların da anlaşılabilirliği ve denetlenebilirliği sağlanmalıdır. Yapay zekâ sistemlerinde hem teknik süreçlerin hem de toplumsal olarak uzlaşmış genel uygulamaların açıklanabilirlik ve şeffaflığı, etik yapay zekâ için önemli bir basamaktır. Bu özelliklere ilave olarak; insan merkezli, toplumsal değerlerin bilincinde ve duyarlı çalışabilen yapay zekâ sistemleri modellemek ve otonom sistemin verdiği kararların kontrol edilebilirliği ve izlenebilirliği açıklanabilirlik kapsamında değerlendirilmelidir. Açıklanabilirlik ilkesi yapay zekâ ve makine öğrenmesi alanlarındaki şeffaf olmayan algoritmalara karşı bir önlem olarak ortaya çıkmıştır (Robbins, 2019). Bu ilke kapsamında otonom sistem karar mekanizmalarındaki işlemlerin doğruluğunun kontrol edilmesi, verilen karara etki eden veri, algoritma ve tasarım süreçlerinin denetlenebilir olması sağlanmalıdır.

4.4.3.2. Zarar vermeme

Asimov tarafından oluşturulan ve kabul gören robot yasalarına göre otonom sistemler insanları zarara uğratmamalı, durumunu olumsuz etkilememelidir (Anderson, 2008). Bu yasanın etik ilkelerde uygulanması çok çeşitli alanlarda çalışma gerektirir. Otonom sistem ile etkileşime geçecek tüm birimlerin incelenmesi ve birbirilerine olan etkilerinin değerlendirilmesi gerekir. Otonom sistemin zarar vermeme ilkesine uygunluğunun denetlenmesi ve ancak güvenli olarak işaretlenmesi ile bir sonraki basamağa devam edebilmesi gerekir. Zarar vermeme ilkesi değerlendirilirken, otonom sistemin insana olan etkisi öncelikli olarak incelenmelidir. Sistemin toplum içerisindeki bireylere duygusal ve fiziksel anlamda zararlı etkileri dikkate alınmalıdır. Sistemi meydana getiren bileşenler de dâhil olmak üzere, her bir alt uygulama ve yazılım kurallarının insana verebileceği olası zararlar analiz edilmeli, güvenlik prosedürlerine uyumlu olması sağlanmalıdır (NIST SP 800-18, 1998). Fiziksel zarar vermektan kaçınma kavramı aynı zamanda duygusal bütünlüğün sağlamasını da kapsar. Bu nokta da yapay zekânın dezavantajlı gruplara karşı pozitif ayrımında bulunup bulunmayacağı da sistem modelleme aşamasında düşünülmesi gereken konulardandır. Zarar vermeme ilkesi kapsamında insanlarla birlikte; doğal çevre ve diğer tüm canlılar da değerlendirilmelidir.

Bir diğer zarar verme yolu, dış etkenler tarafından sistemin istismar edilmesi ve oluşan güvenlik açıkları sonucu, insana ya da topluma zarar vermedir. Dış saldırılara ve güvenlik açıklarına karşı kendini koruyabilir bir sistem modellemek, zarar vermeme ilkesi altında düşünülmeli ve güvenlik sağlanmadan süreç devam ettirilmemelidir. Saldırı karşısında istismara izin verilmemesi adına günümüzde sıklıkla başvuru ve sistemin ilk ayarlarına dönmesini sağlayan planlar, sistem içerisinde mutlaka bulunmalıdır (Floridi ve diğerleri, 2018). Sisteme verilen diğer tüm komutlardan üstün şekilde belirlenecek bir kodun, ne durumda olursa olsun sistemi çıkış ayarlarına dönüşe zorlaması, bunun da sonucu olarak genel güvenlik düzeyine geri dönüşüne olanak sağlanması sistemin güvenliğini artıracak ve olası riskleri en aza indirecektir.

Modelde bulunması gereken diğer bir konu gizliliğin ve verilerin korumasının garanti altına alınmasıdır. Kişisel verileri içeren; genetik ve sağlık bilgileri, ırk ve etnik köken, siyasi veya dini inançlar, cinsel tercihler gibi hususlar hakkında bilgi toplanması ve bu bilgilerin kullanılabilmesinin de belirli esaslara göre yürütülmesi gerekmektedir (Canada, 2000). Bu

kapsamda sistemin kullandığı verilerin istenmeyen üçüncü şahısların eline geçmesi engellenmelidir. Hükümetlerin de durumu ayrıca değerlendirilmeli, konusu suç teşkil eden durumlarda adli makamlarla veri paylaşımı gibi hususlar mevcut yasalar çerçevesinde tespit edilmelidir. Ancak otonom sistemlerin dâhil olduğu suç eylemlerinde cezai sorumluluğun belirlenmesi için kesin hükümler çoğu durumda bulunmamaktadır (Caşın, Al ve Başkır, 2021). Konu ile ilgili mevzuatın olmadığı durumlarda şirket politikaları ve kullanıcıların talepleri göz önüne alınmalı ve sistem bu yapıda tasarlanmalıdır.

4.4.3.3. Adalet

Adalet; hak ve hukuka uygunluk, doğruluk ve adil olma anlamına gelmektedir (TDK, 2022). Yapay zekânın ayrımcılığın tüm türlerine karşı adaleti sağlamak temel görevlerindendir (Montreal Declaration, 2017). Avrupa Bilim ve Yeni Teknolojilerde Etik Grubuna (2018)' e göre yapay zekâ teknolojinin ve getirdiği faydaların tüm topluma eşit şekilde dağıtılmasını sağlamalıdır. Bu nedenle etik yapay zekâ oluşturma sürecinde; otonom sistemden etkilenen tüm unsurların dikkate alınması ve faydaya dâhil edilmesi gereklidir. Doğrudan ya da dolaylı sistemden etkilenebilecek tüm paydaşların doğru şekilde tespit edilmesi ve fikirlerinin alınması sürece olumlu katkı sağlayacaktır. Tasarım esnasında eşit temsil sağlanması, adalet ilkesinin doğru uygulanması ile gerçekleşecektir.

Etik ve güvenilir bir otonom sistem meydana getirmek için; önyargı içermeyen, titiz şekilde belirlenmiş veri toplamak; kısıtlama, gereksinim ve karar noktalarını dikkatlice analiz etmek gerekir. Bunun tersi bir yaklaşım; gerek eğitim gerekse uygulama esnasında kullanılan veri setlerinin önyargı içermesi, eksik hususlar barındırması veya kötü yönetilmesi yapay zekâ sistemlerinin kullanıldığı toplumların düzenlerine zarar verebilir (Çalışkan, Joanna, ve Narayanan, 2017). Zararın önlenmesi için farklı geçmiş, kültür ve disiplinden gelen personelin modelleme sürecine dâhil edilmesi ve girdi çeşitliliği sağlayarak eksik hususların tamamlanması gerekir.

Otonom sistemler kadar, kullanıcılar için de adaletin sağlanması gerekir. Otonom sistemlerin tasarımında, tüm kullanıcıların erişimine olanak sağlayan, kullanıcı dostu ara yüzler oluşturulmalıdır. Sistemler; milliyet, dil, din, cinsiyet, yaş, bir örgüte ait olma, fiziksel veya başka engel durumları gibi hususların toplum içerisinde bir ayrımcılığa yol açmamasını sağlayacak ve herkese eşit mesafede olacak şekilde tasarlanmalıdır. Kapsayıcılık ve adil

erişim imkânı için; en geniş kullanıcı kitlesine hitap eden evrensel tasarım ilkeleri göz önünde bulundurulmalıdır (Floridi ve diğerleri, 2018). Sistemin aynı durumda benzer tepkiler vermesi olarak adlandırılabilen “güvenilirlik” kavramı da bu alanda tasarlanması gereken bir konudur. Sistemin çeşitli olaylara karşı davranışlarının tamamen aynı olmasa bile yakın ve tekrar edebilir olması sağlanmalıdır.

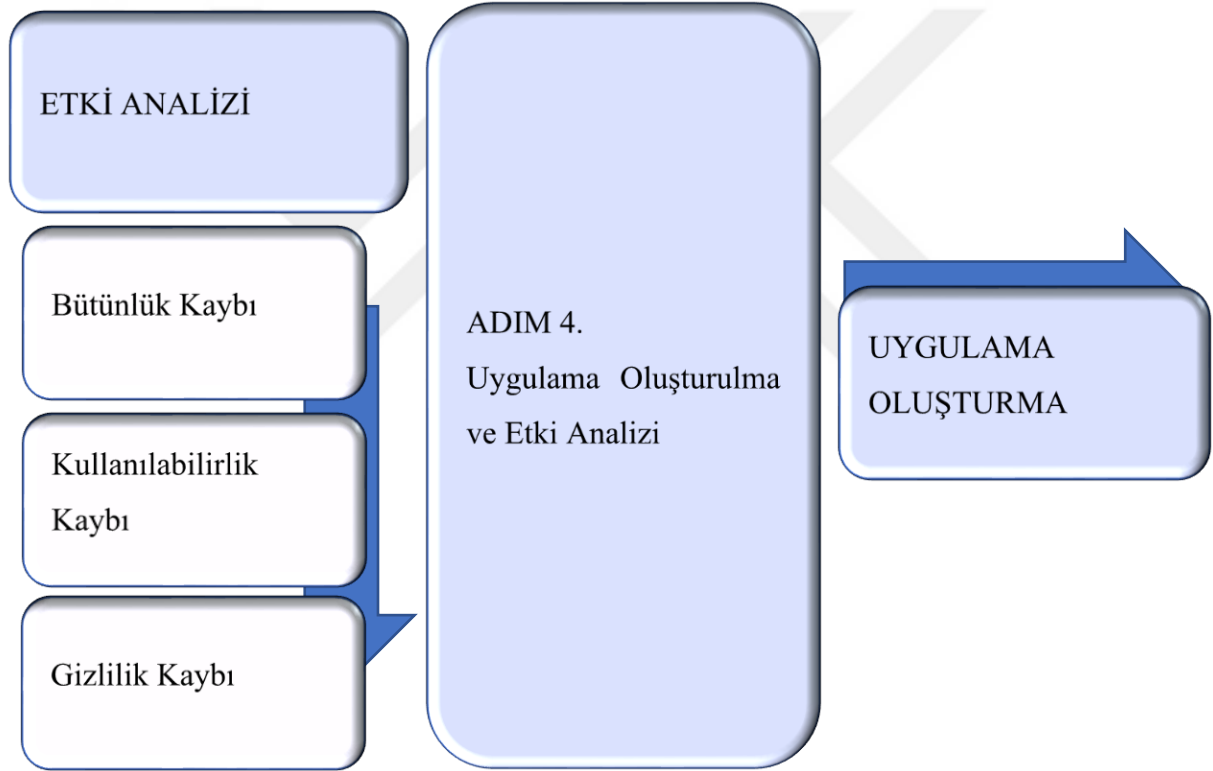
4.4.3.4.Fayda sağlama ve iyilik

Fayda sağlama ve iyilik ilkesinin temelinde zararı önlemek ve fayda sağlamak vardır. Otonom sistemin kullanım amacı göz önünde bulundurularak yapay zekâyâ sahip sistemlerin faydaları ile risk ve maliyetlerinin dengelenmesini de içeren bir grup kurallar bütünü içermektedir (Beauchamp ve Childress, 2012). Bu ilkeye göre sistemin topluma ve insanlara sağlayacağı fayda açık olmalıdır. Sistemin insanlığı ileriye taşıma ve toplumsal refaha olan katkısının tartışmalı olması bu ilkeye uyumsuz bir duruma yol açar. Bu nedenle öncelikli olarak sistemin hangi probleme fayda sağlamak adına modelleneyeceği ortaya koyulmalıdır (Cowls, King, Taddeo ve Floridi, 2019). Toplumsal refahı desteklemek için, sistemler oluşturulmadan önce, kullanıcılardan modelleyicilere kadar sistem ile etkileşime geçecek tüm birimlerin fikirleri alınmalı ve sürece katılımları sağlanmalıdır. Bu aşama aynı zamanda temel hak ve özgürlüklerin korunması, doğaya zarar vermeme, kaynakların etkin kullanılması, kısıtlıkların önlenmesi ve afet yönetimi gibi toplumun refahını artıracak konular açısından da değerlendirilmelidir (AIGGS, 2019).

4.4.4. Uygulama oluşturulma ve etki analizi

Etki Analizi; uygulama oluşturma aşamasında etik ihlallerin potansiyel etkilerinin tespit edilmesi ve değerlendirilmesi aşamasıdır. Etki analizi ile otonom sistemdeki etik ihlalin iş akışını ne şekilde etkileyeceği ölçülür ve potansiyel etkiler belirlenir. Analiz esnasında otonom sistemin kullanım alanları tüm temel bileşenler dâhil edilerek incelenir ve hassas noktalar ortaya çıkarılır. İhlallerin otonom sistem üzerindeki etkisinin belirlenmesi, değerlendirilmesi, önlenmesi veya azaltmaya yönelik yapılacak işlemler algoritmaya aktarılır. Etik ihlalin etki analizi, otonom sistemin önemli iş kesintilerine neden olan durumları, bu etkilerin maddi ve manevi olası maliyetlerini ve mümkün olan kurtarma stratejilerini içermelidir (Stoneburner ve diğerleri, 2002).

Otonom sistemin işleyişine en fazla etkisi olan, toplumsal değerlere en fazla zarar veren ve adli olarak probleme yol açan geçmiş olaylar öncelikli olarak değerlendirilmelidir. Etik bir ihlal sonucu yaşanan istenmeyen durumlarda sistemin ve sürecin nasıl korunacağı düşünülmelidir. Bunu sağlamak için otonom sistemin görevi, sistemdeki verinin içeriği gözden geçirilmelidir. Birden fazla etki tespit edilmek istenirse, ihlaller ile ilgili nitel veya nicel bir değerlendirme yapılarak öncelik belirlenmelidir. Bilgi teknolojilerinde istenmeyen olayların olumsuz etkileri; bütünlük kaybı, kullanılabilirlik kaybı ve gizlilik kayıplarının herhangi birinin meydana gelmesi veya birkaçının bir arada ortaya çıkması olarak tanımlanabilir (McCumber, 2005).



Şekil 4.5. Uygulama Oluşturulma ve Etki Analizi (Stoneburner vd, 2002 ve Nieves vd. 2017’den uyarlanmıştır)

McCumber'in bilgi güvenlik modeli içerisinde yer alan unsurları şunlardır;

➤ **Bütünlük Kaybı:** Veri ve sistemin uygunsuz değişiklikler sonucu korunamamasını ifade eder. Bilgi teknolojilerinde istemli veya istemsiz olarak yapılan yetkisiz değişiklikler verilerde ve otonom sistemde bütünlük kaybına yol açar. Sistemde kullanılan veri kayıpları

düzeltilmezse, otonom sistem hatalı kararlara neden olacağı gibi, dolandırıcılığa hassas hale gelecektir. Bütünlük kaybının düzeltilmemesi otonom sisteminin güvencesini azaltır.

➤ Kullanılabilirlik Kaybı: Otonom sistemin son kullanıcı tarafından kullanılmadığı durumları ifade eder. Sistem işlevselliğinin ve kullanılabilirliğini kaybeder.

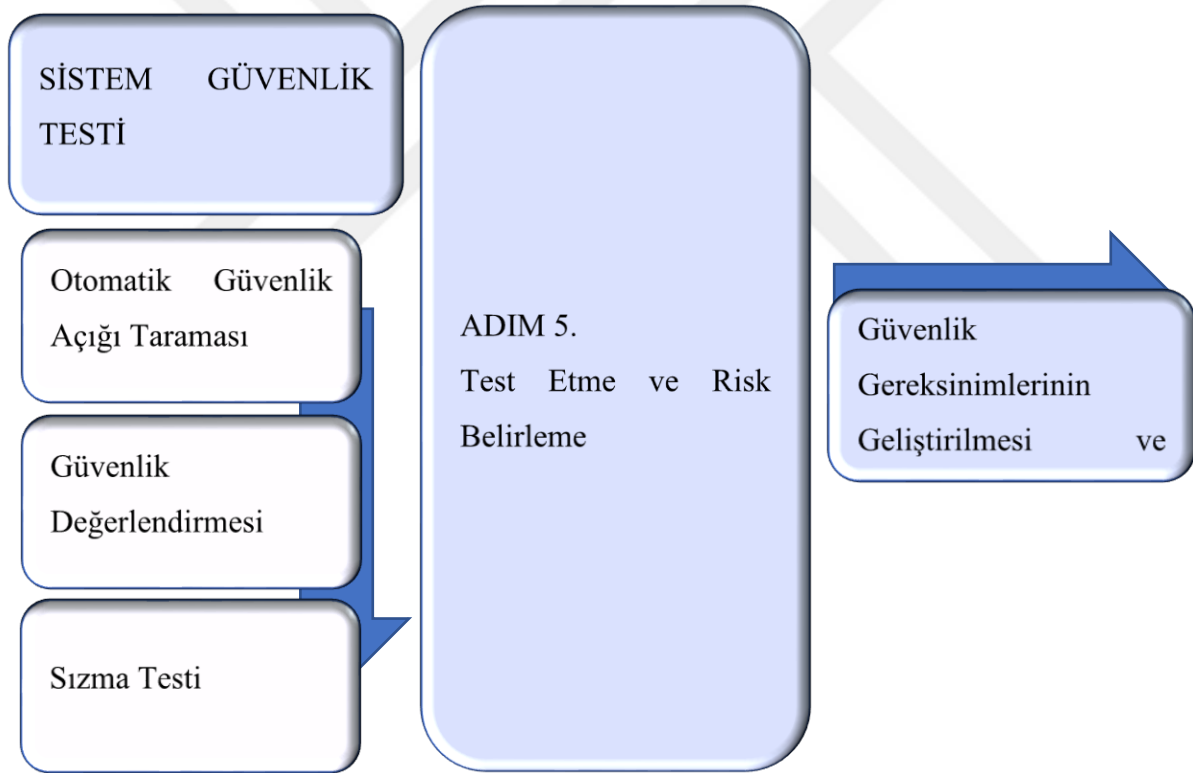
➤ Gizlilik Kaybı: Veri gizliliğinin yetkisiz kişiler tarafından açığa çıkarılması veya elde edilmesi anlamına gelir. Kişisel bilgilerin açığa çıkarılması otonom sisteme olan güveninin kaybolmasına ve sistem üreticileri hakkında yasal işlem yapılmasına neden olabilir. Temel olarak bilgiye yalnızca erişimi olan kişilerin erişebilmesini içerir (Henkoğlu ve Yılmaz, 2013). Her üç durumda da ihlaller yüksek maliyetli kayıplara veya otonom sistemin görevini yerine getirmemesine neden olabilir. Bu maksatla etki analizi sonucu müdahale gerektiren alanlar belirlenmeli ve maliyet-fayda ilişkisi gözetilerek çözüm ortaya koyulmalıdır. Etki büyüklüğünü belirlemek için ayrıca etik risklerin karşılaşımla sıklıkları, yaklaşık maliyetler ve tehdidin şiddeti gibi faktörler de değerlendirilir.

4.4.5. Test etme ve risk belirleme

Otonom sistemlerde etik kontrol, risklerin ortaya çıkması ile devreye giren reaktif veya proaktif yöntemler ile sağlanabilir. Her iki durumda da seçilmesi gereken yöntem; mevcut kaynaklara, etik riskin durumuna, tahsis edilen bütçeye, sahip olunan teknolojik imkânlarla veya testi gerçekleştirecek uzmana sahip olma gibi diğer faktörlere bağlı olarak belirlenebilir (Whittaker, 2000). Seçilebilecek sistemlerden ilki otomatik güvenlik açığı tarama aracıdır. Daha evvel tespit edilmiş güvenlik açıklarını uygulamada taramak ve tespit etmek için kullanılır. Otomatik tarama araçları için unutulmaması gereken; bazı potansiyel risklerin değerlendirilmesinde, otonom sistemin içinde bulunduğu kültürel durumun tam olarak değerlendirilemeyebileceğidir. Sistemin bulunduğu ortama göre, gerçek güvenlik açıklarını tespit edebilme yetenek ve kapasitesine dikkat edilmelidir. Otomatik tarama yazılımları tarafından; kültürel, sosyal veya ekonomik ortam ve bu ortamın getirdiği gereksinimler dikkate alınmadan potansiyel güvenlik açıklarının değerlendirilmesi; aslında probleme yol açmayacak bazı durumların riskli olarak belirlenmesine veya yanlış değerler dizisi ile hatalı olumlu raporlar üretebilmesine yol açabilir (Dustin, 2003). Etik riskleri belirlemede kullanılabilecek diğer teknik kapsamlı bir güvenlik değerlendirme sürecidir. Süreç; asgari olarak test komutlarını, test prosedürlerini ve olası sonuçları içerecek şekilde düzenlenmelidir. Güvenlik değerlendirme süreci, otonom sistemi ve sistem tarafından kullanılan tüm fonksiyonları çevresel ortamdaki etkinlikleriyle yüksek verimlilikle test

etmeyi amaçlar (Bertolino, 2007). Prosedürde amaç; genel ve özel etik standartları karşılamak ve amaçlanan güvenlik politikalarının uygulanmasını sağlamaktır. Etik risklerin gözden geçirilmesi ve aynı zamanda otonom sistemin farklı yönlerinin güvence altına alınması için kullanılır.

Risk değerlendirme sürecinde otonom sisteminin güvenliğini aşmak için yapılan girişimlere dayanma kabiliyeti de değerlendirilebilir. Sisteme tehdit oluşturacak özellikle insanların neden olduğu riskler de dahil olmak üzere birçok durumla alakalı geçmiş bilgiler yetersizdir, bu nedenle risk belirleme alandaki deneyim önemlidir. Risk değerlendirmesi adımı; otonom sistem için öngörülen güvenlik gereksinimlerinin, mevcut veya planlanan kontroller tarafından karşılanıp karşılanmadığı belirlenir. (Nieles ve diğerleri, 2017).

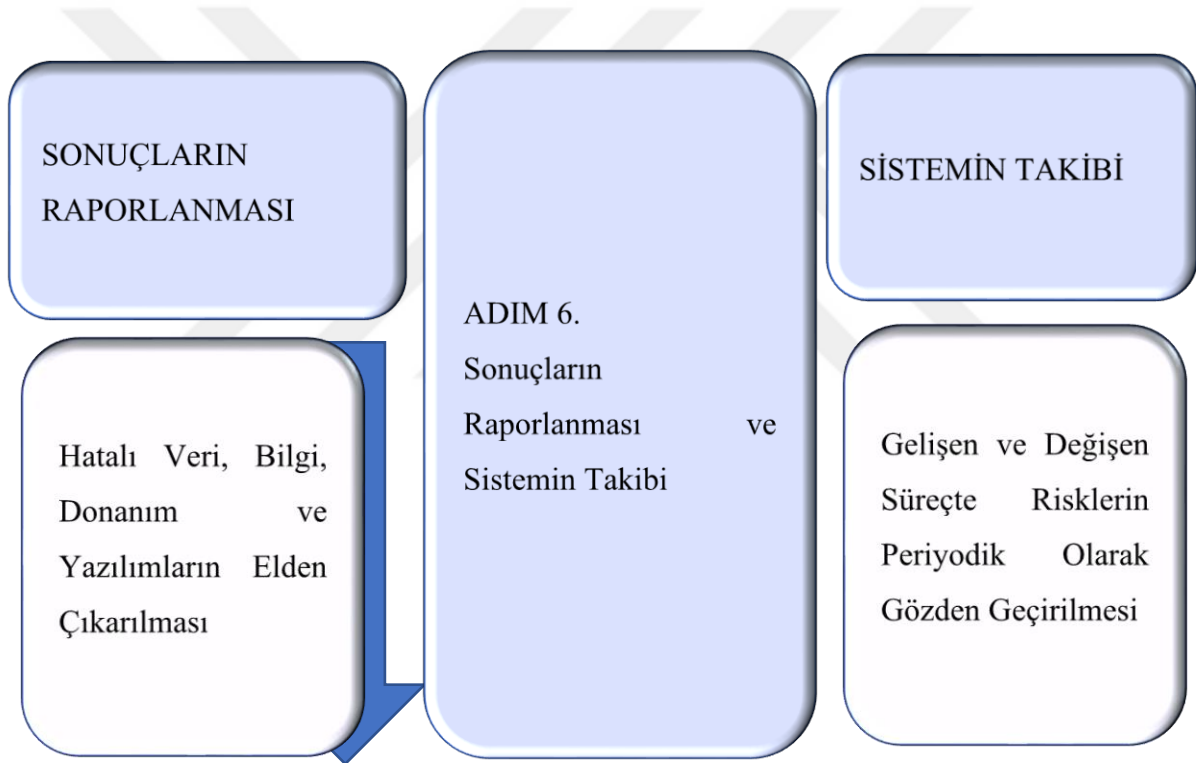


Şekil 4.6. Test etme ve Risk belirleme (Stoneburner vd., 2002 ve Nieles vd. 2017'den uyarlanmıştır)

Yapılan risk değerlendirmesi sonucunda; otonom sistem tasarımının, uygulamadaki kontrol gereksinimini karşılayıp karşılamadığına ilişkin sonuçlar, otonom sistemin ilişkili olduğu veriler, kişiler, donanım ve yazılım, algoritma ve prosedürler ile beraber bütüncül bir yaklaşımla ortaya koyulmuş olacaktır.

4.4.6. Sonuçların raporlanması ve sistemin takibi

Bu adımda risklerin sonuçları değerlendirilir ve birtakım değişikliklerin yol açabileceği olası problemler için sistemin takibi sağlanarak, yukarıdaki süreçlerde ortaya çıkan yüksek öncelikli risklerin giderilmesi sağlanır. Giderilemeyecek hatalı veriler, bilgi, donanım ve yazılımlar var ise sistem dışına çıkarılması gerekir. Süreçte kullanılan kişisel verilerin sistem dışına çıkarılma süreci; hataların arşivlenmesini, uygun olmayan verilerin ayıklanmasını veya imha edilmesini içerebilir. İmha edilecek veya değiştirilecek bileşenler dışında kalan verilerin uygun şekilde işlenmesi için gerekli faaliyetler yürütülür (Nieles ve diğerleri, 2017).



Şekil 4.7. Sonuçların Raporlanması ve Sistemin Takibi (Stoneburner vd., 2002 ve Nieles vd. 2017’den uyarlanmıştır)

Sistemin takibi aşaması; sistemin sahip olduğu güvenlik yöntemlerini daha iyi hale getirebilmeyi ve süreç içerisinde oluşabilecek değişikliklere cevap vermeyi içerir. Geliştirme aşamasında yapılan hatalar, truva atı gibi istenmeyen programlar, hatalı kodlar ve kötü niyetli girişimler meydana gelebilecek tehditler veya güvenlik açıkları arasındadır (Stoneburger ve diğerleri, 2017). Riskli bileşenlerden ayrılan otonom sistem, gerçek hayatta

kullanıma sunulur. Çevresel faktörlerin etkisi ile meydana gelen yeni tehditler değerlendirilmelidir. Sistem kendisinden beklene işlevleri yerine getirmeye başlamıştır.

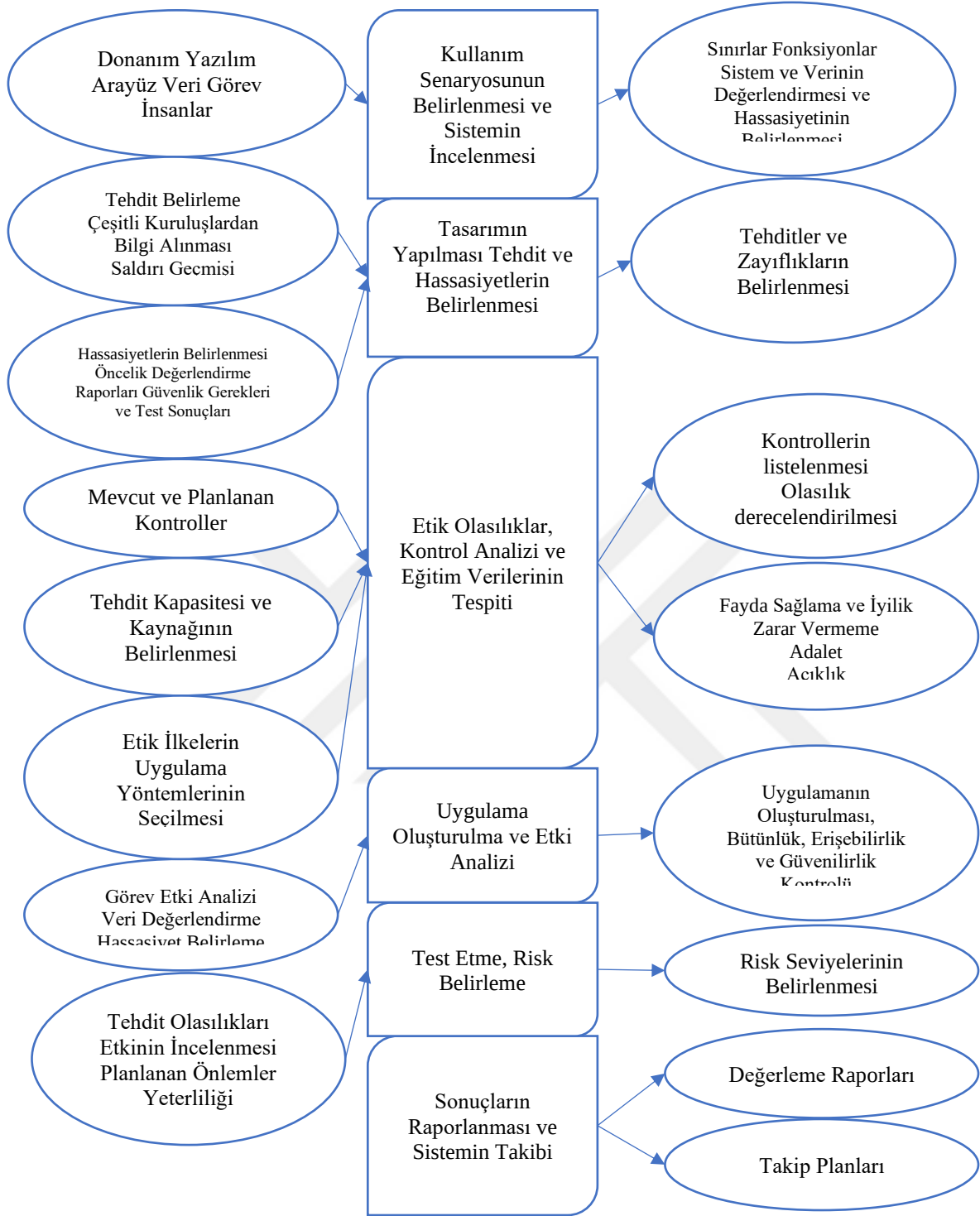
Yapay zekâya sahip sistemler, diğer makinelerden farklı olarak, elde edilen verilerden öğrenebilir ve algoritmik hesaplar ile yorum yapabilir, farklı yollarla daha etkin şekilde hedefe ulaşabilir. Bu süreç sonunda otonom sistem, değişiklikler yoluyla sürekli olarak gelişim gösterebilir (Doğan, 2019). Bu nedenle gelişen ve değişen süreçte; otonom sistemin etik değerlere uygun davranışlarda bulunmama riskini azaltmak için birtakım gereksinimler arasında değiş-tokuş yapılması ihtiyacı ortaya çıkabilir. Risklerin şeffaf olarak belirlenmesi ve dengede tutulması ve bu dengeyi sağlayacak mekanizmaların bulunması önemlidir. Bu nedenle riskli davranışlara karşı güvenlik kontrollerinin zaman içerisinde değişme potansiyeli göz önüne alınarak periyodik olarak gözden geçirilmesi gerekir (Nieles ve diğerleri, 2017).

4.4.7. Etik kontrol modelinin tamamlanması

Model yapılan literatür taraması ve etik ilkeler ışığında hazırlanmıştır. Tanımlanan modelde olası etik dışı durumlar belirlenirken ilk adım, otonom sistemin sınırlarının tespiti ve hangi durumlar için kullanılacağını değerlendirilmesidir. Otonom sistemi oluşturan alt birimler, yapay zekâ sisteminin kullanım sınırlarının belirlenmesi için sistemin alacağı görevler, donanım, yazılım durumu ile elde bulundurulması planlanan veriler ve kullanıcılar beraber değerlendirilmiştir. Kullanım senaryosu oluşturularak, olası etik dışı davranışların daraltılması hedeflenmiştir. Bu durum aynı zamanda alınabilecek tedbirlerin belirlenmesi için gerekli olan birtakım önemli bilgileri de sağlayacaktır. Modelde ikinci adımda iç veya dış faktörler tarafından sistemde etik ihlale neden olabilen tehlikeler ve bu tehlikeler tarafından istismar edilebilecek hassasiyetler belirlenmiştir. Potansiyel tehdit kaynakları ve hassasiyetler belirlenerek bu adım tamamlanmıştır. Üçüncü adım etik olasılıkların kontrolü ve etik sistemin amaçları ile uyumlu hale getirilmesidir. Belirlenen etik riskler ve bu risklerin üstesinden gelmek için gerekli olan güvenlik gereksinimleri bu aşamada geliştirilir. Olasılıklar belirlenirken etik ihlale yol açabilecek tasarımların bu aşamada değerlendirilmesi uygun bulunmuştur. Otonom sistemin belirlenen durumlar karşısında alacağı tedbirleri içeren etik ihlalleri önleme sürecinin sistem kullanılmaya başlamadan önce tamamlanması esastır. Bu aşamada önceden belirlenen fayda sağlama ve iyilik, zarar vermeme, adalet ve açıklık ilkeleri değerlendirilir. Dördüncü adımda etki analizi ve uygulama oluşturma basamakları ayrıntılı olarak ele alınmıştır. Etik ihlallerin potansiyel etkilerinin tespiti ve

analizi yapılarak ihlalin sürece olan olası etkisinin değerlendirilmesi ve önlenmesi bu aşamada algoritmaya aktarılır. Modelin beşinci adımda risklerin ortaya çıkması ile devreye giren birtakım tespit yöntemleri örneklerle anlatılmıştır. Kullanıcı tarafından iki tespit yöntemi arasında seçim yapılabilmesi için verilen ölçütlere göre değerlendirme yapılmalı ve geçerli yöntem uygulanmalıdır. Son adımda risklerin sonuçlarının değerlendirilmesi; hatalı veri, bilgi, yazılımın sistem dışına çıkarılması aşamasında yapılacak işlemler belirlenmiştir. İmha edilecek veya değiştirilecek bileşenler dışında kalan veriler uygun şekilde işlenerek tekrar kullanılabilir.





Şekil 4.8. Etik Kontrol Modeli (Stoneburner ve diğerleri, 2002 ve Nieves ve diğerleri 2017'den uyarlanmıştır.)

5. SONUÇ VE ÖNERİLER

Araştırmada “Asilomar Yapay Zekâ Prensipleri”, Yapay Zekânın Sosyal Olarak Sorumlu Gelişimi Projesi kapsamında Montreal Üniversitesi tarafından geliştirilen “Sorumlu Yapay Zekâ için Montreal Deklarasyonu” (Montreal Declaration, 2017), Otonom ve Akıllı Sistemlerle İnsan Refahını Önceliklendirme Vizyonu tarafından geliştirilen IEEE Prensipleri ve Floridi ve diğerleri (2018) tarafından yapılan kapsamlı çalışmalar ile ABD Siber Güvenliği Geliştirme Yasasında (Public Law, 2014) onaylanan ve özel sektör firmaları ile kritik altyapı sunan işletmeler için kullanılması gerekli görülen ve Stoneburner ve diğerleri (2002) tarafından oluşturulan Risk Yönetim Kılavuzu ve Nieves ve diğerleri (2017) tarafından oluşturulan Bilgi Güvenliği El Kitabı kullanılmıştır. Bu iki kaynağın birleştirilmesi, yapay zekâda etik kontrol modeli oluşturulmasında daha kapsamlı bir yaklaşımın benimsenmesine imkân tanımıştır.

Otonom sistemlerdeki etik ihlallerin toplumsal bir güvenlik açığına neden olması, meydana gelme olasılığı ve etkisi dikkate alındığında günümüz şartlarında beklenen bir durumdur. Temel olarak etik kontrol; güvenlik açıklarına karşı etik standartları belirleme, değerlendirme ve etik dışı davranışları asgari düzeye indirme sürecidir (Treviño ve Weaver, 2001). Modelin tasarlanmasındaki amaç; otonom sistem modellerken ya da kullanırken karşılaşılabilen etik riskleri tespit etmek ve etik dışı davranışların azaltılması için kullanıcılara işlem basamaklarını içeren bir yol haritası sunmaktır. Modelin nihai hedefi, yapay zekâ aktörlerinin ve otonom sistemlerin, toplumsal uygulamalarda etik dışı davranışlarda bulunmasının engellenmesidir. Yukarıda oluşturulan etik denetim modeli kullanarak; kişisel ve toplumsal verileri depolayan, işleyen ve nakleden yapay zekâ sistemleri, sürdürülebilir refah artışı ve toplumsal fayda sağlamaya yönlenebilecektir.

Model ile hem otonom sistemlerin yol açabileceği etik sorunlar ele alınmış hem de bu sorunlara yönelik çözümler geliştirmek için algoritma belirlenmiştir. Bu algoritma ile yapay zekâ teknolojisinin geliştirilmesi ve kullanımı sırasında ortaya çıkabilecek etik riskler azaltılabilir ve toplumun yapay zekâyı olan güveni artırılabilir. Hızla gelişmekte ve hayatımızın birçok alanında kullanılmaya başlayan teknolojilerde etik kontrollerin yapılmaması durumunda ortaya çıkabilecek olumsuz sonuçların önlenmesi, teknolojinin insan hayatına katkı sağlayabilmesi açısından büyük önem taşımaktadır (Floridi ve Cows, 2018).

2019). Çalışmada yapay zekâ etiği konusunda yapılan ayrıntılı analiz sayesinde, yapay zekâ uygulamalarında kullanılması gereken etik kurallar tespit edilerek en yaygın dört ilke belirlenmiştir. Çalışmada belirlenen bu dört temel etik ilkenin uygulanmasına ilişkin öneriler sunulmuş ve akabinde otonom sistemlerin yol açabileceği etik sorunların giderilmesi için bir etik kontrol modeli ortaya koyulmuştur.

Ortaya koyulan model ile yapay zekâ uygulamalarının geliştirilmesi ve kullanımında açıklık, hesap verebilirlik ve eşitliği sağlamak için alınacak tedbirlerin önemi vurgulanmış ve otonom sistemlerde kişisel verilerin gizliliğinin sağlanması için bilginin işlenmesi sürecinde kullanılabilecek yaklaşımlar tespit edilmiştir. Araştırmanın son bölümü kullanıcılara otonom sistem üzerinde daha fazla kontrol sağlayabilecek bir algorithmadan oluşmaktadır. Çalışma temel olarak yapay zekâ ve etik ile ilgili sistemik ilişkilerin ortaya koyulması ve risk faktörlerinin belirlenerek çözüm için model öne sürmesi gibi önemli katkı yapmasının dışında, diğer alanlar ile bütünleştirilerek geliştirilebilecek bir yapıdadır. Yapay zekânın topluma yarar getirmesi ve bunu yaparken ikincil zararlara sebep olmaması için, bu sistemlerin insana ait olan ve toplum içerisinde kabul gören değerleri, inanışları ve algıları dikkate alması gerekir. Öyle ki etik kuralların zaman içerisinde değişebileceği de göz önüne alındığında yapay zekânın kendisini yeni kurallara uyumlama yeteneğinin de olması gerekir. Bu noktada toplumun ahlaki değerlerinin yapay zekânın gelişme sürecine dâhil edilmesi, yaşça birbirinden farklı insanlar arasında zaman zaman görülen ve istenmeyen bir durum meydana getiren kuşak çatışması benzeri bir durumun ortaya çıkmasına engel olacaktır.

Çalışmada sunulan model önerisiyle yapay zekâyâ sahip sistemlerin; insanlara zarar vermeden, adalet ve hakkaniyet prensiplerine uygun olarak kullanımının artması ve sistemlerin toplumsal fayda sağlamaya yönlendirilebilmesi hususlarında alana katkı sağlandığı değerlendirilmektedir. Ortaya koyulan etik kontrol modeli; yapay zekâyâ sahip sistemlerin geliştirilmesinde ve kullanımında, yapay zekâ geliştiricilerinin ve kullanıcılarının etik standartlara uygun hareket etmelerine kılavuz olabilir. Önceden tespit edilecek etik sorunların daha az süre ve daha az maliyetle çözümlenmesiyle yapay zekâ ve etik alanındaki gelişmeler hız kazanabilir. Sunulan etik model, bu alanda çalışan diğer araştırmacılara farklı bir bakış sunması açısından da önemlidir. Alandaki araştırmaların artması ile otonom sistemlerin geliştirilmesi ve toplum tarafından daha fazla kabul görmesi topluma ve insana daha fazla fayda sağlayacaktır.

Model oluşturulurken etik hususların uygulanması hususunda esnek bir yöntem kullanılmış ve algoritma oluşturulmaya başlamadan önce mümkün olan en fazla faydalı düşüncenin sisteme dâhil edilmesi hedeflenmiştir. Tasarım araç ve yöntemleri belirlerken, etik ilkelerin iş ve kullanım durumlarına uygun şekilde sürece aktarılması modelin hata riskini azaltacağı değerlendirilmiştir. Bu yaklaşım ile dijital ortamda etik yapay zekâ unsurlarını belirleyerek, farklı toplumlarda da kullanabilecek ortak bir algoritma oluşturulması planlanmıştır. Modelde önerilen adımlar uygulayanların değerlendirmeleri ışığında genişletilebilir, kısaltılabilir ya da yeni eklemeler yapılabilir.

Model kullanılarak kritik bilgiler işleyen, depolayan ve taşıyan yapay zekâyâ sahip sistemlerin neden olduğu etik riskler, yapılacak sürekli değerlendirmeler ile en aza indirilebilir. Sonuç olarak yapay zekâ teknolojisinin yeni ve aynı zamanda güçlü bir teknoloji olduğu yadsınamaz bir gerçektir. Bu teknolojinin yakın zamanda meydana getirebileceği toplum yaşayışına ve ahlakına aykırı sonuçların iyi analiz edilememesi, toplumsal düzen, eşitlik ve adalet gibi temel kavramlar için dönüşü olmayan sonuçlar doğurabileceği göz önünde bulundurulmalıdır. Toplum psikolojisi, ahlak felsefesi, sosyoloji gibi bilimlerin de yardımı ile yapay zekâ tasarım sürecinde otonom sisteme, semantik analiz yetenekleri ile durumsal farkındalık kazandırılabilir. Önümüzdeki dönemde yeni tasarlanan otonom sistemlerin toplumsal ilerleme sürecine katkı sağlaması bu yolla sağlanabilir.

KAYNAKLAR

- AI Now Institute, (2017). *AI Now Report 2017*. New York: Institute Publish.
- Al Lawati, H.A.H., Sayed, R.T., ve Caldwell, C. (2019). Transformative Ethics and Moving Toward Greatness: Problems and Realities. *Journal of Values-Based Leadership*. XII, 1.
- Alexandre, S. (2017). Artificial Intelligence: A new digital ecosystem. *Journal of Innovation & Knowledge*, 2(2), 64-71. <https://doi.org/10.1016/j.jik.2017.02.001>
- Anderson, M., ve Anderson, S., (2006). Issue on Machine Ethics, *IEEE Intelligent Systems*. 21(4).
- Anderson, M., Susan L.A. and Chris A. (2005). *Towards Machine Ethics: Implementing Two Action-Based Ethical Theories*. 2005 AAAI Fall Symposium on Machine Ethics.
- Anderson, S.L. (2008). Asimov's Three Laws of Robotics and Machine Metaethics, *AI and Society*, vol. 22, no. 4, 477-493.
- Arslan, M. ve Berkman, A. Ü. (2009). Dünyada ve Türkiye'de İş Etiği ve Etik Yönetimi, *TÜSIAD*, 2009, 06-492.
- Atalay, M. ve Çelik, E. (2017). Büyük veri analizinde yapay zekâ ve makine öğrenmesi uygulamaları. *Mehmet Akif Ersoy Üniversitesi Sosyal Bilimler Enstitüsü Dergisi*, 9(22): 155-172.
- Avrupa Bilim ve Yeni Teknolojilerde Etik Grubu. (2018). *Statement on Artificial Intelligence, robotics and 'autonomous' systems*, Belçika: EU Publish.
- Avrupa Parlamentosu. (2020). *Report on a comprehensive European Union policy on artificial intelligence*. Belçika: EU Publish. doi:10.12849/210127
- Awad, E., Dsouza, S., ve Kim, R. (2018). The Moral Machine Experiment. *Nature*, 563, 59-64.
- Barrat, J. (2015). *Our final invention: Artificial intelligence and the end of the human era*. New York: St. Martin's Press.
- Bartfield, C., ve Datta, A. (2018). *Artificial intelligence and human rights*. In *The Cambridge Handbook of Human Rights and Information Technology (187-207)*. Cambridge: Cambridge University Press.
- Başoğlu, B. ve Bulut, M. (2017). Kısa dönem elektrik talep tahminleri için yapay sinir ağları ve uzman sistemler tabanlı hibrit sistem getirilmesi. *Journal of Faculty of Engineering and Architecture of Gazi University*, 32(2): 575-583.
- Baum, S. D. (2020). Social choice ethics in artificial intelligence. *AI and SOCIETY*, 35(1), 165-176.

- Bawden, D. ve Robinson, L. (2012). *Introduction to Information Science*. London: Facet. 351.
- Beauchamp, T.L. ve J.F. Childress. (2012). *Principles of biomedical ethics*. Oxford: Oxford University Press.
- Bellman, R.E. (1978). *An Introduction to Artificial Intelligence: Can Computers Think*. Boston: Boyd and Fraser Publish.
- Béranger, J. (2021). *Societal Responsibility of Artificial Intelligence*. London: ISTE Ltd.
- Berendt, B. (2019). AI for the Common Good. Pitfalls, challenges, and ethics pen-testing. *Paladyn, Journal of Behavioral Robotics*, 10(1), 44-65. <https://doi.org/10.1515/pjbr-2019-0004>.
- Berners-Lee, T. (1998). What the Semantic Web can Represent. *Design Issues*, 1, 120.
- Bertolino, A. (2007). Software testing research: Achievements, challenges, dreams. *Future of Software Engineering*, 2007, 85-103.
- Bolay, S. H. (2013). *Felsefe Doktrinleri ve Terimleri Sözlüğü*, Nobel Yayın Dağıtım, Ankara.
- Bostrom, N. (2016). *Superintelligence: paths, dangers, strategies*. Oxford: Oxford University Press,.
- Brooks, R.A. (2002). *Flesh and machines: how robots will change us*. New York: Pantheon Books.
- Buttarelli, G. (2018). The EU General Data Protection Regulation: Towards a new global standard?. *Computer Law & Security Review*, 34(2), 274-287. <https://doi.org/10.1016/j.clsr.2017.12.005>.
- CAHAI. (2020). *Consultative Committee of the Convention on the Protection of Individuals with regard to Automatic Processing of Personal Data. Guidelines for trustworthy AI*. Belçika: Council of Europe.
- Calayır, G. N. ve Kabak, M. (2021). Bakım için makine öğrenme tekniklerinin analizi ve bir uygulama . *Journal of Turkish Operations Management* , 5 (1) , 662-675.
- Caldwell, D. F., & Clapham, S. E. (2003). Ethics and entrepreneurship: Dead or alive?. *Journal of Business Venturing*, 18(2), 271-288.
- Caliskan, Aylin & Bryson, Joanna & Narayanan, Arvind. (2017). Semantics derived automatically from language corpora contain human-like biases. *Science*. 356. 183-186.
- Callicott, J. B. (2014). *Thinking Like A Planet : The Land Ethic And The Earth Ethic*. Oxford: Oxford University Press.

- Canada Parliament. (2000). *The Personal Information Protection and Electronic Documents Act*. Kanada: Government Press.
- Canbay, P. ve Demircioğlu, Z. (2021). Endüstri 5.0'a Doğru: Zeki Otonom Sistemlerde Etik ve Ahlaki Sorumluluklar . *AJIT-e: Academic Journal of Information Technology* , 12 (45) , 106-123.
- Canbay, Y. ve Sağiroğlu, Ş. (2020). *Büyük Veri ve Yapay Zekâ Uygulamalarında Mahremiyet ve Etik. Yapay Zekâ ve Büyük Veri: Teknolojiler, Yaklaşımlar ve Uygulamalar*, Türkiye: Nobel, 157-186,.
- Capurro, R. (2008). Information Ethics for and from Africa. *Journal of the American Society for Information Science and Technology*, 59 (7), s. 1162-70.
- Caşın, M., Al, D. , Başkır, N. D. (2021). Yapay zekâ ve robotların eylemlerinden kaynaklanan cezai sorumluluk sorunu. *Ankara Barosu Dergisi*, 79: 1-73
- Cellan-Jones, R. (2018). Will AI Kill or Save Humankind? İngiltere: BBC.
- Cerka, P., Jurgita, G. ve Gintarė, S. (2015). Liability for damages caused by artificial intelligence. *Computer Law ve Security Review*, 31, 376- 389.
- Coeckelbergh, M. (2020). *AI ethics*. The MIT Press, Cambridge: MA.
- Committee on Artificial Intelligence CAHAI. (2020). *Feasibility Study 2020*. France: Committee Press.
- Cookson, C. (2018). *Artificial Intelligence faces public backlash, warns scientist*. New York: Financial Times.
- Cowls, J., King, T., Taddeo, M., ve Floridi, L. (2019). *Designing AI for Social Good: Seven Essential Factors*. SSRN. <http://dx.doi.org/10.2139/ssrn.3388669>
- Creemers, R., ve Webster, G. (2021). *Personal Information Protection Law of the People's Republic of China. The Digichina Project*, Stanford University Press.
- Crevier, D. (1993). *AI: The tumultuous history of the search for artificial intelligence*. New York: BasicBooks.
- Doğan, M. (2018). Yapay Zekâ ve Savaş. In Savaş ve Barış Bildiri Kitabı. Bursa: Sentez Yayıncılık.
- Doğan, K. (2019). Sürücüsüz Araçlar, Robotik Cerrahi, Endüstriyel Robotlar ve Cezai Sorumluluk. *Dokuz Eylül Üniversitesi Hukuk Fakültesi Dergisi*, Özel Sayı 2019, 3223.
- Dönerçark, M., & Tecim, V. (2020). Kurumsal Karar Destek Sistemlerinde Yapay Zekâ Kullanımı: Tasarım ve Uygulama. *Yönetim Bilişim Sistemleri Dergisi*, 6(2), 77-103.

- Doshi-Velez, F., ve Kortz, M. (2017). *Accountability of AI Under the Law: The Role of Explanation*. Cambridge: Harvard University.
- Dustin, E. (2003). Automated security assessment program pitfalls. *IEEE Security & Privacy*, 1(5), 20-25.
- Erdal, H. (2018). Yapay Zekâ Teknikleri ve Uzman Sistemleri'nin Karasal Akıllı Ulaşım Sistemlerinin Denetiminde Kullanımı. *Akıllı Ulaşım Sistemleri ve Uygulamaları Dergisi*, 1(1), 32-39.
- Ereken, Ö. F. (2021). *Yapay zekâ tabanlı personel seçim sistemi uygulaması*. İzmir Dokuz Eylül Üniversitesi Sosyal Bilimler Enstitüsü, İzmir.
- Ersoy, Ç. (2018). *Robotlar, Yapay Zekâ ve Hukuk*. İstanbul: On İki Levha Yayıncılık.
- European Commission. (2019). *Ethics Guidelines For Trustworthy AI*. Belçika: EU Publish.
- Floridi, L. (2013). *The Ethics of Information*. Oxford: Oxford University Press.
- Floridi, L. (2021). *Ethics, Governance, and Policies in Artificial Intelligence*. Oxford: Oxford University Press.
- Floridi, L. ve Cows, J. (2019). A unified framework of five principles for AI in society. *Harvard Data Science Review*, 1(1).
- Floridi, L., Cows, J., Beltrametti, M., Chatila, R., Chazerand, P., Dignum, V., Luetge, C., Madelin, R., Pagallo, U., Rossi, F., Schafer, B., Valcke, P., & Vayena, E. (2018). AI4People—An ethical framework for a good AI society: Opportunities, risks, principles, and recommendations. *Minds and Machines*, 28(4), 689–707. <https://doi.org/10.1007/s11023-018-9482-5>.
- Ford, M. (2016). *The rise of robots: Technology and the threat of a jobless future*. New York: Basic Books.
- Güney, İ., & Özkaya, E. (2020). Yapay Zekâ Uygulamaları ve Lojistik Sektörüne Etkileri. *Ege Akademik Bakış*, 20(1), 103-113.
- Güngör, M. (2015). *Ulusal bilgi güvenliği: Strateji ve kurumsal yapılanma* [Uzmanlık tezi, Bilgi Toplumu Dairesi Başkanlığı].
- Güven, H., & Ayvaz Güven, E. T. (2023). Yapay Zekâ Uygulamalarının E-Ticarete Kullanımı. *International Journal of Management and Administration*, 7(13), 69-94. DOI: 10.29064/ijma.1194949
- Harwell, D. (2020). *Managers turn to surveillance software, always-on webcams to ensure employees are (really) working from home*. New York: The Washington Post.
- Haugeland, J. (1985). *Artificial Intelligence: The very idea*. MIT Press.

- Henkoğlu, T., & Yılmaz, B. (2013). Avrupa Birliği Bilgi Güvenliği Politikaları. *Türk Kütüphaneciliği*, 27(3), 451-471.
- Higgins, A., Hodgson, P., & Travis, D. (2014). *Bright ideas for user experience researchers*. Ebook.
- Hukuk Çalıştay Raporu. (2019). Yapay Zekâ Çağında Hukuk: *İstanbul, Ankara ve İzmir Baroları Çalıştay Raporu*. İzmir.
- İnce, H., İmamoğlu, S. E., & İmamoğlu, S. Z. (2021). Yapay zekâ uygulamalarının karar verme üzerine etkileri: Kavramsal bir çalışma. *International Review of Economics and Management*, 9(1), 50-63. <https://doi.org/10.18825/iremjournal.866432>
- İnternet: Asilomar AI Principles (2017). *Principles of AI*. URL: <https://futureoflife.org/aiprinciples>. Son Erişim Tarihi: 29.09.2018.
- İnternet: Evdeki gizli tehlike: Biri sizi gözetliyor olabilir. (2021, Kasım). *Postaktuel*. URL: <https://www.postaktuel.com/90532-2/> Son Erişim Tarihi: 29.03. 2022.
- İnternet: Federal Trade Commision (FTC). *Business Guidance*. URL: <https://www.ftc.gov/business-guidance/small-businesses/cybersecurity/nist-framework>. Son Erişim Tarihi: 28 Ocak 2023
- İnternet: Initiative on Ethics of Autonomous and Intelligent Systems. (2017). *Ethics in Action*, URL: <https://ethicsinaction.ieee.org>, Son Erişim Tarihi: 25 Kasım 2020
- İnternet: İyilik için Yapay Zeka Zirvesi (2019). *AI for Good Global Summit (AIGGS)*. <https://aiforgood.itu.int/>. Son Erişim Tarihi: 28 Kasım 2022.
- İnternet: Shaping the Ethical Dimensions of Smart Information Systems. (2022, Nisan). *SHERPA*. URL: <https://www.project-sherpa.eu/> Son Erişim Tarihi: 28.04. 2022.
- İnternet: TechTarget (2023). 5 Key Elements of a Modern Cybersecurity Framework. *TechT*. URL: <https://www.techtarget.com/searchsoftwarequality/definition/NIST>. Son Erişim Tarihi: 28 Ocak 2023
- İren, E., & Can, Ö. (2017). Bilgi Sistemlerinde Güncel Güvenlik Problemleri ve Önerilen Çözümler. *TÜBAV Bilim Dergisi*, 10(2), 27-42. <https://doi.org/10.5505/tbd.v10i2.57127>
- Isaak, J., & Hanna, M. J. (2018). User data privacy: Facebook, Cambridge *Analytica*, and privacy protection. *Computer*, 51(9), 56–59. <https://doi.org/10.1109/MC.2018.3191268>.
- Jobin, A., Ienca, M., & Vayena, E. (2019). The global landscape of AI ethics guidelines. *Nature Machine Intelligence*, 1(9), 389–399. <https://doi.org/10.1038/s42256-019-0088-2>
- Kaplan, A., ve Kaplan, N. (2020). Yeni İletişim Teknolojileri ve Yapay Zekâ Uygulamaları Bağlamında İhtiyaç Olgusunun Değişimi Ve Toplumsal Dönüşüm [Change in the

- concept of need and social transformation in the context of new communication technologies and AI applications]. *The Journal of Social Sciences*, 46(1), 178-193. <https://doi.org/10.29228/sobider.43976>
- Khan, S. A. (2018). Impact of technology on journalism. *Journalism and Mass Communication*, 8(4), 13-23. <https://doi.org/10.1177/2347083518816727>
- Kızıl, S. (2021). *Disiplinlerarası Girişimcilik ve Yenilikçilik Çalışmaları*. Ankara: Seçkin Yayıncılık. E-ISBN: 978-625-8065-93-0.
- Kizza, J. M. (2016). *Ethical and Social Issues in The Information Age* (7th ed.). Springer.
- Köklü, K. (2018). İş Analizi, İş Analistliği ve İş Zekâsı [Job analysis, business analysis and business intelligence]. *Lectio Socialis*, 2(2), 121-142. <https://doi.org/10.30935/sdusbed.437831>
- Korienek, G., & Uzgalis, W. (2002). Adaptable robots. *Metaphilosophy*, 33(1-2), 83-97. <https://doi.org/10.1111/1467-9973.00210>.
- Kurbanoglu, S. (1992). Uzman sistemler. *Türk Kütüphaneciliği*, 6(4), 189-193.
- Kurzweil, R. (1990). *The age of intelligent machines*. Cambridge: MIT Press.
- Kurzweil, R. (2006). *The singularity is near*. Cambridge: Gerald Duckworth.
- Lake, B. M., Ullman, T. D., Tenenbaum, J. B., & Gershman, S. J. (2017). Building machines that learn and think like people. *Behavioral and Brain Sciences*, 40, e253. doi: 10.1017/S0140525X16001837.
- Liao, S. M. (2020). *Ethics of artificial intelligence*. Oxford University Press.
- Luger, G. F., & Stubblefield, W. A. (1993). *Artificial intelligence*. Cummings Publishing.
- Ma, J., & Liu, Q. (2019). Research on the ethical problems and governance strategies of artificial intelligence. *Frontiers of Engineering Management*, 6(3), 337-346. doi: 10.1007/s42524-019-0051-1
- Maalej, Z., & Ayachi, R. (2021). Understanding AI for Sustainable Development: A Systematic Literature Review. *Journal of Artificial Intelligence and Sustainable Development*, 2(1), 1-13. <https://doi.org/10.46681/jaisd.2021.01.01>
- Maas, M. (2019). *Ethically Aligned Design: A Vision for Prioritizing Human Well-being with Artificial Intelligence and Autonomous Systems*. IEEE Standard Association.
- Maşa, E. ve Karadağ, E. (2019). Sistematiik İnceleme. *İstanbul Üniversitesi İktisat Fakültesi Ekonometri ve İstatistik Dergisi*, 31, 85-112.
- McCarthy, J. (1955). *A proposal for the Dartmouth summer research project on artificial intelligence*. Dartmouth College.

- McCarthy, J. (2007). *What is Artificial Intelligence*. Stanford University: Computer Science Department.
- McCorduck, P. (2004). *Machines who think: A personal inquiry into the history and prospects of artificial intelligence*. Systems. Federal Computer Security Managers' Forum Working Group.
- McCumber, J. (2005). *Assessing and managing security risk in IT systems*. CRC Press.
- McSorley, K. (2003). The secular salvation story of the digital divide. *Ethics and Information Technology*, 5(2), 75-87. <https://doi.org/10.1023/A:1024946302065>
- Miller, K. W., & Coldicott, R. (2019). Ethics and data science: A survey of professional coders and users. *Journal of Business Ethics*, 160(3), 657-676.
- Mockenhaupt, R. (2021). Machine learning for materials science. *Current Opinion in Materials Science*, 27, 128-134.
- Montreal. (2017). *Declaration for a Responsible Development of Artificial Intelligence*. Kanada: Montreal Press.
- Müller, V. C. (2020). *Ethics of artificial intelligence and robotics*. In E. N. Zalta (Ed.), *The Stanford Encyclopedia of Philosophy* (Winter 2020 ed.). Stanford University.
- Nabiyev, V. V. (2005). *Artificial intelligence and intelligent systems: theory and applications*. World Scientific.
- National Institute of Standards and Technology (NIST Law). (2018). *Framework for improving critical infrastructure cybersecurity*. National Institute of Standards and Technology.
- Nieles, M., Dempsey, K. ve Yan Pillitteri, V., (2017). *An introduction to computer security: The NIST handbook (NIST SP 800-12)*. National Institute of Standards and Technology.
- Nilsson, N. J. (1998). *Artificial intelligence: A new synthesis*. Morgan Kaufmann Press.
- NIST Special Publication 800-18. (1998). *Guide for developing security plans for information technology systems*. National Institute of Standards and Technology.
- OECD. (2019). *Recommendation of the Council on artificial intelligence*. OECD Legal Instruments.
- Örselli, E. (2010). Türkiye’de toplumsal ve yönetsel etik değerler ile ikilemler: Uygulamalı bir araştırma. (Doktora tezi, Selçuk Üniversitesi Sosyal Bilimler Enstitüsü, 2010).
- Osoba, O., ve Welser, W. I. V. (2017). *An intelligence in our image: The risks of bias and errors in artificial intelligence*. Rand Corporation.

- Polanyi, K. (1944). *The great transformation: The political and economics of our time*. New York: Farrar and Rinehart.
- Robbins, S. (2019). A Misdirected Principle with a Catch: Explicability for AI. *Minds and Machines*, 29, 495–514. <https://doi.org/10.1007/s11023-019-09509-3>
- Robbins, S., & Henschke, A. (2017). The value of transparency: Bulk data and authoritarianism. *Surveillance and Society*, 15(3/4), 582–589. <https://doi.org/10.24908/ss.v15i3/4.6606>.
- Rodrigues, J. F. (2009). Ethics of emerging technologies: scientific facts and moral challenges. *Journal of Science and Technology Studies*, 8(1), 23-33.
- Rodrigues, R., Panagiotopoulos, A., Lundgren, B., S Lahlé Shaelou ve A Grant (2019). *Regulatory options for AI and big data*. SHERPA.
- Romansky, R. (2015). Social Computing and Digital Privacy. *Communication and Cognition*, 48(34), 65-82.
- Royackers, L., Timmer, J., Kool, L., & van Est, R. (2018). Societal and ethical issues of digitization. *Ethics and Information Technology*, 20(2), 127–142.
- Russell, S. J., & Norvig, P. (2003). *Artificial Intelligence: A Modern Approach (2nd ed.)*. New Jersey: Prentice Hall.
- Samuel, A. L. (1959). Some Studies in Machine Learning Using the Game of Checkers. *IBM Journal of Research and Development*, 3(3), 210-229.
- Santiago, O. C. (2019). *Taking Business Ethics Seriously*. Cambridge: Scholars Publishing.
- Santiago, R. (2019). *The Ethics of Artificial Intelligence*. The Routledge Companion to Digital Ethics, 159-169.
- Say, C. (2018). *50 Soruda Yapay Zekâ*. İstanbul: Bilim ve Gelecek Kitaplığı.
- Schmidt, F. L. (2012). Select on Intelligence. *Handbook of Principles of Organizational Behavior*, 1, 5-15.
- Sharkey, N., & Sharkey, A. (2010). The crying shame of robot nannies: An ethical appraisal. *Interaction Studies*, 11(2), 161-190.
- Silva, N., Xavier, B., & Nathansohn, B. M. (2018). Bilgi Bilimi Alanında Yapay Zekada Bilimsel Üretimin Analizi. *Ulusal Bilgi Bilimleri Araştırma Çalışmaları*. <http://www.brapci.inf.br/index.php/res/v/103730>.
- Simborg, D. (2017). *What comes after Homo Sapiens?*. CA: DWS Publishing.
- Şimşek, N. (2007). *Akıllı Zekâ*. Ankara: Asil Yayınları.
- Singer, P. (2011). *Practical Ethics (3rd ed.)*. Cambridge: University Press.

- Smith, J. D., & Johnson, K. L. (2018, Haziran). Artificial Intelligence in Predictive Maintenance. *IEEE Conference on Prognostics and Health Management*, 25-30. San Diego, CA: IEEE.
- Spinello, R. A., & Tavani, H. T. (2004). Intellectual property rights in a networked world: Theory and practice. *Information Science Publishing*, Hershey, PA.
- Stahl, B. C. (2021). Artificial Intelligence for a Better Future. NYC: SpringerBriefs in Research and Innovation Governance. https://doi.org/10.1007/978-3-030-69978-9_4
- Stahl, B.C. ve Wright, D. (2018). *Ethics and Information Technology: A Case-Based Approach to a Health Care System in Transition*. Cham: Springer.
- Stoneburner, G., Goguen, A., & Feringa, A. (2002). *Risk Management Guide for Information Technology Systems, Special Publication (NIST SP)*, National Institute of Standards and Technology, Gaithersburg: MD.
- Sucu, E. (2020). Yapay Zekâ ve Makine Öğrenmesi. *Journal of Computer and Education Research*, 8(15), 252-265.
- Taddeo, M., & Floridi, L. (2018). How AI can be a force for good. *Science*, 361(6404), 751-752.
- Taşdemir, D. D. (2020). *Kültürel Zekâ, Duygusal Emek ve İş Performansı Arasındaki İlişkilerin Çok Kültürlü Bir Ortamda İncelenmesi*. Yayımlanmamış doktora tezi, Kocaeli Üniversitesi Sosyal Bilimler Enstitüsü, 2020, Kocaeli.
- Tonybee, A. J. (1946). *A Theory Of History I*. Oxford: University Press.
- Treviño, L. K., & Weaver, G. R. (2001). Organizational Justice and Ethics Program ‘Follow-Through’: Influences on Employees’ Harmful and Helpful Behavior. *Business Ethics Quarterly*, 11, 651–671.
- Türk Dil Kurumu. (2022). *TDK Sözlük*. Ankara: Türk Dil Kurumu
- Ülgen, H., & Mirze, K. S. (2004). *İşletmelerde Stratejik Yönetim*. İstanbul: Literatür Yayınları.
- Vakkuri, V., Koivisto, R., & Helander, N. (2019). Ethical design methods for industrial AI. *Journal of Business Ethics*, 160(4), 1015-1032.
- Velasquez, M. G. (2017). *Business Ethics Concepts and Cases*. New York: Pearson Education, pp. 122-147.
- Whittaker, J. (2000). Choosing a method for computing inter-coder reliability. *Advances in Health Sciences Education*, 5(1), 93-113.
- Winner, L. (1980). Do artifacts have politics?, *Daedalus*, 109(1), 121-136.

- Wolkenstein, A. (2018). Artificial Intelligence and Accountability: Using Regulatory Sandboxes to Foster Responsible Innovation. *Journal of Law and Policy for the Information Society*, 14(2), 247-275.
- Yatkın, A. (2009). Ethical change over time. *Journal of Business Ethics*, 89(4), 617-627.
- Yereli, A.B., & Şahin, I. (2019). Yapay Zekanın Ekonomi, Toplumsal Refah ve Vergilendirme Üzerindeki Etkisi. *Vergi Dünyası*. YIL 39, SAYI 459.
- Yick, J., Sharma, M., Lopez, D., & Park, J. (2016). Smart home security: A systematic literature review. *Smart Homecare Technology and TeleHealth*, 4, 1-14.



ÖZGEÇMİŞ

Adı-Soyadı

: D. Devrim TAŞDEMİR

Bilimsel Faaliyetler (Yayınlar, Bildiriler, Katıldığı Projeler)

1. Taşdemir, D.D. (2021). *Etik Girişimcilik*. Kızıl, S. (Ed.) İstanbul: Efe Publish, 168.
2. Taşdemir, D.D. (2021). Çok Kültürlü Çalışma Ortamında Kültürel Zekâ ve Duygusal Emek. İstanbul: Efe Publish, 220.
3. Yıkılmaz, I., Taşdemir, D. D., ve Gündüz Cekmecelioglu, H. (2021). The assessment of the intermediation role of emotional labor dimensions in the relationship between cultural intelligence and individual work performance. *Business and Economics Research Journal*, 12(1), 157- 172.
4. Taşdemir, D.D. (2020). Kültürel Zekâ, Duygusal Emek Ve İş Performansı Arasındaki İlişkilerin Çok Kültürlü Bir Ortamda İncelenmesi, Yayınlanmamış Doktora Tezi, Kocaeli Üniversitesi Sosyal Bilimler Enstitüsü, Kocaeli.
5. Yamen, M., Taşdemir, D.D. ve Demirkaya, H. (2017). Örgütsel Öğrenmenin Örgütsel Değişim Üzerindeki Etkilerine Yönelik Bir Araştırma. *Eurasian Academy of Sciences*, 2017, 74-86.
6. Taşdemir, D.D., Çekmecelioglu H. ve Yıkılmaz, İ. (2019). Çok Kültürlü Ortamda Çalışanların Kültürel Zekâ Düzeylerinin Duygusal Emek Gösterilerine Etkisi. *OPUS International Journal of Society Studies*, 2019, 719-740.
7. Taşdemir, D.D., Yamen, M. ve Demirkaya, H. (2017). Eğitim Yoluyla Liderlik Özelliklerinin Geliştirilmesi: Türk Ordusu'nda Asteğmen Adaylarında bir Uygulama. *Avrasya Bilim Akademisi Avrasya İşletme ve Ekonomi Dergisi*, 2017, 146-157.
8. Taşdemir, D.D. (2015). Piyade Okulu'nda Eğitim Alan Asteğmen Adaylarının Askeri Liderlik Özelliklerinin Geliştirilebilirlik Durumları Üzerine Bir Araştırma, Yayınlanmamış Yüksek Lisans Tezi, Marmara Üniversitesi Sosyal Bilimler Enstitüsü, İstanbul.