

T.C.
SELÇUK ÜNİVERSİTESİ
SOSYAL BİLİMLER ENSTİTÜSÜ
FELSEFE ANABİLİM DALI
FELSEFE BİLİM DALI

HUBERT L. DREYFUS VE YAPAY ZEKÂ
ELEŞTİRİLERİ

Emine YAKIŞ

YÜKSEK LİSANS TEZİ

Danışman

Prof. Dr. Murat ARICI

Konya-2024

Bu çalışma, 14.11.2024 tarihinde ařağıdaki jüri tarafından FELSEFE Anabilim Dalı FELSEFE Bilim Dalı Programında Yüksek Lisans Tezi olarak kabul edilmiştir.

Savunma Tez Jürisi

Prof. Dr. Murat ARICI

Selçuk Üniversitesi
Edebiyat Fakültesi

Prof. Dr. Cengiz İskender ÖZKAN

Selçuk Üniversitesi
Edebiyat Fakültesi

Doç. Dr. Nihat DURMAZ

Selçuk Üniversitesi
Edebiyat Fakültesi

Not: Tez Savunma Jürisi isimlerinin yer aldığı bu sayfa kapaktan hemen sonra gelecek şekilde ve imzasız olarak konulmalıdır.

TEŞEKKÜR

Yüksek lisans sürecim boyunca akademik gelişimime önemli katkılarda bulunan, tez çalışmamın her aşamasında bilgi ve deneyimleriyle bana rehberlik eden, yol gösterici eleştirileri ve yapıcı önerileriyle araştırmanın yürütülmesinde büyük katkıları olan, desteğini ve sabrını esirgemeyen değerli danışmanım Sayın Prof. Dr. Murat ARICI'ya en içten teşekkürlerimi sunarım.

Ayrıca, bugünlere gelmemde emeği büyük olan ve hayatım boyunca verdikleri sevgi ve destekle yanımda olan sevgili anneme ve babama sonsuz şükranlarımı ve sevgilerimi sunarım.



T. C.
SELÇUK ÜNİVERSİTESİ
Sosyal Bilimler Enstitüsü Müdürlüğü



Öğrencinin	Adı Soyadı	Emine YAKIŞ	
	Numarası	214266001005	
	Ana Bilim / Bilim Dalı	Felsefe	
	Programı	Tezli Yüksek Lisans <input checked="" type="checkbox"/>	Tezli Yüksek Lisans(İÖ) <input type="checkbox"/>
	Tez Danışmanı	Prof. Dr. Murat ARICI	
	Tezin Adı	Hubert Dreyfus ve Yapay Zekâ Eleştirileri	

ÖZET

Yapay zekâ, insan beyninin temelde biyolojik bir makine olduğu ve insan zihninin de sembol manipülasyonları ile çalıştığı fikrine dayanan ve doğru bir programlama ile insan gibi düşünebilen, karar verebilen ve eylemde bulunan yapay sistemler oluşturmayı hedefleyen bir alandır. Oldukça iddialı bir hedefi benimseyen ve heyecan uyandıran bu alana ciddi eleştiriler de yöneltilmiştir. Yapay zekânın önemli eleştirilenlerinden biri olan Hubert L. Dreyfus, insan bilincinin modellenmesine imkân vermeyen öznel yönleri bulunduğunu savunarak, yapay zekâ çalışmalarının kaçınılmaz olarak başarısızlığa uğrayacağını öne sürer. Dreyfus, insanın beden temelli

bir varlık olarak sahip olduđu deneyimlerin ve öznel bilinç durumlarının, mevcut yapay zekâ sistemleri tarafından yakalanamadığını ve ileride de yakalanamayacağını savunur. Bu tez, Dreyfus'un söz konusu eleştirilerinin geçerliliğini ve bu eleştirilerin güçlü yapay zekâ hedefleri üzerindeki etkilerini derinlikli bir analize tabi tutmaktadır. Bu bağlamda tez, Dreyfus'un eleştirilerini (i) halihazırda geçerliliğini yitiren eleştiriler (ii) gelecekte geçerliliğini yitirebilecek eleştiriler ve (iii) ilkesel olarak daima geçerli olacak eleştiriler şeklinde üç ana sınıfa ayırmakta; üçüncü gruba giren ve güçlü yapay zekânın önündeki en büyük engellerden olan bilincin belli niteliklerinin niçin aşılabilir statüde olduğunu masaya yatırmaktadır. Tüm bu incelemeler sonucunda tez, Dreyfus'un eleştirilerine kısmen destek vermekte ve insanın bilinçli zihinselliğini oluşturan fenomenal yönlerin aşılabilir ilkesel sınırlılıklar oluşturduđu ve bu nedenle güçlü yapay zekânın insan benzeri bir bilinç geliştirmesinin mümkün olmadığı sonucuna ulaşmaktadır.

Anahtar Kelimeler: Hubert L. Dreyfus, yapay zekâ, bilinç, bedensel varoluş, arka plan.



SELÇUK ÜNİVERSİTESİ
Sosyal Bilimler Enstitüsü Müdürlüğü



Öğrencinin	Adı Soyadı	Emine YAKIŞ
	Numarası	214266001005
	Ana Bilim / Bilim Dalı	Felsefe
	Programı	Tezli Yüksek Lisans <input checked="" type="checkbox"/> Tezli Yüksek Lisans(İÖ) <input type="checkbox"/>
	Tez Danışmanı	Prof. Dr. Murat ARICI
	Tezin İngilizce Adı	Hubert Dreyfus and His Criticisms Of Artificial Intelligence

ABSTRACT

Artificial intelligence is a field that is based on the idea that the human brain is basically a biological machine, and the human mind works with symbol manipulations and aims to create artificial systems that can think, decide and act like humans with the right programming. This field, which adopts a very ambitious goal and arouses excitement, has also been subjected to serious criticism. Hubert L. Dreyfus, one of the important critics of artificial intelligence, argues that artificial intelligence studies will inevitably fail, claiming that human consciousness has subjective aspects that cannot be modelled. Dreyfus also argues that the experiences and subjective states of consciousness that humans have as a body-based being cannot be captured by current artificial intelligence systems and will not be captured in the future. This thesis provides an in-depth analysis of the validity of Dreyfus' criticisms and their implications for strong AI goals. In this context, the thesis categorizes Dreyfus'

criticisms into three main categories: (i) criticisms that are currently outdated, (ii) criticisms that may become outdated in the future, and (iii) criticisms that will always be valid in principle. As a result of all these analyses, the thesis partially supports Dreyfus' criticisms and concludes that the phenomenal aspects that constitute human conscious mentality constitute insurmountable principled limitations and therefore it is not possible for powerful artificial intelligence to develop a human-like consciousness.

Keywords: Hubert L. Dreyfus, artificial intelligence, consciousness, bodily existence, background.



İÇİNDEKİLER

TEŞEKKÜR	iii
ÖZET	iv
ABSTRACT.....	vi
İÇİNDEKİLER	viii
1. GİRİŞ: YAPAY ZEKÂ: TARİHÇE VE TEMEL YAKLAŞIMLAR	1
1.1. Yapay Zekâ Felsefesi Alanında Ana Yaklaşımlar	6
1.2. Güçlü Yapay Zekâya Yönelik Eleştiriler	7
2. BÖLÜM: HUBERT DREYFUS VE ÇALIŞMALARI.....	15
2.1. Dreyfus: Hayatı, Eserleri ve Genel Felsefe Anlayışı	15
2.2. Dreyfus ve Fenomenoloji Üzerine Çalışmaları	17
2.3. Dreyfus'un Yapay Zekâya Genel Yaklaşımı	23
3. BÖLÜM: DREYFUS VE GÜÇLÜ YAPAY ZEKÂ ELEŞTİRİLERİ: BİR SINIFLAMA.....	28
3.1. Dreyfus'un Eleştirilerinin Bir Sınıflaması	28
3.1.1. Halihazırda Geçerliliğini Yitiren Eleştiriler	31
3.1.2. Gelecekte Geçerliliğini Yitirebilecek Eleştiriler	33
3.1.3. İlkesel Olarak Daima Geçerli Olacak Eleştiriler.....	36
3.2. Dreyfus'un Güçlü Yapay Zekâya Yönelik Eleştirilerinin Önemi.....	40
4. BÖLÜM: GÜÇLÜ YAPAY ZEKÂ VE İLKESEL SINIRLILIKLAR.....	42
4.1. Bedensel Varoluş: Gömülülük, Özümsemişlik ve Bedenlenmişlik	42
4.2. Kenar Bilinç.....	45
3.3. Bağlam Anlayışı ve Deneyim	47
3.4. Fenomenal Bilinç	50
3.5. İçgüdü ve Sezgisel Anlayış.....	55
3.6. Öz Farkındalık ve Anlama	58
5. SONUÇ: İLKESEL SINIRLILIKLAR BİZE İNSAN BİLİNCİ HAKKINDA TAM OLARAK NE SÖYLER?.....	64
KAYNAKÇA.....	68

1. GİRİŞ: YAPAY ZEKÂ: TARİHÇE VE TEMEL YAKLAŞIMLAR

Doğanın karmaşıklığından etkilenecek nesnelere inşa eden insan, modern dönemle birlikte gelişen teknoloji ve artan imkânlar doğrultusunda, kendisine benzeyen makineler üretmeye yöneldi. Bilgisayar bilimi, bu dönemde öne çıkan bir alan oldu ve teknolojik gelişmeler, bilgisayarların gücünü ve kapasitesini hızla artırdı. Bununla birlikte yirminci yüzyılın ortalarında insanın merakını kendisine yöneltmesi bilgisayar biliminde önemli bir evrime zemin hazırladı.

Bu noktada, insan zihninin işleyişinin fiziksel süreçlere indirgenmesi, bu süreçlerin programlanarak zeki makineler üretme düşüncesini ortaya çıkarmıştır. Dolayısıyla yapay zekâ, insanın bilinçli yetilerini yapay bir sistemde üretebilmeyi amaçlayan bir bilim dalı olarak ortaya çıkmıştır. Aynı zamanda insan zihninin karmaşıklığını anlama arzusu, yapay zekâ alanının doğuşunda ve gelişimini şekillendirmede bir diğer önemli etken olarak karşımıza çıkmıştır.

Yapay zekâ alanının kurulması ve gelişmesindeki en etkili olay 1956 yılında düzenlenen Dartmouth Konferansı'dır. Alanın ilk resmi toplantısı sayılabilecek bu konferans, Dartmouth Koleji'nde genç bir profesör olan John McCarthy öncülüğünde düzenlenmiştir. John McCarthy, Marvin Minsky, Nathaniel Rochester ve Claude Shannon tarafından Rockefeller Vakfı'na sunulmak üzere 1955 yılında hazırlanan teklif raporu "yapay zekâ" teriminin ilk defa kullanıldığı metin olarak kabul edilmektedir:

"1956 yazında Hanover, New Hampshire'daki Dartmouth College'da iki aylık, on kişilik bir yapay zekâ çalışması yapılmasını öneriyoruz. Çalışma, öğrenmenin her yönünün ya da zekânın diğer herhangi bir özelliğinin prensipte bir makine tarafından simüle edilebilecek kadar kesin bir şekilde tanımlanabileceği varsayımı temelinde ilerleyecektir."¹

¹ John McCarthy, Marvin Minsky, Nathaniel Rochester ve Claude Shannon tarafından sunulan "A Proposal for the Dartmouth Summer Research Project on Artificial Intelligence" başlıklı öneri metni.

Allen Newell ve Herbert Simon dahil olmak üzere on bilim adamının katıldığı ve sonrasında “Yapay Zekâ Yaz Araştırma Projesi” olarak adlandırılacak olan bu çalışmada asıl amaç, bilgisayar bilimi ve bilişsel bilimdeki önemli figürleri bir araya getirerek yapay zekâ alanını tanıtmak ve ilkelerini belirlemektir. Konferans başlangıcındaki temel inanç ise sembol işleme yeteneğinden ibaret olarak görülen zekânın makine tarafından simüle edilebileceği yönündeydi. Dolayısıyla “Dartmouth konferansı, bilgisayarların kullanımını sayıları hesaplamının ve verileri işlemenin ötesine, sembollerini manipüle etmeye kadar genişletme girişimi olarak görülebilir” (Kaplan, 2016: 14). Bu doğrultuda yapay zekâ çalışması “mevcut amaç uyarınca, yapay zekâ problemi, bir makinenin, insan öyle davrandığında zeki olarak adlandırılacak şekilde davranmasını sağlamak olarak ele alınmaktadır.”² Bu amaca yönelik olarak, insan gibi davranabilen bir makine geliştirme çalışmalarının başlangıcını simgeleyen “Mantık Kuramcısı” (Logic Theorist - LT) yapay zekâ tarihinde erken ve etkili bir başarı örneğidir. İlk çıktıları Dartmouth Konferansı’nda tanıtılmış olan bu program, insanın problem çözme yeteneklerini simüle ederek matematiksel teoremleri kanıtlamak üzere tasarlanmıştır.

1955 yılında, Allen Newell ve Herbert Simon, RAND’da bilgisayar programcısı olan J. C. Shaw’un da katkılarıyla, bugün ilk yapay zekâ programı olarak kabul gören Mantık Kuramcısını geliştirmeye başladılar. Daniel Crevier’in aktardığına göre Simon, *Principia Mathematica* eserinden yola çıkarak mantığa yöneldiklerini ancak teoremlerin etkili bir kanıtını bulmanın ötesinde, “insanların seçici sezgisel yöntemlerle bir sonraki adımda yapılacak doğru şeyi nasıl bulduklarıyla ilgilendiklerini” (Crevier, 1993: 44) söylemiştir. Bu görüş doğrultusunda, Newell ve Simon, Mantık Kuramcısını tasarlarken insanlar gibi sezgisel yöntemleri kullanarak matematiksel teoremleri otomatik olarak kanıtlayabilmeyi hedeflediler. LT, “kanıt arayışında her doğrultuyu körlemesine zorlamak yerine, sezgisel yöntemin ya da “sezgisel aramanın” kılavuzluk ettiği bir arama sürecinden faydalanıyordu” (Nillson, 2018: 118). Program, başlangıç noktasından itibaren her adımı değerlendirir ve hedefe ulaşma olasılığı en yüksek olan adımı takip ederek sonuca ulaşır. Sezgisel yöntemleri

² <http://jmc.stanford.edu/articles/dartmouth/dartmouth.pdf>

kullanarak matematiksel teorem ispatlamayı hedefleyen ilk girişim olan Mantık Kuramcısı, “Russell ve Whitehead’in Principia Mathematica’sının 2. bölümündeki ilk elli iki teoremden otuz sekizini kanıtlamayı başardı... Mantık Kuramcısı”nın bir teorem için (2.85 numaralı) yaptığı ispat Russell ve Whitehead tarafından türetilenden bile daha zarıftı” (Crevier, 1993: 46). Bu, yapay zekânın matematiksel düşünce süreçlerini simüle etme potansiyelini gösteren önemli bir adımdı.

Newell ve Simon, Logic Theorist (LT) ile elde ettikleri başarıları, General Problem Solver (GPS) ile devam ettirdiler. LT, matematikteki teoremleri otomatik olarak türetme amacı taşıyan bir sistemdi. Ancak, GPS, LT’den farklı olarak, başlangıç itibarıyla insanların problem çözme kurallarını taklit edecek biçimde tasarlanmış bir programdı.

Newell ve Simon, programın çalışma stratejisini şu şekilde ifade etmişlerdir:

GPS, çeşitli operatörler tarafından dönüştürülebilen nesnelere oluşan bir görev ortamıyla ilgilenir; nesnelere arasındaki farklılıkları tespit eder; ve görev ortamı hakkındaki bilgileri hedefler halinde düzenler. Her hedef, hedefe ulaşmayı neyin oluşturduğunu tanımlayan, hedefe ulaşmayla ilgili çeşitli türdeki bilgileri kullanılabilir hale getiren ve bilgiyi diğer hedeflerle ilişkilendiren bir bilgi koleksiyonudur

....

Temel olarak GPS programı, ulaşılması ilk hedefe ulaşılmasına yol açacak alt hedefler oluşturarak bir hedefe ulaşmanın bir yoludur (Newell & Simon, 1963: 284).

“Araç-amaç analizi” olarak da adlandırılan bu strateji, “problemleri problemlere dönüştürerek, çözülmesi gereken problem ile verili şey arasında bir “fark” hesaplamaktır. Ardından program, baştaki simge yapısına, simgeleri idare eden bir “operatör” uygulayarak bu farkı azaltmaya çalışır” (Nillson, 2018: 124-125).

GPS, belirli bir problemi çözmek için genel stratejiler kullanır. Program, belirli bir hedefe ulaşmak için izlenecek adımları belirleyebiliyor ve bu adımları insanların aynı sorunlara yaklaşma sırasına benzer bir şekilde dikkate alabiliyordu. GPS’in

başarıları, programın alt hedefleri ve olası eylemleri düşünme sırasının, insanların benzer sorunlara yaklaşma sırasına benzer olduğunu gösterdi. Simon ve Newell'in ortak çalışmaları, bilişsel psikoloji ve yapay zekâ arasındaki bağı güçlendirerek, insan düşünme süreçlerini bilgisayar programlarıyla modelleme konusunda önemli bir temeli oluşturmuştur. Sembolik hesaplama ve dil işleme üzerine yoğunlaşılın ve öncü çalışmaların ortaya çıktığı bu dönemin ardından umut vaat eden yapay zekâ alanında beklenen gelişmeler karşılaşılan zorluklar nedeniyle gerçekleşemedi ve “yapay zekâ kıışı” olarak değerlendirilen bir dönem yaşandı. 1980'li yıllara gelindiğinde uzman sistemler ile birlikte yapay zekâ alanı yeniden canlandı. Uzman sistemler, belirli bir alanda insan benzeri mantık kullanarak problem çözmeye odaklanan yazımları ifade eder. Ayrıca, 1982 yılında Japonya Uluslararası Sanayi ve Ticaret Bakanlığı “Beşinci Nesil Bilgisayar Sistemleri” adını verdiği bir proje başlattı. “Bu projenin amacı, büyük veri ve bilgi tabanlarından yapay zekâ çıkarımları yapabilecek ve doğal dil aracılığıyla insanlarla iletişim kurabilecek bilgisayarlar üretmekti” (Nillson, 2018: 353). İlerleyen süreçlerde yapay zekâ çalışmaları örüntü tanıma, doğal dil işleme, makine öğrenmesi gibi alanlarda gelişim göstererek önemli ilerlemeler kaydedildi. Oldukça ses getiren gelişme ise 1997 yılında IBM'in “Deep Blue” adlı programının dünya şampiyonu Garry Kasparov'u yenmesiydi ve bu olay yapay zekânın önemli bir başarısı olarak tarihe geçti.

İlk bilgisayar programlarından günümüzdeki karmaşık öğrenme algoritmalarına kadar olan bu yolculuk, insanın kendi zekâsını anlama ve makinelerde yeniden oluşturma çabasının bir yansımasıdır. Modern teknolojinin en çarpıcı ve etkileyici alanlarından biri olarak yapay zekâ, bilgisayar sistemlerinin insan benzeri düşünce yeteneklerini simüle etme amacı doğrultusunda gelişim göstermektedir. Bu, zihinsel süreçleri, öğrenmeyi ve hatta duygusal deneyimleri taklit etme çabasını içerir ve “bilgi işleme, doğal dil işleme, yapay görme, makine öğrenimi, robotik ve benzeri çok yönlü bilgisayar simülasyonlu uygulamaları kapsayan yüksek ve güçlü simülasyon faaliyetleri olarak tanımlanır” (Yeşilkaya, 2022: 100).

Yapay zekâ, bilim ve teknolojideki etkileyici ilerlemelerle birlikte, felsefe disiplini içinde yeni bir düşünsel çerçevenin doğmasına öncülük etmiştir. “Yapay zekâ, felsefi düşünce ile bilim ve teknolojiyi bir araya getirerek bu disiplinler arasında bir

entegrasyon noktası oluşturmaktadır” (Yeşilkaya, 2022: 99). Yapay zekânın gelişimi, felsefe, bilim ve teknolojiyi birbirine yaklaştırarak, bu alanlardaki etkileşimi ve birlikte çalışmayı kaçınılmaz kılmaktadır. Yapay zekâ her zaman felsefi bakış açısına ihtiyaç duyar; zira alanda karşılaşılan zorluklar aslında temel felsefi sorunları içermektedir. Yapay zekâ felsefesi, teknolojinin ve bilimin gelişimine paralel olarak ortaya çıkan yeni sorulara yanıt arar ve bu teknolojinin insan doğası, bilinç ve varlıkla ilgili temel meselelere etkilerini değerlendirir. İnsan düzeyinde zekâ ve yeteneklere sahip bir makine inşa etmenin mümkün olup olmadığı konusunu içeren bu felsefi alanda, yapay zekânın doğası ve sınırları üzerine çeşitli konular ve argümanlar ele alınmaktadır.

Alan Turing’in 1950 yılında *Mind* dergisinde yayınlanan “Bilgiişlem Makineleri ve Zekâ” (Computing Machinery and Intelligence) adlı makalesinde ele aldığı “Makineler düşünebilir mi?” sorusu, felsefi düzlemde merkezi bir konuma sahiptir. Çünkü Turing bu makalesiyle birlikte yapay zekâ çalışmalarının felsefi düzlemde ele alınmasını sağlayarak yeni bir tartışma alanının doğuşunu sağlamıştır. Ayrıca, yapay zekânın çalışma alanı Darmouth Konferansı ile ifade edilmiş olsa da yapay zekâ düşüncesinin, terimin kullanılmasından daha öncesinde Alan Turing tarafından ortaya koyulduğu görülmektedir. Bahsi geçen makalede makinelerin zeki olduğuna karar vermek için taklitçilik oyunu olarak ifade ettiği, sonrasında ise “Turing testi” adıyla anılan bir oyun geliştiren Turing, bu testi geçen yapay bir sistemin insan zihinselliğine eşdeğer olduğunun kabul edilmesi gerektiğini savunmaktadır. Alan Turing önerdiği oyunu şu şekilde anlatır:

Üç kişi ile oynanır; bir erkek (A), bir kadın (B) ve her iki cinsiyetten de olabilen bir sorgulayıcı (C). Sorgulayıcı, diğer ikisinden ayrı bir odada kalır. Sorgulayıcı için oyunun amacı, diğer ikisinden hangisinin kadın, hangisinin erkek olduğunu belirlemektir. Onları X ve Y etiketleriyle tanır ve oyunun sonunda “X, A”dır ve Y, B”dir” ya da “X, B”dir ve Y, A”dır” der (Turing,1950:433).

Turing daha sonrasında A’nın yerine bir bilgisayarı yerleştirir. Yani aslında Turing testi, bir sorgulayıcı kişinin klavye aracılığı ile diyalog kurarak bir yapay zekâ ile bir insandan hangisinin yapay zekâ olduğunu anlayabilmesi üzerine kuruludur. Eğer hangisinin yapay zekâ olduğu bilinemiyorsa test geçilmiştir ve yapay zekânın bilinçli

olduğunun kabul edilmesi gerekir. Turing'in bu iddiası insan benzeri yapay zekânın mümkün olduğunu düşünenler için oldukça ilgiyle karşılanırken, diğer taraftan yapay bir sistemde insana özgü bilinç halini modellemenin mümkün olmayacağını düşünen filozoflar tarafından çok ciddi eleştirilerin de hedefi olmuştur. Bu eleştirileri değerlendirmeden önce yapay zekâ felsefesindeki temel yaklaşımlar olan zayıf ve güçlü yapay zekâ arasındaki görüş farklılıklarını incelemekte fayda var. Zira bu yaklaşımlar Turing'e getirilen eleştirilerle birlikte ortaya koyularak yapay zekâ felsefesi içinde temel bir konu haline gelmiştir.

1.1. Yapay Zekâ Felsefesi Alanında Ana Yaklaşımlar

Yapay zekâ felsefesi alanında, bilgisayar sistemlerinde insan benzeri düşünce süreçlerinin ne ölçüde taklit edilebileceği, birçok temel kavramı içeren merkezi bir tartışma konusunu oluşturmaktadır. Görüş farklılıklarını anlamak adına üzerinde durmamız gereken temel mesele, yapay zekâ alanındaki iki temel yaklaşım olan güçlü yapay zekâ ve zayıf yapay zekâ ayrımıdır.

İlk kez, Turing testini geçen yapay bir sisteme zekâ atfedilebileceği iddiasında bulunan Alan Turing'e karşı etkili bir eleştiri getiren John Searle'ün ele aldığı bu ayrım, yapay zekâ çalışmalarında ulaşılmak istenen hedefi ve güçlü yapay zekânın imkânını belirlemekte önemli bir noktada yer alır. Zayıf yapay zekâ, filozoflar tarafından çok fazla eleştiri getirilmeyen, yapay bir sistemin belirli uzmanlıkları yerine getirebileceği yönündeki görüşü temsil eder. Bu yaklaşıma göre, belirli görevleri yerine getirmek üzere programlanan yapay zekâ sistemleri; insan gibi davranabilen, ancak düşünce, akıl yürütme ya da deneyim sahibi olabilme gibi zihinsel süreçlere sahip olmayan sistemlerdir. Dolayısıyla zayıf yapay zekâ, spesifik problemleri çözmek ya da insanları belli alanlarda desteklemek üzerine programlanmış kural tabanlı sistemleri ifade eder. Belirli kurallar ekseninde çalışabilen bu sistemler, hedeflenen görevleri insan düzeyinde, hatta çok daha etkili bir şekilde yerine getirebilecektir. Genel anlamda bir bilinç halini hedeflemeyen bu yaklaşım doğrultusunda, yapay zekânın insan zekâsını taklit etmekten öteye geçemeyeceği görüşü desteklenmektedir. Searle'ün ifadesiyle, "Zayıf yapay zekâyâ göre bilgisayarın zihin araştırmasındaki temel değeri, bize çok güçlü bir araç sunmasıdır" (Searle, 1980: 417).

Güçlü yapay zekâ görüşü ise, üretilen bir programın insanın bilinçli yönünü taklit etmekten daha çok, gerçekten bilinç sahibi olabileceği yönündeki fikri savunmaktadır. Güçlü yapay zekâ görüşü, dijital bir bilgisayarın insanlara özgü olduğu düşünülen akıl yürütme, anlama, düşünme gibi bilinçlilik atfedeceğimiz yetilere sahip olacak şekilde programlanabileceğini ileri sürer. Bu görüş temelde işlevselcilik kuramından etkilenmiştir. İşlevselciliğe göre, zihinsel durumları anlamlandırmak için biyolojik yapıya değil, o durumun sistemdeki işlevsel süreçlerine odaklanması gerekir. Bu görüşü benimseyen yapay zekâ kuramcıları, beyni bir bilgisayar donanımı, zihinsel süreçleri ise bilgisayar donanımına yüklenen programlar gibi görürler. “Bu bakış açısına göre, ne olursa olsun, doğru bir programı ve doğru girdi-çıkı ilişkileri olan herhangi bir fiziksel sistem, tam olarak sizin ve benim zekâlarımızla aynı anlamda bir zekâyâ sahip olacaktır” (Searle, 1996: 39).

Günümüzdeki yapay zekâ çalışmalarının merkezinde, insan benzeri düşünce yetenekleri, algılama kapasiteleri ve duygu durumlarına sahip yapay zihinlerin üretilmesi hedeflenmektedir. Yani hedef güçlü yapay zekâyı gerçekleştirmektir. Bu hedefe yönelik çalışmaların karşısında insanın zihinsel içeriklerinin çeşitliliği ve karmaşık yapısı çok büyük bir engel oluşturmakla birlikte, yapay zekâyâ bilinç kazandırma çabası ile insan zihninin doğasını anlama isteği bilim ve felsefe dünyasında önemli tartışmaların odağını oluşturmaktadır.

1.2. Güçlü Yapay Zekâyâ Yönelik Eleştiriler

Yapay zekâ felsefesinin derinliklerine doğru bu yolculukta, güçlü yapay zekâ perspektifinin eleştirel bir incelemesi önem kazanır; zira bu bakış açısı, temel varsayımları ve sonuçları hakkında sorgulanabilir bir dizi eleştiriyi beraberinde getirir. Yapay zekâ alanında önemli felsefi sorulara odaklanan ve güçlü yapay zekâ görüşüne yönelik sorgulamalarıyla ön plana çıkan üç önemli düşünür; Searle, Gödel ve Dreyfus'tur. Getirdikleri çeşitli eleştirilerle yapay zekâ alanına önemli felsefi katkılarda bulunan bu düşünürlerin çalışmalarını irdelemek, yapay zekâ alanının felsefi boyutunu ve yapay zekânın sınırlılığı iddiasını daha iyi kavramamızı sağlayacaktır.

Kuşkusuz, alana yönelik en etkili eleştiri, John R. Searle tarafından geliştirilmiştir. İlk kez “güçlü” ve “zayıf” yapay zekâ ayırımına değinen Searle’ün asıl eleştirisi “doğru şekilde programlanmış bir bilgisayarın kelimenin tam anlamıyla anladığı ve diğer zihin durumlarına sahip olduğu” (Searle, 1980: 417) iddiası olarak ele aldığı güçlü yapay zekâ görüşüne yöneliktir. Turing testini geçmenin, bir bilgisayarın düşünme yetisine sahip olduğunu söylemek için yeterli olduğunu savunan Alan Turing’e karşı John Searle, bilgisayarın zihne sahip olabilmesi için testi geçmesinden başka özelliklere sahip olması gerektiğini savunur. *Zihinler, Beyinler ve Programlar* adlı makalesinde ele aldığı ve “Çince Odası Argümanı” olarak bilinen ünlü deneyinde Searle, yalnızca soruların geldiği giriş ve cevapların verildiği çıkış deliği olmak üzere iki delikten oluşan bir odada kilitli olduğunu varsayar. Searle tek kelime Çince bilmiyordur ancak odanın içinde Çince sorulara verilecek yanıtları gösteren kitaplar ve karşılaştırmayı yapabileceği İngilizce kurallardan oluşan bir kılavuz vardır. Dışardan gelen Çince sorulara karşılık gelen doğru yanıtı kitaplar aracılığıyla bularak cevabını çıkış deliğinden geri gönderir. Doğru cevap karşısında dışardaki insanlar odanın içindeki kişinin Çince bildiğini düşünür ancak tek kelime Çince bilmeyen Searle’ün tek yaptığı belirtilen kurallar doğrultusunda sembollerini bir araya getirmektir. Cevaplarının anlamından ve bunları çözmenin içerdiği zihinsel süreçlerden tamamen habersizdir. “Hiçbir bilgisayar programının akla denk olamayacağının nedeni çok basittir: Bilgisayar programı yalnızca bir sözdizimdir; akılsa, sözdizimden çok başka bir şeydir. Akıl, anlamsaldır ve içeriği vardır” (Searle, 1996: 44). Bizim zihinsel durumlarımız, tanımları dolayısıyla belli içeriklere sahiptirler.

Searle’ün; bir bilgisayar programının, kendisine yüklenen ve bir noktada insani davranışlar diyebileceğimiz belirli görevleri yerine getiremeyeceği iddiasında olmadığını belirtmekte fayda var; “ancak programın, bu faaliyetlerle uğraşırken insan zihninde meydana gelen süreci kopyalamak yerine, düşünmeyi simüle ettiğine inanmaktadır” (Kaplan, 2016: 73). Bu doğrultuda, sembol manipülasyonlarından ibaret olan bir program, dışardan zekice davranışlar sergilese dahi, sistemin kendisi hiçbir farkındalığa sahip olamayacaktır, yani gerçekleştirdiği hiçbir davranışını anlamlandıramayacaktır. Sonuç olarak, güçlü yapay zekâ imkansızdır, çünkü sembol

manipülasyonları anlamdan yoksundur. Yalnızca sentakstan oluşan bilgisayar programı, anlama ve içeriğe sahip olan akla denk olamayacaktır; salt sentakstan semantiğin çıkması ilkece mümkün değildir.

Özetlemek gerekirse, Searle, sadece sembol manipülasyonu ile zihinsel içeriği açıklamanın ve anlama durumunun gerçekleşmesinin mümkün olamayacağını, gerçek anlamda bir bilinç sahibi olabilmek için de fiziksel bir beyin sürecinin gerekliliğini savunmuştur. Bu bağlamda, anlam deneyiminin insana özgünlüğüne vurgu yapan Searle'ün eleştirisi, yapay zekâ alanında önem arz eden sembolik hesaplama modeline yönelik etkili bir eleştiri olmuştur. Güçlü yapay zekâ görüşüne yönelik Searle'ün eleştirilerinin ardından, şimdi de çalışmalarıyla yapay zekâ alanının temel sorularını ve sınırlılıklarını farklı bir perspektiften değerlendirmemizi sağlayan; matematik, mantık ve matematik felsefesi alanlarında önemli bir figür olan Kurt Gödel'e odaklanalım.

19. Yüzyılın sonlarına gelindiğinde, aritmetiğin tümüyle mantığa indirgenebileceği fikri önem kazanmaya başlamıştır. Russell ve Whitehead, bu görüşü savunarak *Principia Mathematica* adlı eserlerini sunmuşlar ve dönemin düşünce dünyasında önemli bir etki yaratmışlardır. Aynı dönemde, David Hilbert'in temsilcisi olduğu matematiğin mutlak bir biçimselleştirilmesini amaçlayan biçimci (formalizm) yaklaşım da öne çıkmış ve matematik felsefesinde yeni bir dönemin başlamasına yol açmıştır (Gözkân, 2010: xiii-xiv). Aslında o dönem hâkim olan genel düşünce matematiğin temellerini sağlamlaştırmaktı. Bu düşünce neticesinde Hilbert, kendi adıyla anılan bir program da geliştirmişti, ancak bazı paradoks ve tutarsızlıklarla karşı karşıya kalmıştı (Eyim, 2019: 222). “Hilbert paradoksların çözümü olarak, eksiksiz ve kendi içinde tutarlı bir biçimde belirlenen sonlu sayıdaki aksiyom dizilerini kullanarak matematikteki tüm önermelerin kanıtlanabilir olduğunun gösterilebileceğini öne sürmüştür” (Eyim, 2019: 222). Bu çalışmaların etkisinde kalan Kurt Gödel, matematik sistemlerin içsel tutarlılığına odaklanmış ve Hilbert'in programını incelemesinin ardından da programın eksikliklerini ve sınırlamalarını fark etmiştir. 1931 yılında Gödel, matematikteki formel sistemlerin sınırlamalarına dair geliştirdiği “Über formal unentscheidbare Sätze der Principia Mathematica und verwandter Systeme I” (“Principia Mathematica ve Benzeri Sistemlerin Formel Olarak Çözümsüz Teoremleri

Üzerine I”) adlı makalesini yayımlamıştır. Gödel bu makalesi ile birlikte, döneme egemen olan, “matematikselse düşünceenin her dalının, sonsuz sayıda doğru önermeler topluluğuna olanak verecek biçimde dizgesel olarak bir aksiyomlar kümesi ile kurulabileceği” (Nagel & Newman, 2010: 37) yönündeki genel kanının, aksiyomatik yöntemin bazı içsel sınırlamaları olduğu sonucunu ortaya koyarak desteklenemez olduğunu göstermiştir.

Bu çalışmasında Gödel iki önemli sonuca, “Gödel’in Eksiklik Teoremleri” olarak da bilinen iki teoreme ulaşmıştır;

- (i) “Gödel’in Birinci Eksiklik Teoremi: Yeterince ifade edilebilir herhangi bir matematik sistem eksik ya da tutarsız olmak zorundadır” (Eyim, 2019: 224). Yani, matematikte belirli bir formel sistem içsel tutarlılığını kanıtlamaz.
- (ii) “Gödel’in İkinci Eksiklik Teoremi: Tutarlı bir matematiksel sistem kendi tutarlılığını ispatlayamaz” (Eyim, 2019: 224). Yani, bir formel sistem içinde tutarlılığı göstermek için dışardan başka bir sisteme ihtiyaç duyar.

Gödel’in teoremi; “yeterli güce sahip herhangi bir mantıksal (logical) sistemin durumları, muhtemelen sistemin kendisi tutarsız olmadıkça, sistem içinde ne ispatlanabilen ne de ispatlanamayan şekillerde formüle edilebilir” (Turing, 1950: 444) olduğunu göstermektedir. Bu sonuç, matematikte bir sistem içinde tüm gerçekleri kapsayan bir sistem oluşturmanın imkânsız olduğunu, yani her sistemde kanıtlanamayan doğru ifadelerin var olduğunu gösterir.

Gödel’in teoremleri, yapay zekâ alanındaki düşünceyi etkilemiş ve bilgi işleme sistemlerinin sınırlamalarını anlamak için önemli bir çerçeve sağlamıştır. Zira, “Gödel’in biçimsel dizgelerde kendisinin ya da değıllemesinin doğru olduğu gösterilemeyen en az bir önermenin her zaman var olacağı savı, yapay zekânın belirgin bir kısıtlamasının olduğunu göstermektedir” (Eyim, 2019: 226) Bilgisayarlar ve bilgisayar programları, bir sorunu çözmek veya bir sonuca ulaşmak için, önceden belirlenmiş komutları izleyerek işlem yaparlar. “Ancak Gödel, tam olamama teoreminde gösterdiği gibi, temel sayılar kuramında belirlenmiş aksiyomatik yöntemin alanının dışına düşen sayısız sorun bulunmaktadır; ve bu tür aletler, işleyişleri ne kadar

karışık, dâhice olursa olsun, işlemleri ne kadar hızlı yaparlarsa yapsınlar, bu sorunlara yanıt veremezler.” (Nagel & Newman, 2010: 116).

Aslında Alan Turing de 1936 yılında teorik bir hesaplama modeli geliştirerek, tamamen farklı bir biçimde de olsa, Gödel’in ulaştığı sonuca benzer bir sonuca ulaşmıştır. Günümüzde ‘Turing Makinesi’ olarak bilinen bu varsayımsal model, hesaplanabilir fonksiyonları belirleme motivasyonu ile geliştirilmiştir. Basit bir yapıya sahip olan Turing makinesi, hücrelere bölünmüş ve her hücreye bir sembolün (0 ya da 1’in) yerleşebildiği sonsuz uzunlukta bir bant ve bir bant sürücüsünden oluşur (Nillson, 2010: 60). Sonsuz uzunluktaki bu bant sonsuz bir depolama alanına da sahip olacaktır. Bant üzerine konumlanmış olan okuma-yazma kafası, bant üzerindeki sembolleri okuyabilir, üzerine yazabilir ve sola ya da sağa hareket edebilir. Turing makinesi, bir başlangıç durumundan başlar; okuma-yazma kafası bant üzerindeki sembolü okur ve bu sembol makinenin içinde bulunduğu durum neticesinde bir kurala karşılık gelir; bu kurala göre makine kafayı hareket ettirecek ya da sembolü değiştirecektir ve sonuçta bir hesaplama gerçekleşecektir. Bu süreç makine bir duruma ulaşana kadar devam edecektir. Yani Turing makinesi, belli bir duruma ulaşana veya durması durumuna gelene kadar belli kurallara göre semboller üzerinde işlem yapar. “Turing, kendi makinesi için belli bir mantık biriminin (makinenin edimlerine karar veren aksamın) daima belirlenebileceğini ve bu birim sayesinde, hesaplanabilir herhangi bir fonksiyonun hesaplanabileceğini ispatlamıştır” (Nillson, 1980: 61). Ayrıca Gödel gibi Turing de bazen insanlara önemsiz görünen ve hiçbir Turing makinesinin asla gerçekleştiremeyeceği belirli türde hesaplamaların mevcut olduğunu belirtti (Crevier, 1993: 23). Durma problemi en önemli hesaplanamayan problemlerden biridir; “bir bilgisayarın program dilini oluştururken biçimsel dizgeye ispatlanamayacak önermeler yerleştirebilir ve sistemde bir probleme neden olup algoritmanın dışına çıkmadığından bilgisayar duracaktır” (Eyim, 2019: 225). İnsanlar, matematikteki bazı gerçekleri sezgisel olarak anlayabilirken, bir formel sistem içinde bu gerçeklerin kanıtlanamayacağını gösteren teoremler, insan zekâsının formel sistemlerden öte bir yeteneğe sahip olduğunu gösterir. Bu bağlamda matematiksel sistemlerde tam doğruluk iddialarının mümkün olmadığını gösteren Gödel’in teoremleri ile Turing’in farklı bir açıdan ulaştığı hesaplanamayan

fonksiyonların olduğu sonucu, güçlü yapay zekâ görüşünün karşısındaki bazı temel sınırlılıkların görülmesini sağlamıştır.

Güçlü yapay zekâ perspektifinin sınırları ve içsel karmaşıklıkları konusundaki Searle ve Gödel'in özgün bakış açılarını değerlendirdiğimiz bu noktada, bu görüşe karşı ilk ciddi eleştirileri getiren isimlerden biri olan Hubert Dreyfus'un önemli bir rol oynadığını belirtmeliyiz. Dreyfus, yapay zekâ alanındaki ilk eleştirilerden biri olmasıyla birlikte, ortaya koyduğu eleştirilerin etkisiyle bu alanda önemli bir konuma ulaşmıştır. Onun bakış açısıyla yapay zekânın önemli sınırlılıklarını keşfetmek, güçlü yapay zekâ görüşünün belirli noktadaki zayıflıklarını anlamak adına kritik bir adım olarak karşımıza çıkmaktadır.

Yapay zekâ çalışmalarının ilk zamanlarında, insan bilincinin fenomenal yönüne dikkat çeken ve yapay zekâyâ karşı etkili bir eleştiri getiren Amerikalı filozof Hubert Lederer Dreyfus, fenomenal bilincin yapay zekâ çalışmaları için ciddi bir zorluk olduğunu fark ederek önemli çalışmalar ortaya koymuştur. Güçlü yapay zekânın mümkün olmadığına inanan Dreyfus, yapay zekâ araştırmacılarının iyimser iddialarının ilk eleştirilerinden biridir ve dikkat çektiği noktalar, yapay zekâ alanı için zorluk oluşturmaya devam eden insanın fenomenal yönü üzerinedir.

Dreyfus'un yapay zekâ eleştirisi, fenomenolojik bir perspektiften yapay zekânın temel sınırlamalarını ve insan zekâsının özgün niteliklerini vurgulamaktadır. Dreyfus'un eleştirileri, yapay zekâ araştırmalarının başlangıcında özellikle sembolik hesaplama modellerine yönelik olarak ortaya çıkmıştır. Dreyfus, hesaplamalı sistemler aracılığıyla insan benzeri zekâyâ ulaşma olasılığını sorgulamış ve zekânın açık kurallar ve algoritmalarla tam olarak yakalanabileceği yönündeki yaygın görüşe meydan okumuştur. Dreyfus'a göre, sembolik mantık ve kural tabanlı sistemlere dayanan yapay zekâ yaklaşımları, insan uzmanlığının doğasında var olan incelikleri yakalamakta başarısızdır.

Zekânın sembol işleme yoluyla elde edilebileceği fikrine karşı çıkan Dreyfus, insan zekâsının soyut, sembollerle manipüle edilebilir bir süreç olmadığını, aksine bedensel deneyimler, çevresel etkileşimler ve bağlamsal faktörlerle derinden bağlantılı olduğunu vurgular. İnsan zekâsı sembol manipülasyonuna indirgenemeyen, pratik,

günlük anlayışa dayalı bir arka plan bilgisine sahiptir. Ayrıca Dreyfus, bilgisayarların bağlamı anlayamaması, insan sezgisini ve sağduyusunu taklit edememesi, belirsizlik ve muğlaklıkla başa çıkamaması gibi çeşitli sınırlılıklarının bulunduğunu ortaya koymaya çalışmıştır.

Dreyfus'un eleştirileri yapay zekânın bazı temel insan özelliklerini veya yeteneklerini modelleyebilme konusunda sınırlamaları olduğunu ortaya koymuştur. Tartışmalı olmasına rağmen, Dreyfus'un fikirleri AI araştırmalarının sınırlamalarının daha incelikli bir şekilde anlaşılmasını sağladı ve daha esnek ve insan benzeri yollarla öğrenebilen ve uyum sağlayabilen akıllı makineler tasarlamaya yönelik yeni yaklaşımlara ilham kaynağı oldu.

Hubert Dreyfus, bu alandaki öncü isimlerden biridir ve getirdiği eleştirilerle dikkat çekmiştir. Yapay zekâ disiplininin henüz gelişme aşamasında ortaya koyduğu argümanları, zaman içinde birçok eleştiriye maruz kalmış olmasına rağmen, günümüzde bile bazı eleştirileri hala etkisini sürdürmektedir. Bu bağlamda, Dreyfus'un eleştirilerinin uzun vadeli etkileri ve günümüz yapay zekâ araştırmalarına olan potansiyel etkileri üzerine düşünmek, yapay zekâ alanının evrimini anlamak ve gelecekteki gelişimini öngörmek açısından oldukça önemlidir. Eleştirilerinin günümüz yapay zekâ sahnesindeki yeri, bu alandaki yenilikleri ve ilerlemeleri değerlendirmek adına kilit bir rol oynamaktadır. Bu kapsamda bu çalışma, Dreyfus'un yapay zekâ eleştirilerini kategorik bir biçimde ele almayı, bu eleştirileri sistematik bir incelemeye tabi tutmayı ve söz konusu inceleme sonucunda elde edilen sonuçların günümüz yapay zekâ çalışmalarına ne ölçüde katkı sunabileceğini ve bu çalışmaların gelişimine hangi açıdan ilham olabileceğini ortaya koymayı amaçlamaktadır.

Bu minvalde, bu tezde öncelikle Bölüm 1'de Hubert Dreyfus'un yaşamını, eserlerini ve genel felsefi yaklaşımını ele alacağız. Bölüm 2'de, Dreyfus'un güçlü yapay zekâyâ ilişkin eleştirilerini sınıflandırarak, geçerliliğini yitirmiş, gelecekte geçerliliğini yitirebilecek ve ilkesel olarak her zaman geçerli olacağını düşündüğümüz eleştiriler üzerinde duracağız. Bölüm 3'te, insanın bilinçli zihinselliğine özgü yönlerini detaylı bir biçimde inceleyecek ve önceki bölümde ele aldığımız bu ilkesel sınırlılıkların yapay zekâ açısından neden temel engeller teşkil ettiğini göstermeye

çalıřacađız. Sonu blmnde ise, Dreyfus'un eleřtirilerinin insan bilincine iliřkin hangi ilkesel sınırlılıkları ortaya koyduđunu ve gl yapaı zekânın bu sınırları ařmakta neden zorlandıđını deđerlendireceđiz.



2. BÖLÜM: HUBERT DREYFUS VE ÇALIŞMALARI

Giriş bölümünde yapay zekânın tarihçesini ana hatlarıyla ele aldıktan sonra, bu bölümde güçlü yapay zekâyâ yönelik eleştirileriyle tanınan Hubert Dreyfus'un yaşamını, çalışmalarını ve yapay zekâyâ bakış açısını inceleyeceğiz. Dreyfus, insan zihninin bilgisayarlara indirgenemeyeceğini ve bilişsel süreçlerin yalnızca bedensel ve bağlamsal bir temelde anlaşılabilirliğini savunan bir yaklaşıma sahiptir. Burada, Dreyfus'un fenomenolojiye dayalı eleştirileriyle yapay zekâyâ dair ortaya koyduğu temel argümanlara odaklanacağız.

2.1. Dreyfus: Hayatı, Eserleri ve Genel Felsefe Anlayışı

Hubert Lederer Dreyfus, 15 Ekim 1929 tarihinde Indiana'nın Terre Haute kentinde dünyaya geldi. Bilimle olan başlangıç eğitimi 1947 yılında Harvard Üniversitesi fizik bölümünde başladı; ancak C.I. Lewis'in verdiği epistemoloji dersine katılması, onu felsefeye yönlendirdi. 1951 yılında Harvard'dan en yüksek onur derecesiyle mezun olan Dreyfus tezini "Nedensellik ve Kuantum Teorisi" üzerine yazdı. 1952'de Master of Arts derecesini alarak Harvard Üniversitesi'nde öğretim üyeliğine başladı.

Harvard Sheldon gezici bursu ile Avrupa'da eğitim gördü. Bu süreçte, Basel'de Karl Jaspers'ı ve Freiburg'da Heidegger'in takipçilerini dinleyerek farklı düşünce biçimleriyle tanıştı. 1957-60 yılları arasında Brandeis Üniversitesi'nde felsefe eğitmeni olarak görev yaptı ve 1960 yılında Massachusetts Institute of Technology'de (MIT) yardımcı doçent olarak felsefe dersleri vermeye başladı.

Dreyfus, 1964 yılında "Husserl'de Algı Fenomenolojisi" adlı teziyle Harvard Üniversitesi'nden doktorasını tamamladı. 1967 yılında doçent unvanını aldı ve bir yıl daha MIT'de görev yaptı. Daha sonra California Üniversitesi, Berkeley'de doçent olarak ders vermeye başladı ve 1972 yılında profesör unvanını kazandı. 1994 yılında Enstitü Profesörü olan Hubert L. Dreyfus, emekli olduğu 2016 yılına kadar görevine devam etti. 22 Nisan 2017 tarihinde hayatını kaybetti.

Hubert Dreyfus, 20. yüzyıl kıta felsefesinin önde gelen isimlerinden biridir ve geniş bir felsefi yelpazede etkileyici çalışmalara imza atmıştır. Özellikle varoluşçuluk,

fenomenoloji ve yapay zekâ alanlarında önemli katkılarda bulunmuştur. Hayatı boyunca, insanın varoluşsal deneyimini ve bu deneyimin bilgi, teknoloji ve insan ilişkileriyle olan etkileşimlerini derinlemesine incelemiştir. Kariyeri boyunca aldığı ödüllerle ve etkili felsefi düşüncesiyle bilim dünyasında saygın bir konum elde etmiştir.

Dreyfus'un ilk çalışması, James Broderick ile birlikte hazırladığı 1957 yılında yayınlanan "Curds and Lions in Don Quixote" isimli makalesidir. Çalışmalarının başlangıcında bulunan bu makalesinin ardından sayısız makale yazmıştır. Bunun yanı sıra kitapları, antolojileri, video kasetleri, röportajları ve birçok yorumu da bulunmaktadır.

En çok ses getiren çalışması, RAND'da danışmanlık yaptığı zamanda, yapay zekâya karşı hazırladığı raporudur. Yapay zekâ ile ilk kez MIT'de ders vermeye başladığı yıllarda karşılaşılan ve sonrasında bu konudaki düşünceleriyle öne çıkan Dreyfus; öğrencilerin dil öğrenen, anlayan ve problem çözen programlar geliştirmeye çalıştıklarını ancak bu çabaların başarılı olmadığını gözlemlemiştir. 1964 yılında kardeşi Stuart Dreyfus'un aracılığıyla, yapay zekânın araştırılıp geliştirildiği RAND Şirketi ile temasa geçmiştir. Bu süreç, Dreyfus'un yapay zekâya dair kritik görüşlerini daha da şekillendirmesine katkı sağlamıştır. RAND, Dreyfus'u yapay zekânın fizibilitesini felsefi açıdan değerlendirmek üzere danışman olarak işe almıştır (Crevier, 1993: 121). Bu görevi sırasında Dreyfus, "Alchemy and Artificial Intelligence" (Simya ve Yapay Zekâ) adlı bir rapor sunmuş ve bu alanda ciddi eleştirilere imza atmıştır. Elbette yapay zekâ çalışanları arasından bu eleştirilere karşı tepki gösterenler olmuştur. Ancak Dreyfus'un düşünceleri ve eleştirileri, yapay zekâ konusunda daha derin düşünmeye ve daha etraflı tartışmalara yol açmıştır. Her ne kadar Dreyfus'un söz konusu çalışması yapay zekâ alanında dikkat çekmiş olsa da bu çalışmanın yayınlanması yapay zekâ çalışanları tarafından engellenmeye çalışılmıştır. Dreyfus'un bu makalesinde, yapay zekâ görüşüne karşı oldukça etkili eleştiriler bulunmaktadır. MIT'teki ve RAND şirketindeki görev sürecinde, yapay zekâ çalışanları arasında kendisini öne çıkaran bu eleştirel tavrı, Dreyfus'u sektördeki ilk etkili eleştirmenlerden biri yapmıştır.

Dreyfus daha sonrasında raporu genişleterek *What Computers Can't Do: A Critique of Artificial Reason* adlı bir kitap haline getirmiş ve 1972 yılında yayınlamıştır. On iki dile çevrilen bu eseriyle etkili bir çıkış yakalayan Dreyfus, yapay zekâ alanında önemli bir figür haline gelmiştir. Daha sonrasında kitabı gözden geçirip geliştirmiş, ikinci ve üçüncü baskılarında ismini *What Computer Still Can't Do* olarak düzenlemiştir.

Hubert Dreyfus, felsefi kariyeri boyunca bir dizi etkileyici eser kaleme almıştır. Bu eserler arasında öne çıkanlar arasında, kardeşi Stuart Dreyfus ile birlikte kaleme aldığı "Mind over Machine: The Power of Human Intuitive Expertise in the Era of the Computer" (1986) bulunmaktadır. Dreyfus'un diğer önemli eseri ise "Being-in-the-World: A Commentary on Heidegger's Being and Time, Division I" (1991) adlı eseridir. Bu eser, Heidegger'ın temel eseri "Varlık ve Zaman" hakkında bir yorum içerir. Dreyfus, bu çalışmasında Heidegger'ın fenomenolojik düşüncesini anlamak ve derinleştirmek amacıyla kapsamlı bir analiz sunar. Bu eserler, Dreyfus'un genel felsefe anlayışını ve düşünce tarzını daha iyi anlamamıza yardımcı olurken, aynı zamanda çağdaş felsefeye önemli katkılarda bulunmuştur.

2.2. Dreyfus ve Fenomenoloji Üzerine Çalışmaları

Fenomenolojik yaklaşımı Hubert Dreyfus'un yapay zekâ konusundaki eleştirilerinin temelini oluşturur. Dreyfus'un fenomenolojik görüşü, insan bilişselliğinin ve anlamlandırma kapasitesinin bedenlenmiş varoluşu ve bağlamsal anlayışı ile ilişkili olduğunu iddiası ile temellenmiştir. Fenomenolojik perspektife göre, deneyimlerimiz bedensel varoluşumuz ve dünya ile olan etkileşimimizle şekillenmektedir. Dreyfus, insan zihninin sembol manipülasyonları yoluyla çalışmadığını ileri sürer ve zekâmızın dünyayla olan pratik ve bedensel etkileşimlerimiz yoluyla geliştiğini savunur. Bu bakış açısı, yapay zekâ çalışanlarının da benimsediği indirgemeci modele karşı çıkar ve insan bedeninin önemini vurgular. Dreyfus'un bu fenomenolojik yaklaşımı, başlangıçta alanın kurucusu olan Edmund Husserl'in fikirleri ile temellenmiştir. Ancak daha sonrasında Dreyfus, Heidegger fenomenolojini benimsedi. Dolayısıyla Dreyfus'un fenomenolojik yaklaşımı,

Heidegger ve Merleau-Ponty'nin bedenlenmiş doğamızın önemini vurgulayan ve dünya ile iç içe geçmiş insan deneyimlerini ele alan görüşlerinden etkilenmiştir.

Heidegger ortaya koyduğu fenomenolojisinde varlık sorunu üzerine odaklanır. Heidegger'e göre asıl var olan insandır ve diğer her şey sadece vardır. Heidegger, "varlık sorusunu soran, varlığı anlamaya çalışan insan varlığına, "orada-olma", "oradaki varlık" anlamına gelecek şekilde Dasein adını verir (Cevizci, 2009: 1126). Dasein, kendi varlığını sorgulayan ve anlamlandıran bir varlıktır ve en önemli özelliği "dünya içinde olmaktır". "Buna göre insan olmak bir dünyası olmaktır" (Cevizci, 2009: 1126). İnsanlar dünyada yalnızca nesnelere karşılaşmakla kalmaz, aynı zamanda bu nesnelere anlamlı ilişkiler kurarlar. Bu ilişkiler, pratik ve günlük etkileşimlerle şekillenir. Dolayısıyla Heidegger fenomenolojisinde var olan olarak insan, dünya içinde bir varlıktır ve dünya ile sürekli etkileşim içindedir.

Heidegger'in fenomenolojisinden etkilenen Dreyfus ise insan zihninin, dünya pratik, bedensel ve bağlamsal etkileşimler yoluyla anlam kazandığını iddia eder. Bilgi; soyut kurallar yerine pratik deneyim ve becerilerle elde edilir. Ona göre, insanların dünyayı anlamlandırma şekli ve etkileşimleri, sadece zihinsel bir işlem olarak değil, bedensel varoluşun bir parçası olarak da gerçekleşir. Dreyfus'a göre, gerçek zekâ, fenomenolojik bir yapıya sahip olmanın ötesinde, deneyimlerden gelen içgüdüsel bilgilerin ve sezgilerin bir birleşimiyle ortaya çıkar. Bu bağlamda, batı analitik felsefesinin derinlerinde Platonik insan doğası anlayışı olduğunu savunan Dreyfus, insanların rasyonel, bireysel failer olduğu görüşünü reddeder. Dreyfus insanların *bireysel, eylemsel ve rasyonel* olmaktan daha çok *gömülü, somutlaşmış ve bedenlenmiş* olduğunu savunur (Gehrman, 2020: 123). Hubert Dreyfus'un bu görüşü, insan deneyiminin bedensel varoluş ve çevresel etkileşimlerle sıkı bir şekilde bağlantılı olduğunu ifade eder. Deneyimlerimiz sadece öznel değildir, aynı zamanda kültürel geçmişimiz, dilimiz ve tarihsel bağlamımız tarafından da şekillendirilir. Bu, bir olguyu bağlamını anlamadan tam olarak anlayamayacağımız anlamına gelir. Dreyfus, insan anlayışının ve anlam verme sürecinin bağlamsal olduğunu vurgular. Bir şeyi anlamak, onun belirli bir bağlam içindeki yerini ve işlevini bilmekle mümkündür. Bu,

Heidegger'in "el altındalık" kavramıyla da ilişkilidir; bir nesnenin anlamı, onun günlük yaşamımızdaki pratik kullanımıyla belirlenir.

“Heidegger’e göre, dünyadaki şeyler Dasein’in belirli şekillerde kullanabileceği aletler ya da gereçler olarak var olur. Şeylerin bu şekilde, elde mevcut olan şeyler olarak değil de el altında olan şeyler olarak deneyimlenmesi, varoluşun, yani Dasein’in temel bir özelliği olarak ortaya çıkarken, onun dünya ile olan ilişkisinin teorik bir ilişki değil de tamamen pratik bir ilişki olduğunu gözler önüne serer” (Cevizci, 2009: 1128). Dolayısıyla bu özellik ile Heidegger, insanların araçları ve nesnelere nasıl kullandığını, onların işlevlerine ve bağlamlarına göre nasıl anlamlandırdığını açıklar. Örneğin, bir marangoz için bir çekiç, sadece bir nesne değil, iş yapma sürecinin ayrılmaz bir parçasıdır. Çekiç, marangozun elinde anlam kazanır ve bu anlam, sadece soyut bilgi ile açıklanamaz. Dreyfus, bu tür pratik bilgi ve deneyimlerin, yapay zekânın şu ana kadar başaramadığı bir şekilde insan bilişini şekillendirdiğini savunur.

Dreyfus, dünyayla etkileşimimizin büyük kısmının bilinçli düşünce tarafından değil, doğrudan ve somutlaşmış pratik bilgi tarafından şekillendirildiğini vurgular. Bu nedenle, zekânın birincil odağı olarak bilişten yola çıkmak ve bir eylem açıklaması geliştirmek yerine; Dreyfus, becerikli faaliyetin kendisinin insan zekânının en mükemmel biçimi ve temeli olduğu önermesiyle başlar. Dreyfus’un “becerikli başa çıkma” olarak ifade ettiği bu durum, insanların sezgisel, kavramsal olmayan eylemler yoluyla çevreleriyle etkileşim kurma biçimidir. İnsan davranışları ve becerileri otomatikleşir ve bilinçli düşünme gerektirmeden gerçekleştirilir. “Fenomenoloji, birçok uzmanlık biçiminin, eylemi yönlendirmek için nedenlere ihtiyaç duyulan bir aşamadan geçmesine rağmen, çok sayıda ilgili deneyimden sonra, öğrencinin nedenlerin rol oynamadığı bir başa çıkma yöntemi geliştirdiğini öne sürmektedir” (Dreyfus, 2005: 53). Bu başa çıkma becerisi, bisiklete binmek, klavyede yazı yazmak veya kalabalık bir caddede gezinmek gibi günlük aktivitelerde açıkça görülmektedir. Bu eylemler, soyut akıl yürütme veya resmi talimatlar yerine, dünyaya dair somutlaşmış bir aşinalığa dayanarak kusursuz ve akıcı bir şekilde gerçekleştirilir. Ustaca başa çıkma sırasında, o anın farkında olmam, bilinçli bir değerlendirmeden daha çok, dünya ile aramdaki derin bağdan kaynaklanır. Dreyfus’un belirttiği gibi,

ustaca başa çıkma, dünyayla etkileşimin en yüksek ve en temel biçimidir (Wrathall, 2014: 4). Dünya ile aramızdaki bu etkileşim, duruma yönelik eylemde bulunma becerisi kazandırır. Bilinçli bir planlama süreci olmaksızın bedenimiz durum neyi talep ederse o doğrultuda harekete geçme eğilimindedir. “Becerikli başa çıkma, somutlaşmış kapasitelerimize ve etkileşimde bulunduğumuz fiziksel ve sosyal çevrelerin özelliklerine bağlı olan doğrudan, kendini kaptıran ve kendini unutan bir tepki verme yeteneğinden oluşur” (Gehrman, & Schwenkler, 2020: 124).

fenomenolojik inceleme, çok çeşitli durumlarda insanların, eylemin neyi başarmayı amaçladığını belirleyen temsili bir durumun sürekli eşliğinde olmaksızın, dünyayla organize bir amaç doğrultusunda ilişki kurduğunu göstermektedir. Örnekler arasında tenis oynamak gibi beceri gerektiren aktiviteler; ofise gitmek veya dışlerini fırçalamak gibi alışılmış faaliyetler; yatakta yuvarlanmak veya konuşurken jestler yapmak gibi gündelik, düşünmeden yapılan faaliyetler; ve sıkıcı bir ders sırasında kişinin parmaklarını kıpırdatması ve davul çalması gibi spontane aktiviteler. Tüm bu eylem durumlarında, kişinin eylemi gerçekleştirirken ne yaptığına dair herhangi bir temsilin olmaması mümkündür (Dreyfus, 1993: 27).

İnsanlar yaşamlarında sergiledikleri her zeki eyleme belli kuralları izleyerek ulaşmazlar. Dreyfus, insanların sahip olduğu ve modellemeye uygun olmayan fenomenolojik boyuttaki bilgi birikiminin, insanların anlayış yeteneğinin kaynağı olarak görmektedir. Dolayısıyla, insanların problemleri anlama ve çözme becerisi, bilinçli hesaplamalardan daha çok, sezgi yoluyla hareket edebilmesinden kaynaklanır ve o an ki koşulda gereken eylemi gerçekleştirmeye her zaman hazır bulunurlar. Bu bağlamda, Dreyfus’a göre, insanlar, karşılaştıkları durumları bilinçli olarak analiz etmek yerine, bedenlerinin ve deneyimlerinin bir parçası haline gelmiş olan pratik bilgileri kullanarak anında tepki verirler. Dreyfus insan zekâsının ve becerikliliğinin büyük ölçüde arka planda çalışan, somutlaşmış kapasitelerden kaynaklandığını belirtir. “Arka plandaki uygulamalar dünyayı genel olarak bizim için anlaşılır kılar. Bir şey, bir olay, bir kişi, anlamlı bir amaç veya hedef vb. olmanın ne olduğunu anlamamıza yardımcı olurlar” (Wrathall, 2017: 9).

Arka plan, örtük bilgiyi kullanarak bireylerin durumlara uygun eylemler gerçekleştirmesine olanak tanır. Örtük bilgi, bilinçli olarak farkında olmadığımız, ancak günlük yaşamımızda ve pratik faaliyetlerimizde sürekli kullandığımız bilgi türüdür. Örneğin, bir marangozun bir tahta parçasını sezgisel olarak nasıl keseceğini bilmesi, örtük bilgisine dayanır ve bu bilgi, marangozluk pratiğinin arka planı tarafından şekillendirilmiştir. Bu örtük bilgi, günlük yaşamımızın ayrılmaz bir parçasıdır ve genellikle bilinçli farkındalık olmadan işler. Dreyfus, bu arka planın, sosyal ve kültürel normlar, bedensel alışkanlıklar ve pratik beceriler gibi unsurlardan oluştuğunu savunur. Bu unsurlar, bize çeşitli durumlarda nasıl davranmamız gerektiğini sezgisel olarak öğretir ve anlamlı eylemler gerçekleştirmemizi sağlar. Dreyfus, bilginin soyut mantıksal kurallarla değil, pratik deneyimlerle edinildiği fikrini savunur. Buna göre, insanlar somut durumları, deneyimleri ve örnekleri yoluyla bilgiyi anlar ve bu bilgiyi gerçek hayattaki durumlara uygularlar. Dreyfus, özellikle uzmanlık alanlarında, pratik deneyimlerin bilgi ve beceri gelişiminde temel bir rol oynadığını vurgular. Bu nedenle, karmaşık problemleri çözme veya ustalık gerektiren faaliyetlerde, deneyim ve içgüdülerin daha önemli olduğunu savunur. “Gerçek olgu, uzman olabilmek için bağımsız kurallara uymaktan daha kapsayıcı ve duruma özgü bir başa çıkma yoluna geçmemiz gerektiğini gösteriyor” (Dreyfus, 2005: 52).

İnsan becerilerinin gelişimini ve uzmanlaşma süreçlerini daha iyi anlamak ve açıklamak amacıyla Hubert Dreyfus ve kardeşi Stuart Dreyfus, ustalık kazanma süreçlerini açıklayan bir model geliştirmişlerdir. Bu modele göre, bir kişinin herhangi bir alanda uzmanlık kazanması acemilik, deneyimlilik, yetkinlik, ustalık ve uzmanlık olmak üzere beş aşamalı bir süreçte gerçekleşir. Acemilik aşamasında bireyler bağlamdan bağımsız olarak belirli kural ve yönergeleri öğrenirler ve bu doğrultuda davranışlar sergilerler. Örneğin, bir kişi ilk kez bisiklete binmeye başladığında, acemi aşamasındadır; temel kuralları öğrenir ve bu kurallara göre hareket eder. İlerleyen zamanlarda, bisiklet sürmeyi deneyimleyen kişi, bu beceriyi daha da geliştirerek kuralların hala önemli olduğu, ancak belirli durumları tanıma ve buna uygun hareket edebilme esnekliğinin kazanıldığı ileri düzey bir acemi haline gelir. Yetkinlik aşamasına geldiğinde, bireyler belirli durumlara karşı verilmesi gereken uygun tepkileri planlayabilirler. Yani sistematik bir yaklaşımla problem çözme becerileri

gelişir. Ustalık seviyesine ulaştığında ise, kurallardan daha çok deneyim ve sezgi ön plana çıkar; sezgisel karar verme yetisini kullanmaya başlayan bisiklet sürücüsü aniden çıkan engelleri hızlıca gözlemleyip tepki verir ve karmaşık trafik durumlarını rahatça yönetebilir. Son olarak, uzmanlık seviyesinde, “hepsi aynı perspektiften görülen ancak farklı taktiksel kararlar gerektiren çeşitli durumlarda yeterli deneyime sahip olan uzman, farkında olmadan bu durum sınıfını, her biri belirli bir yanıt gerektiren alt sınıflara ayırır. Bu, karakteristik olan başarılı sezgisel durumsal tepkiyi mümkün kılar” (Dreyfus, 2005: 53). Bisiklet sürme örneği ile ifade etmek gerekirse, uzmanlık seviyesine ulaşan birey, bisiklet sürmeye dair derin bilgi ve sezgiye sahip olarak, zorlu yarışlarda rekabet edebilir ve önceden görülemeyen koşullarla başa çıkabilir. Uzmanlar, durumsal farkındalıkları ve deneyimlerine dayanarak, karmaşık görevleri hızlı ve doğru bir şekilde yerine getirebilirler. Uzmanlık aşaması beceri edinim sürecinin en son aşamasıdır ve tam anlamıyla ustalaşmayı temsil eder.³ Dreyfus’un ifadesiyle, “uzmanlar asla düşünmez ve her zaman haklıdır” (Dreyfus ve Dreyfus, 1986: 31).

Uzmanlık, eylemlerine karar vermek veya eylemlerini gerekçelendirmek için kurallara ve standartlara güvenmek yerine, mevcut somut durumun taleplerini ayırt etme ve bunlara akıcı bir şekilde yanıt verme yeteneğidir (Wrathall, 2017: 11). Uzmanlık seviyesine ulaşan bireyler, deneyimleri ve pratikleriyle birlikte gelişen içgüdüsel tepkileri ile karmaşık sorunları çözme ve yeni durumlara uyum sağlama kabiliyetlerini büyük ölçüde artırır. Dreyfus’un vurgulamaya çalıştığı şey, belli aşamalardan geçerek deneyim kazanan insanın uzmanlık kazandığı noktada artık eylemlerini, rasyonalist bir zihin görüşünün savunduğu gibi kuralları işleyerek değil, sezgi yoluyla sergiliyor oluşudur. Dolayısıyla insanın zeki davranışı, sürekli işleyen kurallarla ortaya çıkmamaktadır. Güçlü yapay zekânın temel argümanı olan sembol işlemlenin, insan bilişinin belli bir kısmını oluşturduğunu söyleyen Dreyfus’a göre, insanın en önemli özelliğinin sezgisel olarak neyin nasıl olması gerektiğinin farkında olması ve sağduyu ile akıl yürütebilmesidir. “Bir uzmanın becerisi o kadar onun bir

³ Hubert L. Dreyfus ve Stuart E. Dreyfus, ustalık kazanma sürecini *Mind Over Machine* (1986) adlı eserlerinde detaylandırmışlardır.

parçası haline gelmiştir ki, bunun kendi bedeninden daha fazla farkında olması gerekmez” (Dreyfus & Dreyfus, 1986: 30).

Dreyfus’un fenomenolojik görüşleri, insan varoluşunun derinliklerini anlamamıza yardımcı olurken, bilginin ve anlamın yalnızca zihinsel süreçler aracılığıyla değil, aynı zamanda bedensel deneyimler ve çevresel etkileşimler yoluyla şekillendiğini vurgular. Dreyfus, bedensel varoluşu aracılığıyla, insanların dünyayla sürekli bir etkileşim içinde olduğunu ve bu etkileşimlerin, pratik bilgi ve becerilerin kazanımında merkezi bir rol oynadığını savunur. Onun düşünceleri, teknolojinin ve yapay zekânın insan yaşamındaki rolünü değerlendirmemizde de önemli bir perspektif sunar. Dreyfus’un fikirleri, insan zekâsının bedensel ve çevresel bağlamını önemseyen bir perspektif sunarak yapay zekânın insan zekâsını taklit etme yeteneğinin köklü bir şekilde sorgulanmasına yol açmıştır.

2.3. Dreyfus’un Yapay Zekâya Genel Yaklaşımı

Dreyfus’un yapay zekâya genel yaklaşımı, zihinsel süreçlerin karmaşıklığını ve insan deneyiminin derinliğini vurgulayan bir temele dayanır. Dreyfus, “Simya ve Yapay Zekâ” başlıklı makalesinde bu felsefi perspektifi ustalıkla açıklar. Makalede, yapay zekâ alanındaki geleneksel yaklaşımların insan zekâsını tam olarak anlayamayacağı ve simyanın sırlarını çözmeye benzediği savunulur. Öncelikle yapay zekâ araştırmasının iyimser varsayımlara dayandığını ancak bu varsayımların kendi içlerinde sorunlar barındırdığını iddia etmektedir. İyimserliklerinin temelinde ise insan bilgi işleminin dijital bir bilgisayar gibi sembol işlemekten ibaret olduğu iddiası yer almaktadır ve araştırmacılara göre doğru bir programlamayla dijital bir bilgisayarın zeki davranışlar sergilemesi sağlanabilir. Yapay zekâ düşüncesinin temel inancı, “Minsky’nin ifadesiyle, biz ‘et makineleri’ olduğumuza ve pinpon oynayabildiğimize göre, prensipte ya da pratikte metal bir makinenin de aynı şeyi yapamaması için hiçbir neden olmadığı” düşüncesine dayanmaktadır (Dreyfus & Dreyfus, 1986: 27).

Bu makalesinde, yapay zekâ alanını oyun oynama, problem çözme, dil çevirisi ve örüntü tanıma olarak dört alana ayıran Dreyfus, bu alanlardaki gelişmelere odaklanarak eleştirel bir değerlendirmede bulunmuştur. Oyun oynama alanında Samuel’in dama programı ilk yıllarda önemli bir başarı elde etti. Bu programda “Bir

hamle ağacı, son pozisyona bağlı olan bir derinliğe kadar aranır ve ardından, belirli parametrelerin değerlendirilmesine dayanarak bir hamle seçilir” (Dreyfus, 1965: 9). Bu yöntem satranç için yeterli olmayan bir yöntemdi, ancak erken dönemde başarılı satranç programları da geliştirildi. Ardından geliştirilen Newell, Shaw ve Simon’un programı iyimser iddiaları da beraberinde getirdi. Simon; on yıl içinde dijital bir bilgisayarın satranç şampiyonu olabileceği, önemli matematik teoremlerini keşfedebileceği ve kanıtlayabileceği, müzik besteleyebileceği ve psikolojik teorileri bilgisayar programlarına dönüştürebileceği (Dreyfus, 1974: 22) yönünde tahminlerde bulunmuştu ancak hedeflenen zaman içinde beklenen başarı sağlanamadı.

Newell, Shaw ve Simon tarafından geliştirilen Mantık Kuramcısı adlı yazılım ve GPS programı problem çözme alanı için önemli başarılarıdır. Ancak bu alanda da beklenen hedefler gerçekleştirilemedi. Dil çevirisi ise en zor olan alandı. Yine başarılı ilk adımların gerçekleşmesine karşın en büyük başarısızlık da bu alanda gerçekleşti. Dreyfus’un ifade ettiği gibi “Sözcüklerin parçaları, bütün sözcükler ya da sözcük grupları olsun, dilbilimsel öğelerin bağımsız olarak işlenip birbiri ardına başka bir dildeki karşılık gelen öğelere dönüştürülebileceği mekanik bir sözlüğün kolayca oluşturulabileceği baştan beri açıktı” (Dreyfus, 1965: 12). Ancak sözdizimi ve anlambilimdeki karmaşıklık ilk başarının ardından ilerleyebilmeyi oldukça zorlaştırdı. İnsan okuyucuların kolaylıkla anlayacağı bir cümle için benzersiz bir sözdizimsel yapı oluşturmanın ortaya çıkardığı zorluk, alan için oldukça problem oluşturdu.

Örüntü tanıma, yapay zekâ alanının gelişiminde kritik bir role sahiptir. Dreyfus’un belirttiği gibi, diğer alanlardaki ilerleme durgunluğunun üstesinden gelmek için örüntü tanıma alanındaki başarıyı sağlamak gerekmektedir. Bu alanda etkili çalışmalar yapılmış olmasına rağmen, mevcut durumda aşılamayan önemli zorluklarla karşılaşmıştır. Dreyfus yapay zekânın karşılaştığı zorlukları şu şekilde ifade etmiştir:

Oyun oynarken, alternatif yollar ağacının katlanarak büyümesi, takip edilebilecek yollarda bir kısıtlama gerektirir; satranç gibi karmaşık oyunlarda programlar en umut verici yolları seçemezler. Problem çözmeye sorun, seçici aramanın nasıl yönlendirileceği değil, arama sürecini başlatmak için problemin nasıl yapılandırılacağıdır. Dil çevirisinde, doğal bir dilin içsel belirsizliği nedeniyle

manipüle edilecek unsurlar bile net değildir; örüntü tanımada, her üç zorluk da ayrılmaz bir şekilde iç içe geçmiştir (Dreyfus, 1965: 46).

Yapay zekâ alanında ilerleme kaydedildiği yönündeki bir iddiaya verdiği yanıt ise Dreyfus'un mevcut duruma yönelik görüşünü ortaya koymaktadır; “[İlerlemenin] bu tanımına göre, ağaca tırmanan ilk kişi, Ay'a uçma yönünde gerçek bir ilerleme kaydettiğini söyleyebilir” (Dreyfus, 1965: 17).

Dreyfus, alana yönelik bu değerlendirmesinin ardından sembol manipülasyonunun bilgi işlemenin üç temel insan niteliğini (kenar bilinci, öz/ilinek ayrımı ve belirsizlik toleransını) sistematik olarak dışladığını görür. Kenar bilinci, bilinçli deneyimlerimize eşlik eden arka plan farkındalığını ifade eder. Eylemlerimizin ardındaki belirli ayrıntıların veya nedenlerin açıkça farkında olmaksızın, dünyayı ve içinde faaliyet gösterdiğimiz bağlamı sezgisel olarak anlamamızı içerir. Öz/ilinek ayrımı, Dreyfus'un insanların sahip olduğu ancak makinelerin kopyalamakta zorlandığı bir beceriyi tanımlamak için kullandığı bir kavramdır. Bir durum ya da nesnenin temel özellikleri ile tesadüfi ya da temel olmayan özellikleri arasında ayırım yapma yeteneğimizi ifade eder. Bu ayırım, ilgisiz veya tesadüfi ayrıntıları göz ardı ederken, belirli bir görev veya hedefle ilgili kilit yönere odaklanmamızı sağlar. Dreyfus, bu becerinin insan uzmanlığı için çok önemli olduğunu ancak makinelere programlanmasının zor olduğunu savunur. Dreyfus'a göre belirsizliğe tolerans ise, birden fazla olası yorum veya çelişkili bilgi içeren durumlarla başa çıkma ve bu durumları yönetme kapasitesidir. İnsanlar günlük yaşamlarında sıklıkla belirsiz veya muğlak durumlarla karşılaşır ve arka plan bilgilerinden, bağlamdan ve sezgisel anlayışlarından yararlanarak bu durumların üstesinden gelebilirler. Dreyfus; insanların, net ve kesin cevapları olmayan durumları anlamlandırmakta zorlanan makinelere kıyasla belirsizliğe karşı daha yüksek bir toleransa sahip olduğunu öne sürer. Dreyfus, bu bilişsel yeteneklerin insan zekâsının doğasında olduğunu, ancak makineler için kopyalama açısından zorluklar teşkil ettiğini savunur.

Başarılı bir örüntü tanıma için bu bilgi işleme biçimlerinin her biri gereklidir. İnsanın bir şeyi bütün olarak algılamasının yanı sıra makinenin belli kuralları üzerinden ve parça olarak algılaması gerekecektir. Örneğin bir meyvenin önce büyüklüğünü, sonra rengini algılayarak anlamına ulaşması gerekecektir. Ancak insan

algısının, bütünü görmesinin yanı sıra bağlamdan bağımsız da düşünülmesi mümkün değildir. Bir nesnenin anlamına bağlamından hareketle ulaşılmaktadır. Bağlam, bazen bir nesne veya kişinin tanınmasını sağlayan belirli kalıpların fark edilmesine yardımcı olabilir; ancak bazı durumlarda, bir nesne veya kişi yalnızca bağlam içinde tanınabilir ve bağlamdan ayrıldığında tanınmaz (Dreyfus, 1965: 42). Örneğin bir insanın yüzündeki bir uzvunun karakteristiğini tüm yüzüne bağlı olarak tanımlayabiliriz. Ancak, dijital bir bilgisayar programı, tanımlanmış kurallara göre tamamen belirli veriler üzerinde çalışmak zorundadır (Dreyfus, 1974: 25). “İnsan örüntü tanınmasının mekanik tanımadan radikal bir şekilde farklı olduğunun ilk göstergesi, insanların (ve hayvanların) yön ve boyut değişikliklerine, eksiklik ve bozulma derecelerine ve arka plandaki gürültü miktarına karşı toleransında görülür” (Dreyfus, 1965: 38).

Dreyfus, insanların zeki davranışlarının bağlamı içinde anlamlandırılması gerektiği görüşündedir. Bu bağlam; insanın deneyimleri, duyguları, fiziksel çevresi ve toplumsal etkileşimleri gibi birçok faktörü içerir. Dreyfus’a göre, yapay zekâ modelleri genellikle bu bağlamı ihmal eder. Örneğin, bir insanın bir olayı değerlendirmesi veya karar vermesi genellikle o olayın geçmiş deneyimlerine, duygusal durumlara ve çevresel faktörlere bağlıdır. Ancak, yapay zekâ sistemleri genellikle bu tür bağlamsal faktörleri hesaba katamazlar. Dolayısıyla, Dreyfus’un bağlam anlayışı, yapay zekânın insan benzeri zeki davranışlarını gerçekleştirmesinin zorluğunu ve yapay zekânın insan bilişini tam olarak taklit etme yeteneğinin sınırlarını vurgular. Dreyfus, insana özgü olduğunu savunduğu sınır bilinç, içgörü ve bağlam anlayışını Wittgenstein’den esinlenerek “belirgin gruplama” adını verdiği özel bir kombinasyon içerisine alır. Bu sürecin örtük şekilde olduğunu ileri süren Dreyfus için insan bilgi işlemesi çoğu durumda sezgisel bir süreçtir.

Dreyfus’a göre, sembolik mantık ve kural tabanlı sistemlere dayanan yapay zekâ yaklaşımları, insan uzmanlığının doğasında var olan incelikleri ve örtük bilgiyi yakalamakta başarısızdır. Dreyfus, insan zekâsının soyut, sembollerle manipüle edilebilir bir süreç olmadığını, aksine bedensel deneyimler, çevresel etkileşimler ve bağlamsal faktörlerle derinden bağlantılı olduğunu vurgular. Dolayısıyla Dreyfus, hesaplamalı sistemler aracılığıyla insan benzeri zekâyâ ulaşma olasılığını sorgulamış ve zekânın açık kurallar ve algoritmalarla tam olarak yakalanabileceği yönündeki

yaygın görüşe meydan okumuştur. Bilgi işlemenin zekâ için gerekli ancak yetersiz bir koşul olduğunu ve daha bütüncül ve somutlaştırılmış bir zekâ biçiminin yalnızca bir yönü olduğunu savunmuştur. Bu doğrultuda Dreyfus, ortaya koyduğu yapay zekâ eleştirisinde, insan zekâsının “modellenebileceği” şeklindeki rasyonalist anlayışı hedef almaktadır (Kenaw, 2008: 230). “Sembolik işleme modeline karşı çıkan Dreyfus, doğal bilişin sembolik bir kodla gösterilebilir bir şey olmadığını ileri sürmektedir” (Kutlusoy, 2006: 730). Dreyfus, insan zekâsının bilinçli sembolik işlemeden daha çok bilinçsiz süreçlere de bağlı olduğunu ileri sürerek, insan zihinselliğinin kurallara dönüştürülemeyen yönlerinin de olduğunu göstermeye çalışmıştır. Ona göre, zihinsel süreçler sadece mantıksal algoritmalarla açıklanamaz; insanlar duygusal, sosyal ve bedensel bağlamlar içinde anlam üretirler. Dreyfus, yapay zekâ araştırmalarının sıklıkla insan zekâsını basitleştirdiğini ve insan deneyiminin karmaşıklığını göz ardı ettiğini ileri sürer. Dolayısıyla yapay zekâ çalışmaları, insan düşüncesinin kurallara dönüştürülebileceğini varsayan rasyonalist öğretinin gerçekleştirilmeye çalışıldığı başarısız bir proje olarak karşımıza çıkmaktadır.

Ona göre, insan deneyimi sadece bilgi işleme değil, aynı zamanda anlam verme, duyguları deneyimleme ve bağlamsal içgörü gerektirir. Dreyfus, insanların dünyayı anlama ve onunla başa çıkma konusunda makineler tarafından taklit edilemeyecek benzersiz bir kapasiteye sahip olduğuna inanıyordu. Bu bağlamda, yapay zekânın tam anlamıyla insan zekâsını taklit etmesi veya onunla rekabet etmesi pek olası değildir.

Dreyfus’un “Simya ve Yapay Zekâ” başlıklı çalışması odağında yapay zekâyâ genel yaklaşımını incelediğimiz bu bölümün ardından; bir sonraki bölümde Dreyfus’un yapay zekâ eleştirisinde ileri sürdüğü sınırlılıkların bu alanda kaydedilen ilerlemeleri de göz önünde bulundurarak bir sınıflandırmasını sunmaya çalışacağız. Amacımız, Dreyfus’un eleştirilerinin son gelişmeler tarafından ne ölçüde geçersiz kılındığını görmek ve aynı zamanda geçerliliğini koruyan ve yapay zekâ için kalıcı diyebileceğimiz sınırlamaları belirlemektir. Dreyfus’un eleştirilerinin bu derinlikli analizi, yapay zekânın içkin sınırlarını görebilmemize yardımcı olacaktır inancındayız.

3. BÖLÜM: DREYFUS VE GÜÇLÜ YAPAY ZEKÂ ELEŞTİRİLERİ: BİR SINIFLAMA

Dreyfus'un güçlü yapay zekâya yönelik eleştirilerinin geçerliliğini ve yapay zekâ hedefleri üzerindeki etkilerini sistematik bir sınıflandırmaya tabi tutmak anlamlı olacaktır. Böylesi bir sınıflandırma, Dreyfus'un güçlü yapay zekâya dair ileri sürdüğü argümanların kapsamını ve yapay zekâ konusundaki felsefi tartışmalara katkısını daha ayrıntılı bir biçimde anlamamıza olanak sağlayacaktır.

3.1. Dreyfus'un Eleştirilerinin Bir Sınıflaması

Hubert Dreyfus'un eleştirileri özellikle erken dönem yapay zekâ anlayışına yöneliktir. Erken dönem yapay zekâ araştırmacılarına göre insan zihni, semboller ve kurallar kullanarak bilgiyi işler ve bu süreç, bilgisayarlar tarafından taklit edilebilirdi. Bilgisayimsal zihin kuramını temel alan bu yaklaşım, sembollerin belirli kurallar çerçevesinde manipüle edilmesiyle zeki davranışların ortaya çıkabileceğini öngörür. Bilgisayarlar, belirli kurallar kullanarak mantıksal çıkarımlar yapabilir ve problemleri çözebilirlerdi.

Dreyfus'a göre, erken dönem yapay zekâ araştırmaları, insan zihninin karmaşıklığını ve derinliğini yeterince anlamadan, semboller ve kurallar aracılığıyla insan benzeri zekâ oluşturmanın mümkün olduğunu varsayar. Bu yaklaşımlar, Dreyfus'a göre, insan deneyiminin bağlamsal, sezgisel ve bedenlenmiş doğasını göz ardı eder ve bu nedenle insan benzeri zekâyı tam olarak taklit edemez. Zira insan zekâsı bilişsel işlemlerin ötesinde sezgisel ve pratik bir yapıya sahiptir. Bu yüzden, yanlış varsayımlara dayanan yapay zekâ çalışmaları, erken dönemde başarılı ve heyecan verici sonuçlar elde etmiş olsalar da insan zekâsının bu öznel yönlerini modelleyemedikleri için kaçınılmaz olarak başarısızlığa uğramışlardır. Dreyfus, yapay zekâ araştırmacılarının bu başarısızlıklara rağmen iyimser olmalarının nedenini, rasyonalist yapay zekâ çalışmalarının temelinde yatan biyolojik, psikolojik, epistemolojik ve ontolojik dört varsayıma dayandırır. Ancak Dreyfus, araştırmacılar tarafından oldukça iyimser bir tavırla kabul edilen bu varsayımların, kendi içinde sorunları olması nedeniyle, hiç de akla yatkın olmadıklarını göstermeye çalışmıştır.

What Computers Can't Do: A Critique of Artificial Reason adlı kitabında bu varsayımları irdeleyen Dreyfus, bu dört varsayımı şu şekilde açıklamıştır:

1. Beynin –genellikle nöronları gibi olduğu varsayılan— belirli bir işlem seviyesinde, açma/kapama anahtarlarının bazı biyolojik eşdeğerleri yoluyla bilgiyi ayrık işlemlerde işlediğine dair biyolojik bir varsayım.
2. Zihnin, biçimsel kurallara göre bilgi parçaları üzerinde çalışan bir cihaz olarak görülebileceğine dair psikolojik bir varsayım. Bu nedenle, psikolojide bilgisayar, Hume gibi ampiristler (bitleri atomik izlenimler olarak kullanan ampiristler) ve Kant gibi idealistler (kuralları sağlayan programları öne çıkaran idealistler) tarafından tasarlanan bir zihin modeli olarak hizmet eder. Hem ampiristler hem de idealistler, “işlemcinin” katılımının önemli bir rol oynamadığı üçüncü şahıs bir veri işleme süreci olarak bu düşünme modeli için zemin hazırladılar.
3. Tüm bilginin formalize edilebileceğine (biçimsel bir dile ile ifade edilebileceğine), yani anlaşılabilen her şeyin mantıksal ilişkilerle, daha doğrusu Boolean işlevlerle, bitlerin birbiriyle ilişkili olma biçimini yöneten mantıksal hesapla ifade edilebileceğine dair epistemolojik bir varsayım.
4. Son olarak, dijital bilgisayarları besleyen tüm bilgilerin bit cinsinden olması gerektiğinden; bilgisayarlı zihin modeline göre, dünyayla ilgili tüm bilgilerin, akıllı davranışın üretimi için gerekli olan her şey, ilke olarak, durumdan bağımsız belirli öğeler kümesi olarak analiz edilebilir olmalıdır. Bu, var olanın, mantıksal olarak tüm olgulardan bağımsız bir dizi gerçek olduğuna dair ontolojik varsayımdır (Dreyfus, 1972: 68).

Bu varsayımlar güçlü yapay zekâ araştırmacıları tarafından hiçbir şüphe ve belirsizliğe yer bırakmaksızın kabul gördüğünden hiçbir sorgulamaya tabi tutulmamıştır. Ancak tüm bu varsayımları irdeleyen Dreyfus, biyolojik varsayımın artık kanıtlara uymadığını ve diğer varsayımların ise kavramsal zorluklara yol açtığını iddia etmektedir (Dreyfus, 1972: 137). Dreyfus’a göre araştırmacılar insan zekâsının modellenmeye uygun olmayan yönünü irdelemekten kaçınarak sonuca ulaşmayı hedefledikleri programlar yazmışlardır ve bu durum geçici çözümler getirmiştir. Ancak karşılaşılan zorluklar ve beraberinde gelen başarısızlıklar göstermektedir ki “bilgi işlem düzeyinde insan davranışları bir dizi öğenin kurallara dayalı

manipülasyonu açısından analiz edilemeyecektir” (Dreyfus, 1972: 138). Zira insan ve mekanik bilgi işleme tamamen farklı şekillerde ilerleme kaydetmektedir (Dreyfus, 1972: 137). Bu bağlamda Dreyfus güçlü yapay zekâya yönelik eleştirilerinde yapay zekânın insan benzeri bilinç ve yetenekler geliştirebileceği iddialarını derinlemesine sorgulamıştır. Dreyfus’un eleştirileri, yapay zekânın sınırlılıklarını anlamamıza ve bu teknolojinin gerçek potansiyelini gerçekçi bir değerlendirmeye tabi tutmamıza yardımcı olmaktadır.

Büyük miktarda hesaplama gücü ve veri setleri kullanarak problemleri çözme yöntemi, yapay zekânın ilk çalışmalarında belirli ölçüde başarılı olmuştur. Örneğin, satranç oynayan bilgisayarlar veya erken dönem dil çeviri programları, bu yöntemle bazı başarılar elde etmiştir. Ancak ilerleyen süreçlerde belli alanlarda karşılaşılan zorluklar, bu tür bir bilgi işleme yönteminin sınırlamalarını ortaya koymuştur. Öte yandan, günümüzde yapay zekâ çalışmalarının geldiği noktayı ve potansiyel başarılarını irdelediğimizde Dreyfus’un eleştirilerini artık geçersiz kılacak başarıların gerçekleştiğini ve gerçekleşebileceğini de görmekteyiz. Dolayısıyla Dreyfus’un eleştirilerinin günümüzdeki geçerliliğini ve beraberinde yapay zekâ araştırmalarının dinamik doğasını anlayabilmek için söz konusu eleştirileri bir sınıflandırma ile incelemek faydalı olacaktır. İlkel olarak daima geçerli olacak eleştiriler, gelecekte geçerliliğini yitirebilecek eleştiriler ve halihazırda geçerliliğini yitiren eleştiriler olmak üzere üç kategoride ele alacağımız bu sınıflandırma yapay zekânın sınırlılıklarını daha iyi almamıza yardımcı olacaktır.

Hubert Dreyfus’un eleştirilerinin sembolik ve kural tabanlı erken dönem yapay zekâ çalışmalarına yönelik olduğunu göz önüne aldığımızda, mevcut durumda yapay zekâ araştırmalarında kaydedilen ilerlemeler bazı eleştirilerin geçerliliğini yitirdiğini iddia etmek için yeterli neden sunmaktadır. Günümüzde yapay zekâ araştırmacıları, geleneksel sembolik yapay zekâdan (kural tabanlı sistemlerden) farklı olarak, veri odaklı ve öğrenme tabanlı yöntemlere dayanan, daha esnek ve güçlü algoritmalar ve teknikler kullanan bir yaklaşıma sahiptirler. Bu yaklaşım ile yapay zekânın büyük veri setleri ve güçlü hesaplama kaynakları kullanarak, karmaşık problemleri çözme ve insan benzeri zekice davranışlar sergilemesi, yani güçlü yapay zekâ görüşünün başarıya ulaşması hedeflenmektedir. Bu doğrultuda insan benzeri bir zekâ için yapay

zekâ sistemlerinin aşması oldukça güç sınırlamaların hala insanlığın karşısında olmasına rağmen, erken dönemde başarısız sonuçlara yol açan birçok çalışmanın beklenenin ötesinde ilerlemeler kaydettiği de görülmektedir. Dolayısıyla yapay zekâ çalışmalarında gelinen nokta ve umut vaat eden hedefler doğrultusunda, Dreyfus'un eleştirilerini etraflıca değerlendirerek yapay zekâ için aşılması mümkün olmayan sorunları merkeze alarak insan bilinçliliğinin kendine özgü yönlerini irdeleyeceğiz.

3.1.1. Halihazırda Geçerliliğini Yitiren Eleştiriler

Yapay zekâ alanının son yıllarda katettiği teknolojik ilerlemeler, Hubert Dreyfus'un yapay zekâ çalışmalarındaki ilk adımlarda karşılaşılan başarısızlıklar temelinde yaptığı bazı eleştirilerinin, mevcut durumda artık geçerli olmadığını göstermiştir.

Öğrenme: Yapay zekâ alanındaki gelişmelerin temelinde kendi içerisinde de oldukça büyük bir gelişim gösteren makine öğrenmesi vardır. Makine öğrenmesi ile amaçlanan, “bilgi işlem aracılığıyla verili bir girdiden arzu edilen bir çıktı elde eden, bunu da nasıl yapılacağına dair açık ve net bir tarif almaksızın yapabilen programlara ulaşmaktır” (Wooldridge, 2022: 145). Dolayısıyla kendi başına öğrenen yapay sistemler hedeflenmişti. Yapay zekâ içinde bir alt alan oluşturan makine öğrenimi kendi içinde kayda değer bir gelişim sağlayarak, bugün çok daha kapsamlı olan derin öğrenme yaklaşımını ortaya çıkarmıştır:

Derin öğrenme teknikleri, insan beyni ve sinapslarla bağlanan sayısız nöronun eşi benzeri olmayan hesaplama gücünü taklit etmek için tasarlanmış yazılım devreleri olan nöral ağ denilen teknolojiye dayanır. Nöral ağlarda pek çok basit işlemci bir araya gelir, böylece birinin çıktısı diğerinin girdisi olarak iş görür. Bu girdilerin ne kadar etkili olduğu tartışılır; ana fikir girdileri tartmak için çıktılarını kullanarak ağın kendi kendine konuşmasıdır, temelde tıpkı beyenin yaptığı gibi işi sürdürürken öğrenmesidir (Learning, 2021: 49).

Derin öğrenme, insan gibi düşünebilen bir sistem için insanın biyolojik doğasının dikkate alındığı bir çalışmadır ve problem çözme, karar alma, örüntü tanıma, oyun oynama gibi alanın ilk zamanlarında başarısızlıkla sonuçlanan bu çalışmalar makine öğrenimindeki gelişmeler ışığında artık başarıya ulaşmıştır.

Oyun Oynama: Yapay zekâ alanında kaydedilen bu gelişmeler, oyun oynama konusunda üstün performans gösteren sistemlerin geliştirilmesini sağlamıştır. Yapay zekâ için satranç, dama gibi oyunları oynayabilen sistemler geliştirmek önemlidir, çünkü yapay bir sistemin stratejik hamlelerini belirleyebilme, karar verme ve deneyimden öğrenme becerisi kazandırmak için ideal bir çalışmadır. Dreyfus, oyun teorisi ve stratejik oyunlar gibi alanlarda yapay zekânın insanların sezgisel düşünme yeteneğini ve hızlı karar alma kapasitesini tam olarak taklit edemediğini öne sürmüştür. Başlangıçta Dreyfus'u doğrulayacak başarısızlık örnekleri sergilenense de yukarıda işaret ettiğimiz gelişmeler doğrultusunda yapay zekânın stratejik oyunlarda (örneğin, Go ve satranç gibi oyunlarda) dünya şampiyonlarına karşı galibiyetler elde etmesini sağlayacak kadar önemli başarılar elde edilmiştir. Özellikle AlphaGo'nun Go oyununda dünya şampiyonunu yenmesi bu alanda önemli bir başarı örneğidir. Bu, yapay zekânın stratejik düşünme ve karar verme yeteneklerinin önemli ölçüde geliştiğini göstermiştir ve Dreyfus'un bu yöndeki eleştirilerinin bir anlamda geçersizliğini ortaya koymuştur.

Dil İşleme ve Örüntü Tanıma: Dreyfus, erken dönem yapay zekâ araştırmalarına yönelik eleştirilerinde, özellikle dil işleme ve örüntü tanıma gibi karmaşık bilişsel yeteneklerin sembolik ve kural tabanlı sistemler tarafından yeterince anlaşılamayacağını ve taklit edilemeyeceğini vurgulamıştır. Dreyfus'a göre, kural tabanlı bir sistem olan yapay zekâ, insanın zekâsının özelliklerini tam anlamıyla kavrayamayacaktır. İnsan zekâsının karmaşık doğasına vurgu yapan Dreyfus, bağlam anlayışımızın dil işleme ve örüntü tanıma için önemli bir etken olduğunu, yapay bir sistemin bu becerileri gerçekleştirmesi için bağlam farkındalığına yönelik eksikliğin önemli bir eksiklik olduğunu vurgulamıştır. Bağlam anlayışını geliştirmesi, yapay zekâ için ilkesel bir sınırlılık teşkil etse de belli ölçülerde başarılı olduğu çalışmaların mevcut olduğunu da göz ardı etmemeliyiz. Apple tarafından geliştirilen, sesle çalışan kişisel asistan Siri, dil işleme alanında önemli bir başarı olarak kabul edilmektedir (Learning, 2021: 30). Sorulara yanıt verebilen, önerilerde bulunan, talimatları izleyen Siri ile dil işlemede de ciddi bir başarı elde edilmiştir. Ayrıca IBM tarafından geliştirilen Watson adındaki bilgisayar, bir televizyon programı olan Jeopardy!'de iki insan şampiyonu yenmeyi başardı (Learning, 2021: 31). Bu gelişmeler yapay

sistemlerin derin öğrenme teknikleri sayesinde oldukça problemlili olan dil işleme konusunda oldukça etkili başarılar ortaya koymuştur. Örüntü tanıma alanında ise yapay sistemlerin görüntü tanıma ve yüz tanıma gibi uygulamalarda insanlardan daha hızlı ve doğru tanımlar gerçekleştirilebilir hale gelmiştir.

Sonuç olarak, yapay zekâ alanında kaydedilen büyük ilerlemeler, geçmişte var olan birçok sınırlamayı aşmamızı sağlamıştır. Birçok alanda öngörülenin ötesinde başarılar elde edilmiş ve insan becerilerine eşdeğer hatta bazı noktalarda çok daha iyi performans sergileyebilen sistemler ortaya çıkmıştır. Elbette bu durum insan zekâsına eşdeğer yapay sistemlerin mevcut olduğu iddiası için yeterli gelişmeler değildir. Yapay zekânın tüm potansiyelini gerçekleştirmek için hala aşılması gereken çok ciddi sınırlamalar bulunmaktadır. Ancak yapay zekâ araştırmalarının belli alanlarda sürekli gelişim gösterdiği düşünüldüğünde, ilerleyen süreçte birtakım mevcut engellerin de gelecekte aşılabileceğini düşünmemek için hiçbir neden olmadığı iddia edilebilir.

3.1.2. Gelecekte Geçerliliğini Yitirebilecek Eleştiriler

Günümüzde, yapay zekâ çalışmalarının bazı hedeflerinin gerçekleşme potansiyeline sahip olduğunu ileri sürebiliriz. Makine öğrenme ve derin öğrenme alanında kaydedilen başarılar, yapay zekâ çalışmalarında karşılaşılan bazı engellerin aşılabileceğini göstermektedir. Yapay zekâ sistemlerinin deneyimden öğrenme yeteneğini kazanması, önceden belirlenmiş kural ve algoritmaların sınırlılığında çıkarak çok daha kapsamlı hale gelmesine olanak tanımıştır. Dolayısıyla daha öncesinde önemli bir sınır oluşturan bazı durumların yeniden değerlendirilmesi, potansiyel başarıları daha iyi anlamamızı sağlayacaktır. Bu bağlamda, Hubert Dreyfus'un da yapay zekâ için önemli zorluklara işaret ettiğini düşündüğü bazı eleştirilerinin söz konusu çalışmalar doğrultusunda aşılabileceği görülmüştür.

Fiziksel Dünya ile Etkileşim: Hubert Dreyfus'un insanın bedensel varoluşu sayesinde dünya ve çevresiyle etkileşimde olduğu ve bu doğrultuda bir algılama becerisi geliştirdiği yönündeki görüşlerinin yapay zekâ için hala önemli bir konumda yer almasına rağmen, bu yönde bazı gelişmeler de kaydedilmiştir. Fiziksel dünya ile etkileşimde yapay sistemler gelişim aşamasındadır. Geliştirilen sensörler ve kameralar ile yapay zekâ sistemlerinin, çevrelerini daha iyi algılaması sağlanmaktadır. Örneğin,

otonom araçlar çevrelerini sürekli olarak tarayarak yol koşullarını belirleyebilmekte, çevredeki diğer araçları ve yayaları algılayarak güvenli bir sürüş gerçekleştirebilmektedirler. Benzer şekilde geliştirilen robotlar, çevrelerini algılayabilir ve belirlenen görevleri çeşitli sensörleri sayesinde gerçekleştirebilir. Bu tür gelişmeler, yapay zekâ sistemlerinin, teknolojinin ilerleyişi doğrultusunda çok daha karmaşık görevlerde başarılı olabileceğini göstermektedir. Bu doğrultuda, Dreyfus için oldukça önemli olan fiziksel çevreyle olan etkileşimimiz yapay zekâ tarafından belirli ölçüde gerçekleştirilebilir gibi görünüyor

Sosyal Yetenekler: Yapay zekânın insan benzeri öğrenme, duygusal tepki verme (en azından davranışsal açıdan) ve etkileşimde bulunma yetenekleri de önemli ölçüde gelişmiştir. Bu yöndeki çalışmalar, yapay zekânın insan duygularını tanıma, uygun tepkiler verme ve bağlamsal olarak anlamlı sosyal yanıtlar üretebilme yeteneğini artırmaktadır. Örneğin, doğal dil işleme ve yüz tanıma teknolojilerindeki ilerlemeler, yapay zekânın konuşma esnasında insan yüz ifadelerini algılayarak empati kurmaya benzer tepkiler geliştirmesine olanak tanımaktadır. Ayrıca, sosyal robotlar ve etkileşimli asistanlar, duygu tanıma ve uygun tepki geliştirme alanlarında kaydedilen gelişmeler sayesinde insanlar arası etkileşimlerde giderek daha doğal hale gelmektedir. Bu gelişmeler, Dreyfus'un yapay zekânın duygusal ve sosyal yeteneklerden yoksun olduğu yönündeki eleştirilerinin kısmen geçerliliğini yitirmeye başladığını göstermektedir. Ancak, yapay zekânın yalnızca davranışsal olarak benzer tepkiler verdiği ve bu tepkilerin henüz gerçek bir "duygu" içermediği göz önünde bulundurulmalıdır.

Yaratıcılık: Yapay zekâ ve insan bilinçliliği arasındaki ayrımın önemli bir konusu da yaratıcılık becerisi üzerinedir ve yapay zekânın yaratıcılık sınırlarını aşması konusu oldukça tartışmalıdır. Şu anda yapay zekâ sistemlerinin, veri setleri ile belirli kuralları işleyerek yaratıcı eserler ürettiği iddia edilebilir. Geliştirilen yapay zekâ sistemlerine belirli kavramları ilettiğinizde size bir resim yapabilir, hikâye yazabilir ya da bir şarkı besteleyebilir durumdadır. Tabi bunun gerçekten ne ölçüde yaratıcı olduğu tartışmalı bir konudur. Bazıları tarafından yaratıcı kabul edilen bu çalışmalar çoğunluk tarafından insan özgünlüğünde bir yaratıcılık olarak düşünülüyor. Yapay zekâ

sistemleri, mevcut veriler üzerinden orijinal eserler ortaya koyabiliyor ancak bu süreçte hiçbir duygu durumuna sahip değildir. İnsanlar, ortaya çıkardığı eserlerle derin bir bağ kurarlar, çünkü o eseri üretirken yaşadıkları her türlü duygu ve deneyim yalnızca o eser nezdinde gerçekleşmiştir. Dolayısıyla bir ressamın hisleriyle ortaya koyduğu eser ile yapay zekânın tasarladığı eser aynı yaratıcılık seviyesinde değildir. Gelecekte, yapay zekâ teknolojilerindeki ilerlemeler ve yeni yaklaşımlar sayesinde, belki de daha yaratıcı ve özgün eserler üretebilecek sistemler geliştirilebilir. Ancak, insan yaratıcılığının tüm boyutlarıyla yakalanması ve aşılmasının ne kadar mümkün olduğu konusunda tartışmaların bir sonuca bağlanması pek mümkün görünmüyor.

Dolayısıyla, makine öğrenimi ve sinir ağları gibi derin öğrenme teknikleri sayesinde yapay zekâ alanındaki çalışmalar etkileyici başarılar ortaya koyabilmektedir. Bu gelişmeler, yapay zekâ alanında şu anda hayal edilemeyen yeteneklerin ilerleyen zamanlarda mümkün olabileceğini göstermektedir.

Günümüzde insan hayatının pek çok alanında yapay zekâ sistemlerinin gelişimini ilgili alanla ilgili çalışmaların ilk adımlarında tahmin etmek olası bile değildir. Dolayısıyla mevcut durumda, yapay zekâ teknolojisindeki ilerlemeler her detayıyla oldukça etkileyici araştırmaların bir sonucunu oluşturmaktadır. Bununla birlikte araştırma yelpazesinin giderek büyümesiyle çok daha ciddi adımların atılabileceği de görülmektedir. Ancak bu noktada, bu çalışmanın merkezi teması olan, güçlü yapay zekâ ya da genel zekâ olarak da ifade edilebilen “insanın bilinçli zihinselliğine eşdeğer hatta insanı aşan yapay bir sistemin modellenmesi” imkanına odaklandığımızda, mevcut yapay zekâ sistemlerinin en gelişmiş versiyonlarının bile hayal edilenden oldukça uzakta olduğunu söylemek yanlış olmayacaktır. Dreyfus’un erken dönem yapay zekâ araştırmalarına yönelik eleştirilerinin temelinde, insan doğasına ait yetilerin modellerle tam olarak temsil edilemeyeceği inancı yatmaktadır. Bugün de bu öznel ve karmaşık yönlerin, mevcut yapay zekâ araştırmalarına ciddi sınırlamalar getirmeye devam ettiğini söylemek mümkündür.

Sinir ağları dijital bilgisayarinkinden tamamen farklı bir mimariye sahiptir. Bir dijital bilgisayarın merkezi işlemcisinden bir tane transistörü çıkartıp alırsanız, bilgisayar çalışmaz. Ancak bir insan beyninden kocaman bir parçayı çıkartsanız

bile, geri kalan parçalar eksik olanların görevini devralır ve beyin yine de çalışmasına devam edebilir. Ayrıca, dijital bilgisayarın tam olarak nerede düşündüğünü belirlemek mümkündür: Merkezi işlemcisinde. Oysa insan beyninin taramaları, düşünme eyleminin beyin büyük parçaları üzerine dağıldığını açık bir şekilde göstermiştir. Düşünceler bir pin-pon topu gibi oradan oraya sıçıyorlarmışçasına, farklı bölgelerin kesin bir silsile içinde aydınlandığı görülür (Kaku, 2016: 92).

Bu bağlamda, insan bilinçliliğinin doğasının oldukça belirsiz ve karmaşık oluşu göz önüne alındığında, insana özgü olan ve insana özgü olarak kalacak olan birçok yetiye sahibiz. Dreyfus'un fenomenoloji çalışmalarında ve yapay zekâ eleştirilerinde de vurguladığı şey, tamamen fiziksel varoluşa sahip insanın kendine özgü yetenekleri olduğudur ve bu yeteneklerin sembol manipülasyonuna indirgenemeyecek olan bilinçliliğimizi oluşturduğudur. Sonuç olarak, ifade ettiğimiz birçok başarı yapay zekâ alanı için devrim niteliğinde çalışmalar olsa da aşağıda değineceğimiz unsurların da işaret edeceği üzere, hedeflenen güçlü yapay zekâ tam anlamıyla gerçekleşemeyecektir.

3.1.3. İlkesel Olarak Daima Geçerli Olacak Eleştiriler

Hubert Dreyfus'un fenomenolojik görüşü temelinde ileri sürdüğü, bedensel bir varoluşa sahip olan insanın sahip olduğu bazı yetilerin, yapay zekâ alanında gerçekleşen teknolojik ilerlemelere direndiği görülmektedir. Bedensel varoluşa sahip insanın, zekâsını deneyimleri ve çevresel etkileşimleri ile geliştirdiğini savunan Dreyfus, bilinçli durumlarımızın ötesinde tamamen sezgisel olarak kullandığımız becerilerimize vurgu yapar. Bu doğrultuda insan zihninin bazı yönlerinin, güçlü yapay zekâ için aşilamaz sınırlılıklar oluşturduğu görülmektedir.

Bedensel Varoluş: Dreyfus, bir bedene sahip olarak dünyaya geldiğimizi ve bu beden ile dünya arasındaki etkileşimin zekâmızı belirleyen önemli bir faktör olduğu görüşündedir. Ona göre zekâ, sembol manipülasyonunun ötesinde, bu etkileşimler ile kazandığımız sağduyulu ve sezgisel bir süreci de içerir. Dolayısıyla, gerçekten bir bedene sahip olup bu bedensel deneyimin farkında olan bir yapay sistem geliştirilemeyeceğinden, böyle bir sistemin insanlarda olduğu gibi bilgi ve beceri

edinmesi de mümkün olmayacaktır. “Bedenlerimiz bizi anlamlı ilişkiler dünyasına yerleştirir, bu ilişkileri bizim için önemli kılar, onları (ve onlarla ilişkili olarak kendimizi) anlamamızı sağlar ve onlar içinde ve onlar aracılığıyla faaliyetlerimize rehberlik eder” (Susser, 2013: 282).

Bağlam Anlayışı: Dolayısıyla bedensel varoluşumuz aynı zamanda anlamlandırma sürecinde de önemli bir etkiye sahiptir. Fiziksel bir varlık olarak bizler, bağlam anlayışına da sahibiz ve bu bize olayları bağlamları çerçevesinde anlayabilme olanağını sağlar. Bağlam anlayışı önemlidir, çünkü gerçekten doğru anlamlandırma ancak bağlam aracılığı ile gerçekleşir. Bir ifade, farklı bağlamlarda farklı anlamlara gelebilir ya da biz ona farklı bir anlamlar yükleyebiliriz. Bu nedenle, bağlam anlayışı geliştirilmeden, yapay bir sisteme anlama becerisi kazandırmak mümkün değildir. Ancak, bağlam anlayışı kazandırmaya yönelik bir çaba, her seferinde daha geniş bir bağlam ihtiyacına yol açarak sonsuz bir gerileme problemine neden olur. İnsan, tamamen belirsiz bir durumla karşılaştığında bile bir bağlam anlayışı geliştirerek anlam kurabilir ve duruma uygun bir tepki verebilir.

Çoğu zaman eylemlerimizi düşünmeden gerçekleştirdiğimizi fark etmişizdir. Yani özellikle karşılaştığımız duruma açıklık getirerek şu eylemde bulunmamız gerekir şeklinde bilinçli bir düşünce gerçekleştirmeyiz. Zaten nasıl bir bağlam içerisinde olduğumuzu bilir ve bu doğrultuda pratik eylemlerde bulunuruz. Dolayısıyla bağlam anlayışımızın ve bu doğrultuda anlamlandırma kapasitemizin, ayrıca neyin alakalı olduğunu belirleyebilme becerimizin, zihnin sembol işleme düzeyinin ötesinde tamamen sezgisel bir doğaya sahip olduğunu görebiliriz.

İçgüdü ve Sezgisel Anlayış: Hubert Dreyfus, insanların çevrelerinde olup bitenleri içgüdüsel olarak algılayabilme becerisine sahip olduğu görüşündedir. İnsanlar, içgüdü ve sezgileri sayesinde içine gömülü oldukları dünyayı anlayabilirler. Oldukça belirsiz bir durum karşısında ne yapmamız gerektiğini içinde bulunduğumuz bağlamda değerlendirir, içgüdüsel anlayışımız ile belirleriz. Bir deprem deneyimimiz esnasında kendimizi korumaya almak için harekete geçmemiz tamamen içgüdü ve sezgisel bir süreçtir. Her adımımız planlı bir şekilde gerçekleşmeyecektir. Bu tür davranışlar, insan zihninin esnekliğini ve pratikliğini yansıtarak, belirsiz ve karmaşık

durumlara hızlı ve uygun tepkiler verme yeteneğini gösterir. Ancak yapay bir sistemi düşündüğümüzde, olası her durumun verilerine sahip olunması gerekliliğinin yanı sıra her adım için kural izlenmesi gerekecektir. Dolayısıyla yapay zekâ sistemleri genellikle bu tür içgüdüsel yeteneklerden yoksundur.

İnsanlar kurallı bir öğrenme sürecinin yanı sıra deneyimlerinden de bilgi edinen varlıklardır. Edinilen bilgiler ise bilincimizin ötesinde Dreyfus'un arka plan olarak adlandırdığı bir alanda yer edinir. Gerektiği durumlarda bu bilgilere yine sezgisel olarak ulaşırız. Buradan hareketle becerikli başa çıkanlar olarak bizler, içgüdüsel olarak eylemde bulunabilme yeteneğini geliştiririz, yani uzmanlık kazanırız. Araba kullanmayı öğrenmek öncelikle araba sürmenin tüm kurallarını öğrenmeyi gerektirir. Ancak zamanla bu süreç bilinçli olmaktan çıkar ve tamamen sezgisel bir eyleme dönüşür. Hatta başka bir şey düşündüğümüz durumda gideceğimiz mesafeyi nasıl katettiğimizin bile farkında olmayabiliriz. Bilincimiz farklı bir duruma odaklandığında dahi, eğer araba sürme becerisinde belli bir uzmanlık kazanmış isek, araba sürme eylemimiz bilinçli bir düşünmeyi gerektirmeyecektir.

Öz farkındalık: Yapay zekânın bir diğer önemli sınırlılığı, insanın farkındalık sahibi olması ile ilgilidir. Öz farkındalık insanın deneyimin öznesi olarak deneyimlediğinin de farkında olması demektir. Yani deneyim bilinçli bir özneyi gerekli kılar. Ancak mevcut yapay zekâ sistemleri bu öz farkındalık durumundan oldukça uzaktır ve buna yaklaşabilecek gibi de görünmemektedir. Örneğin geliştirilen yapay konuşma asistanlarını düşündüğümüzde, veri bankalarında kayıtlı olan ifadeleri, kodlamalarına uygun olarak sese dönüştürmekte ve dışarıdan bakıldığında gerçekten karşılıklı bir sohbet halindeymiş gibi görünmektedir. Ancak bu tür sistemlerin gerçekleştirdikleri eylemin ne olduğuna dair herhangi bir farkındalıkları yoktur. Yani yalnızca önceden belirlenmiş kuralları doğru şekilde işlemeleri söz konusudur. Dolayısıyla, insan beyninin simülasyonu konusunda kaydedilen bazı ilerlemelere rağmen, bu simülasyonlar hâlâ insan zihninin karmaşıklığını ve derinliğini yansıtmakta yetersiz kalmaktadır. Kaku'nun da belirttiği gibi;

İnsan beynini simüle etmek 2020 dolaylarında mümkün olsa da, onunla basit bir sohbete girmek pek mümkün olmayacak. Pariyetal lobu olmadığı için;

duyulardan, kendi ve başkaları kavramından ve etrafındaki dünyadan mahrum boş bir yazı tahtası gibi olacak. Temporal lobu olmadığı için, konuşmayı başaramayacak. Limbik sistemi olmadığı için, herhangi bir duygusu olmayacak. Aslına bakılırsa yeni doğmuş bir bebekten bile daha az zihinsel güce sahip olacaktır. (Kaku, 2015: 303).

Bilinç, sadece eylemleri gerçekleştirmekle kalmayıp, aynı zamanda bu eylemlerin ne olduğuna dair bir farkındalık geliştirebilmekle de ilgilidir. Bu şekilde herhangi bir öz farkındalığa sahip olmayan bir sisteme de bilinç atfetmek doğru görünmemektedir.

Bu noktada bir değerlendirme yapmak gerekirse, insan beden sahibi olarak dünya içerisine konumlanmıştır ve dünya ile sürekli etkileşim içerisinde olması zekâsının gelişmesinin en önemli etkenidir. Bu sayede bir bağlam oluşturarak, mevcut durumun neyi gerektirdiğini sezgisel olarak anlayabiliriz. Bu bağlam sayesinde anlamları oluştururuz. Ve çoğu zaman, bağlamın gereksinimlerine uygun olarak içgüdüsel bir şekilde eyleme geçeriz. Dolayısıyla insan varoluşu ve bilinçliliği sembol işlemlerin ötesinde sezgisel bir kavrayış gerektirir. Bu bağlamda düşündüğümüzde, tamamen veri işlemlemeden oluşan yapay bir sistem ne kadar gelişmiş olursa olsun, işaret ettiğimiz tüm unsurlar nedeniyle tam bir bilinç haline sahip olamayacaktır. Yapay zekâ sistemleri bağlam anlayışı geliştiremeyeceğinden, karşılaşılabileceği belirsiz durumlarla nasıl başa çıkması gerektiğini de belirleyemeyecektir. Yapay zekâ sistemleri verilerini belli algoritma ve kurallar doğrultusunda işlerken, insan zihni dünya ile doğrudan etkileşime sahiptir.

Sonuç olarak Hubert Dreyfus'un ileri sürdüğü, insanın bilinçli doğasının doğal bir sonucu olan tüm bu yetiler, güçlü yapay zekâ görüşünün insan gibi düşünebilen ve eylemde bulunabilen yapay bir sistem geliştirme hedefinin ciddi sınırlılıkları olduğunu ortaya koyar. İnsanın sahip olduğu, açıklanması zor, sezgilerimizle ilintili bu bilinçli yetiler yapay zekânın bir gün ulaşabileceği konumda değildir. Dolayısıyla teknoloji ne kadar ilerlerse ilerlesin, bilincin modelleme imkânı bulunmayan bu yönleri, insan zihinselliğinin derinliğini kavramamıza engel oluşturduğundan, gerçekten bilinç

sahibi bir yapay sistemin mümkün olamayacağını söylemek temelsiz bir iddia olmayacaktır.

3.2. Dreyfus'un Güçlü Yapay Zekâya Yönelik Eleştirilerinin Önemi

Hubert Dreyfus'un yapay zekânın gelişiminin ilk adımlarında, alanın kapsamı bile daha tam anlamıyla oturmamışken ve bir taraftan gerçekten mevcut koşullarda başarılı çalışmalar da ortaya koyulmuşken, bu derece güçlü eleştiriler geliştirmesi kendisinin oldukça fazla eleştirilmesine sebep olmuştur. Dolayısıyla Dreyfus'un çalışmalarına yöneltilen görüşler çoğunlukla olumsuz olmuştur. Dreyfus'un yapay zekâ düşüncesi ile ilgili görüşlerini ilk kez ortaya koyduğu "Simya ve Yapay Zekâ" başlıklı raporu, saygı bulmadığı gibi saçmalık olarak görülmüştür. Ona karşı görüşler, bilgisayarların neler yapabileceğine dair bir bilgisinin olmamasından dolayı yanlış değerlendirmelere ulaştığı ve bunları kesin sonuçlar olarak gördüğü yönündeydi (McCorduck, 2004: 217). Dreyfus'un bilgisayarlar hakkındaki bilgisinin yetersizliği, 'bilgisayarlar satranç oynayamazlar' gibi, mevcut durumda başarı sağlanamayan bir şeyin hiçbir zaman başarıya ulaşamayacağını varsaymasına neden olmuştur. Dolayısıyla hazırladığı raporda birçok hatalı görüş ortaya koyduğu gerekçesiyle itirazlar yöneltilmiş, hatta Simon, Newell ve beraberindeki araştırmacılar tarafından yayınlanması engellenmeye çalışılmıştır. Sonrasına da Dreyfus'a karşı cephe alınmıştır ve MIT'deki meslektaşları da raporun yayınlanmasından hoşnut olmamışlardır. Dolayısıyla çok fazla reddedilmeyle karşılaşan Dreyfus'un eleştirileri gereken saygıyı görmediği gibi ciddiye de alınmamıştır.

Ancak, geçen zaman, bu ölçüde sert tepkiler alan Dreyfus'un bazı eleştirilerinin doğruluğunu kanıtlamıştır. Yapay zekâ çalışmalarının ilk zamanlarında karşılaşılabilecek zorlukları gerekçeleriyle ortaya koyması ve dikkat çektiği noktalar oldukça önemliydi. Bu bağlamda, Dreyfus'un eleştirilerinin yapay zekâ alanında hâlâ önemli engeller teşkil ettiği göz önüne alındığında, polemik tartışmaların bir kenara bırakılması ve Dreyfus'un çalışmalarına yapıcı bir perspektifle odaklanması gerektiği açıkça görülmektedir.

Hubert Dreyfus'un fenomenolojik görüşüyle temellendirdiği yapay zekâ eleştirisi, bazı noktalarda geçerliliğini yitirmiş olsa da mevcut durumda hala insan

doğasının özgünlüğünü ifade eden eleştirilerin, güçlü yapay zekâ için önemli sınırlamalara işaret ettiği görülmektedir. İnsanın bilinçli zihinselliğinin karmaşıklığına dikkat çeken Dreyfus'un eleştirileri, yapay zekânın karşılaşıcağı zorlukları daha iyi değerlendirmemize ve anlamamıza olanak sağlamaktadır.

Dreyfus'un eleştirilerinin sembolik mantık temelli algoritmik yapay zekâ modeline yönelik olduğunu unutmamalıyız. Yapay zekâ alanındaki çalışmalar, günümüzde çok farklı teknikler geliştirmiş ve bu teknikler sayesinde çok önemli gelişmeler ortaya konmuştur. Dolayısıyla, Dreyfus'un çalışmalarında karşılaştığımız ve tümüyle insan doğasına özgü olduğunu gördüğümüz bazı unsurların, günümüzde güçlü yapay zekânın hedefi için önemli ilkesel sınırlılıklar oluşturduğunu gözlemlediğimizden, Dreyfus'un hala haklı yönlerinin olduğunu ileri sürüyoruz.

4. BÖLÜM: GÜÇLÜ YAPAY ZEKÂ VE İLKESEL SINIRLILIKLAR

Bir önceki bölümde Dreyfus'un güçlü yapay zekâ görüşüne yönelik eleştirilerini kategorik bir sınıflamaya tabi tuttuk ve bu eleştirileri sistematik bir şekilde inceleme teşebbüsünde bulunduk. Bu noktada Dreyfus'un bazı eleştirilerinin ilkece her zaman geçerli olacağına vurgu yaptık. İşte bu bölümde söz konusu her zaman geçerliliğini koruyacak olan eleştirilerin gerekçelerini ve bu gerekçelerin felsefi zeminini açıklama girişiminde bulunacağız. Bunu 6 alt bölüm şeklinde gerçekleştirmek yerinde olacaktır

4.1. Bedensel Varoluş: Gömülülük, Özümsemişlik ve Bedenlenmişlik

Güçlü yapay zekâ kavramı, insan seviyesinde ya da ötesinde zekâyâ sahip bir yapay zekâ oluşturma hedefini ifade eder. Bu hedefin gerçekleşebilmesi için, yapay zekâ sistemlerinin insan deneyimlerini ve bilişsel işlevlerini modelleyebilmesi gerekmektedir. Ancak, bu noktada, bedensel varoluşun önemi ortaya çıkar. Dreyfus'un da belirttiği gibi, insan zekâsı sadece zihinsel süreçlerden ibaret değildir; aynı zamanda bedensel deneyimlerden ve çevre ile etkileşimlerden de büyük ölçüde etkilenir. Bu bağlamda, Dreyfus'un fenomenolojik görüşlerini incelerken rastladığımız bedensel varoluşu açıklayan üç önemli kavrama işaret etmeliyiz: gömülülük, özümsemişlik ve bedenlenmişlik.

“Heidegger ve Merleau-Ponty'nin gömülü-bedenleşmiş başa çıkma anlayışı, zihnin bazen dünyaya yayılması değil, daha çok, en temel varoluş tarzımızda, hiç de zihin olmadığımız, dünyayla bir olduğumuz yönündedir” (Dreyfus, 2014: 259). Dreyfus'a göre de insan dünyaya gömülüdür. İnsanlar, yaşadıkları dünya ve çevreleriyle iç içe olan varlıklardır ve bu bağlamda deneyimlerini etkileyen unsurlarla sürekli etkileşim halindedirler. “Aralarında yaşadığımız anlamlı nesnelere, zihnimize veya beynimize saklanan dünyanın bir modeli değildir; onlar dünyanın kendisidir” (Dreyfus, 1972: 177-8). Dreyfus'a göre insanlar öncelikle dünyayla doğrudan etkileşim içinde olan pratik faillerdir. Dreyfus, dünyayla temel etkileşim tarzımızın, bağımsız, soyut akıl yürütme yerine pratik, somutlaşmış faaliyetler yoluyla olduğunu savunur. Bu bilgi birikimi, deneyim ve gündelik pratikler yoluyla edinilen, somutlaşmış, sezgisel bir dünya anlayışıdır. Dreyfus'a göre insan zekâsı yalnızca dünyada bedenlenmiş olma tarzımızın ışığında anlaşılabilir. Karakteristik

düzenlememize baktığımızda, insan zekâsının her şeyden önce ve en temelde teorik değil, pratik olduğunu görürüz (Gehrman & Schwenkler, 2020: 124).

Dreyfus'a göre, beden temelli bir varlık olarak insan, dünyayı doğrudan deneyimler ve zihinsel süreçlerini bu deneyimler üzerinden şekillendirir. Dreyfus'a göre bir bedene sahip olarak dünyaya gelmek, anlama yeteneğimiz ve zekâmızın önemli etkenlerindedir. Yaşam içerisinde kazandığımız deneyimler insanda arka plan bilgisi oluşturmakta ve Dreyfus anlamlandırabilme becerimizin işlemlemeye uygun olmayan bu yönle ilgili olduğunu savunur. "Dreyfus için bilme ve öğrenme, Merleau Ponty'nin sunduğu gibi, bedenlenmiş bir öznenin sabitlenemez ve dolaylımsız kavrayışına gönderme yapan, bu nedenle de sabitlenmiş bir zihin temsiline indirgenemez bir süreçtir" (Aydın, 2020: 88).

Bedenlenmiş öznenin dünyayla ilişkisi, yönelimsel bağlar aracılığıyla kazandığı yeteneklerden oluşur. Bu yetenekler, bedenlenmiş öznenin dünyadaki durumlara karşılık verme eğilimiyle ve bu eğilimi tekrar ederek kazandığı pratiklik ve alışkanlıklarla tanımlanır. Bedenlenmiş öznenin uyumlanma süreci, dünyayla etkileşimde bulunma ve bu etkileşimleri alışkanlık haline getirme sürecidir. Bu süreçte saklanan ya da depolanan şey, bir temsil içeriği değil, eğilimlere bağlı olarak kazanılan becerikli baş etme yetenekleri ve alışkanlıklardır. Bilme ve öğrenme yetenekleri, alışkanlıklara bağlı olarak kazanılır ve bu, sürekli bir süreçtir.

Dünya içinde gömülü ve bedenlenmiş bir varlık olarak insan aynı zamanda özümsemiş de bir varlıktır. Bu kavram, insan davranışlarının ve becerilerinin otomatikleştiğini ve bilinçli düşünme gerektirmeden gerçekleştirildiğini ifade eder. İnsanlar, birçok beceriyi zamanla öğrenerek ve tekrarlayarak özümsemişinde, bu beceriler bilinçdışı düzeyde gerçekleşir. Örneğin, bisiklete binmek, yazı yazmak veya yüzmek gibi beceriler, başlangıçta öğrenme ve pratik gerektiren süreçlerdir. Ancak, zamanla bu beceriler, düşünmeden ve otomatik olarak gerçekleştirilir hale gelir. Buradan anlaşılacağı üzere, özümsemişlik kavramı, Dreyfus'un "becerikli başa çıkma" olarak da ifade ettiği, deneyim yoluyla kazanılan örtük bilgi sayesinde uzmanlık kazanan insanların, bu uzmanlık sonucunda sergiledikleri davranışların bilinçli değil, sezgisel bir süreçle gerçekleştiği iddiasıyla alakalıdır. Dreyfus'a göre,

uzmanlık, tekrarlanan pratikler ve deneyimler sonucunda örtük bilginin içselleştirilmesiyle oluşur. Bu durumda, uzmanlar karmaşık görevleri bilinçli bir düşünme süreci olmaksızın, sezgisel ve otomatik bir şekilde yerine getirebilirler. Örneğin, bir usta şefin yemek pişirirken malzemeleri göz kararıyla ölçmesi ya da bir cerrahın ameliyat sırasında hızlı ve doğru kararlar alması, bu sezgisel ve özümsemiş bilgiye dayalı süreçlere örnek teşkil eder. Özümsemişlik, insanın çevresiyle olan etkileşimlerinde büyük bir rol oynar. Bu kavram, bireylerin belirli durumlarda hızlı ve etkili bir şekilde hareket edebilmesini sağlar. Özümsemiş beceriler, bilinçli düşünme sürecini devre dışı bırakarak hızlı ve otomatik tepkiler verilmesini mümkün kılar. İnsanlar, tekrar eden pratikler ve deneyimler aracılığıyla becerilerini geliştirir ve bu beceriler, zamanla otomatikleşir. Bu süreç, bireylerin yeni beceriler öğrenme ve mevcut becerilerini geliştirme kapasitesini artırır. “Becerilerle ilgili önemli olan şey, her ne kadar bilim, beceri gerektiren performansın kurallara göre tanımlanmasını gerektirse de bu kuralların hiçbir şekilde performansın üretilmesinde yer almasına gerek olmamasıdır” (Dreyfus, 1967: 29). Temelde, bir beceriyi gerçekleştirmek için belirli kuralların tanımlanması önemlidir. Ancak, bir kez bu kurallar öğrenildikten ve içselleştirildikten sonra, gerçek performans sırasında bilinçli bir şekilde hatırlanmalarına veya uygulanmalarına gerek yoktur. Bu durum, içinde bulunduğu durum neyi gerektiriyorsa o şekilde eylemde bulunan bir insanın eyleminin arka planında bulunan zihinsel süreçlerin tam anlamıyla özümsemiş olduğunun ifadesidir.

Bedenin yetenekleri, hareketliliği, algılamaları ve duyuları insan zekâsının temel bileşenleridir. Bedensel varoluş, insanların dünyayı deneyimleme şeklini belirler ve zekânın bedenle etkileşimi, insan davranışlarının ve deneyimlerinin temel bir parçasıdır. Bu, uzmanlığın sadece bilgi ya da zekâ meselesi değil, aynı zamanda bedensel beceri meselesi olduğu anlamına gelir. Örneğin, yetenekli bir müzisyen sadece bir enstrümanı nasıl çalacağını bilmekle kalmaz, aynı zamanda yüksek seviyede performans göstermek için gerekli kas hafızasını ve sezgiyi de geliştirmiştir. Dolayısıyla “Bedenlenmiş özne ve dünya arasında dolaylı ve temsil içermeyen bir ilişki söz konusudur. Bu bağlamda, öznenin dünyadaki şeylerle ilişkisi sabitlenemez, belirlenemez, süreç içinde ve etkileşimsel bir ilişkidir” (Aydın, 2020: 87).

İnsan, gömülü, özümsemiş ve bedenlenmiş bir varoluşa sahip bir varlıktır. Gömülü olarak, çevresiyle derinden etkileşim halindedir ve deneyimleri bu etkileşimlerin bir ürünüdür. Özümseyerek zamanla edinilen bilgi ve becerileri içselleştirir ve bilinçli düşünme olmadan da bu bilgileri kullanabilir hale gelir. Bedenlenmiş olarak ise, bedeniyle dünyayla etkileşimde bulunur ve bu etkileşimler zihinsel süreçlerini ve algılarını derinden şekillendirir. İnsan tüm bu derin varoluşuyla birlikte, karmaşık düşünme yetileri ve zekâ geliştirir. Bu bağlamda düşünüldüğünde, yapay zekânın hedeflediği insan seviyesindeki veya insan zekâsını aşan seviyede zekâ modellerinin oluşturulması önemli bir sınırlama ile karşı karşıyadır diyebiliriz.

4.2. Kenar Bilinç

Hubert Dreyfus'un insana özgü bilgi işleme biçimi olarak ele aldığı, bilincin kenar bölgeleri olarak bilinen kenar bilinç (fringe consciousness) ya da "sarkaç bilinç," genellikle dikkatimizin merkezinde olmayan, belirsiz ve arka planda kalan bilinç durumlarını ifade eder. Yapay zekâ çalışmaları için ciddi bir sınırlılık teşkil eden bu özellik, Amerikalı psikolog ve filozof William James'in çalışmalarına dayanır.

James'in bilincin doğası hakkındaki görüşleri, bilincin dinamik, sürekli değişen ve ilişkisel bir süreç olduğu yönündedir (Mangan, 2007: 673). Bilinç, ayrık ve tekrarlayan nesnelere ziyade, yoğunlukları değişen bir olaylar akışı olarak görülür. James'e göre, bilincin içerikleri iki geniş kategoriye ayrılır: Dikkatle tutulabilen ve incelenebilen deneyimler ile gözlemlenemeyen ilişki duyguları (Mangan, 2007: 674). İlk kategori, farklı duyuşsal içeriklere sahip net ve belirgin deneyimleri içerirken, ikinci kategori, soyut ve belirsiz ilişki duygularını kapsar ve genellikle dikkatli kavramamıza direnirler. "İlişkisel deneyimler (...) duyuşsal deneyimler kadar temel ve çeşitlidir. James'e göre, bilinçteki bilişsel içeriğin çoğu, belirli bir anlam duygusundan rasyonalite ve zamansal bütünleşmeye dair genel farkındalığımıza kadar, ilişki duyguları tarafından oluşturulur" (Mangan, 2007: 674). İlişkisel deneyimler, bilinçteki herhangi bir karmaşık bilişsel aktivite için hayati öneme sahiptir, ancak genellikle fark edilmezler. Bu kenar deneyimlerin belirgin bir duyuşsal içeriği yoktur ve içsel gözlemlere erişime direnç gösterirler. "Anlam duygumuz düşüncenin tamamen kendine özgü bir unsurudur. İç gözlemin (...) etrafında dönüp izole edemeyeceği ve inceleme için

tutamayacağı, zihnin o uçucu ve “geçişli” olgularından biridir. (...) Özel durumun “kenarına” aittir. (...) Kendi başına imge, çekirdek, işlevsel olarak düşüncenin en az önemli parçasıdır” (James, 1983: 446-47).

Hubert Dreyfus bu tür bir marjinal farkındalığı yalnızca durduğunda fark ettiğimiz bir saatin tik takları ya da bir arkadaşımızı ararken kalabalıktaki yüzlere dair belli belirsiz farkındalığımız şeklinde örneklendirir (Dreyfus, 1965: 21). Ayrıca Dreyfus, bilincin gerisinde kalan ve farkındalık gerektirmeyen bu yönümüzü Micheal Polanyi'nin bilincin kenarları ya da daha farklı ifadeyle bilincin saçaklarına yönelik yaptığı açıklamanın daha anlaşılır kıldığını ifade ederek Polanyi'den şu aktarımda bulunur:

Bu güç, dikkatimizin merkezindeki nesnenin etrafında belirsiz bir şekilde uzandığı için bir arka plan işlevi görme eğiliminde olan alanda bulunur. Bu şekilde gözümüzün ucuyla gördüğümüz ya da zihnimizin arka tarafında hatırladığımız bu alan, odaklandığımız nesneyi görme biçimimizi zorlayıcı bir şekilde etkiler. Hatta o kadar ileri gidebiliriz ki, bu ikincil olarak fark edilen alanın esas olarak dikkatimizi verdiğimiz nesnenin görünümünde farkında olduğumuzu söyleyebiliriz (Aktaran: Dreyfus, 1965: 22).

Kenar bilinç, insan zihninin işleyişinde oldukça incelikli ve etkili, ancak genellikle gözden kaçan bir yöndür. William James'in çalışmalarında belirttiği gibi, bu bilinç durumu, dikkatimizin merkezinde olmayan ancak bilinçli deneyimlerimizi şekillendiren belirsiz ve arka planda kalan unsurları içerir. “Bir şey hakkındaki bilgi, onun ilişkilerinin bilgisidir. Onunla tanışıklığımız, yarattığı çıplak izlenimle sınırlıdır. İlişkilerinin çoğundan, sadece onunla ilgili eklemlenmemiş yakınlıkların “saçaklarının” yarım yamalak bir şekilde farkındayız” (James, 1983: 250). İlişkiler genellikle kısmen farkında olduğumuz, açıkça ifade edilemeyen ve sadece sezgisel olarak hissettiğimiz bir “saçakta” yer alır. Bu saçak, bilincimizin kenarında bulunan ve anlamamızda önemli bir rol oynayan belirsiz bağlantılardır.

İnsan bilişselliğinin bu çalışma şekli, yapay zekânın insan benzeri bilinç seviyelerine ulaşma çabasına köklü ve çetrefilli bir engel oluşturur. Dreyfus'un “arkaplan bilgisi” olarak nitelendirdiği ve satranç ustalarının hamlelerinin sezgisel keşiflerini gerçekleştirdiğini düşündüğü yön de bu kenar bilinç olarak

nitelendirdiğimiz insana özgü ve oldukça karmaşık yönle ilgilidir. Dreyfus'un ifadesiyle; "böyle bir listeyi ararken karşılaşılan zorluklar, en azından insanlar için, tüm olası ilgili özelliklerin bir dizi veya paralel olarak ele alınıp bir tür karar verme yapmak için kullanılmadığını, bunun yerine ayırım yapmada önemli olan birçok özelliğin hiçbiri açıkça ele alınmadan, bilinç dışında kalarak işlerini yaptığını bir kez daha gösteriyor." (Dreyfus, 1972: 34). Kenar bilinç genellikle gözlemlenemeyen ve içsel gözleme direnç gösteren deneyimleri içerir ve bu bilinç hali odaklanma durumunda dahi net bir şekilde kavranmaya açık değildir. Bu durum, deneyimlerin belirgin bir duyuşal içeriğe sahip olmamasından kaynaklanır ve kenar bilincin doğası gereği belirsiz ve yakalanması zor olduğunu gösterir. İnsanlar için problem çözme, algılama, anlama, karar verme gibi birçok durumda kritik bir role sahip olan ve anlaşılması oldukça güç olan kenar bilincin yapay bir sistemde modellenmesi mümkün görünmemektedir. Doğal olarak, yapay zekâ sistemlerinin tüm yönleriyle insan bilinçliliğine sahip olabileceği yönündeki görüş zayıflamaktadır.

3.3. Bağlam Anlayışı ve Deneyim

Bağlam anlayışı, bir durum veya ifadeyi anlamlandırmak için onu çevreleyen diğer bilgilerle birlikte değerlendirme yeteneğini ifade eder. İnsanlar, doğal olarak bu yeteneğe sahipken, yapay zekâ sistemleri için bu yetenek önemli bir zorluktur. Dünya içine konumlanmış varlıklar olarak insanlar, yaşadıkları kültür ve çevre ile sıkı bir bağlam ilişkisi içindedirler ve bu sayede deneyim ve anlam geliştirirler. "Dreyfus'a göre bağlam, temelde insanların olgularla, ifadelerle, sorunlarla vb. durumlarla karşılaştıkları ortamdır; bağlamın temel özelliği, kurallarla tanımlanabilir ya da herhangi bir şekilde kolayca temsil edilebilir olmamasıdır (Andler, 2000: 141).

Bağlam, insan bilişinin ve anlamlandırma süreçlerinin merkezinde yer alır. İnsanlar, çevrelerindeki dünyayı ve bu dünyada meydana gelen olayları, sadece izole bilgilerle değil, bu bilgileri çevreleyen bağlamla birlikte değerlendirirler. Bağlam, bir olayın, ifadenin veya eylemin gerçekleştiği koşulları, geçmiş deneyimleri ve kültürel, sosyal normları içerir. Bu şekilde, bağlam, bir durumun veya bilginin anlamını zenginleştirir ve derinleştirir.

Bağlam olgusu, bilişsel ya da zihinsel yaşamlarımızın temel bir özelliğidir ve hiçbir zaman bir bağlam dışında herhangi bir görevle karşı karşıya kalmadığımız gerçeğinden oluşur: Bir sözcüğü anlamak, bir cümleyi çevirmek, bir sorunu çözmek, bir talebe uygun yanıtı karar vermek; sözcüğün, cümlelerin vb. aslında ortaya çıktığı bir bağlamdan bağımsız olarak söz konusu olamaz: İnsanlar için işaretler, talepler, görevler hiçbir zaman tek başına ortaya çıkmaz. Bu nedenle, belirli bir işarete, talebe, göreve vb. özel bir şekilde eklendiği düşünülen sabit anlam, kural, bilgi, temsil veya bilgi ne olursa olsun bağlamdan gelen “bileşenlerle” bir şekilde “karıştırılmalıdır” (Ander, 2000: 141).

Dolayısıyla bağlam anlayışımız, bilişsel yeteneklerimiz ve günlük yaşamdaki deneyimlerimiz için temel bir işlevselliğe sahiptir. Bu yetenek sayesinde, dünyayı anlamlandırır ve etkileşim kurabiliriz. Çevremizde gördüğümüz bir olayı ya da bir bireyi değerlendirirken de içinde bulunulan bağlamı dikkate alırız. Mesela, her toplumun kendine özgü normları, değerleri ve kuralları vardır. Bu bağlamda, insanlar uygun davranışları ve tepkileri öğrenir ve uygularlar. Örneğin, bir kültürde bir jest saygı anlamına gelirken, başka bir kültürde farklı bir anlama gelebilir.

İnsanlar, bağlama bağlı olarak davranışlarını ve tepkilerini uyarlayabilirler. Bu, içinde bulunulan durumu anlama ve ona uygun şekilde tepki verme yeteneğimizin bağlamla ilişkilendirildiği anlamına gelir. Dreyfus’un “alaka düzeyi” olarak adlandırdığı, gerekli olanı gereksiz olandan ayırt etme ve duruma en uygun olanı seçme yeteneğimiz, belirsizlikle başa çıkma kapasitemiz gibi pek çok becerimiz tümüyle bağlama bağımlıdır. Dreyfus’un kullandığı örnekle açıklamak gerekirse, “bıçak içeren nesnelerin ve ifadelerin alaka düzeyine dikkat çekmek ve belirsizliğini gidermek için kişinin ev içi, tıbbi veya kavgacı bir bağlamda (diğerlerinin yanı sıra) olup olmadığını bilmek gerekir” (Dreyfus, 1972: 201). Bıçağın mutfakta olması, saldırıdan daha çok yemekle alakalı olduğu anlamına gelir. Bir nesnenin varlığı, içinde bulunduğu bağlam içinde anlam kazanır. “Nesnelerin parçaları, kendilerini oluşturdukları nesnelerin içinde deneyimlenir; nesneler, buldukları yerleri doldurur; bir yer, kendi yerel çevresinde bulunur ve bu çevre de insan dünyasında olası durumların ufkunda yer alır” (Dreyfus, 1972: 174). Dreyfus’un bu ifadesi, her şeyin daha büyük bir bütünün parçası olduğunu ve anlamın bağlam içinde ortaya çıktığını vurgular.

İnsan zekâsı, karşılaşılan farklı durumları anlamak, bunları bir bağlama yerleştirmek ve bu bağlamda genelleme yaparak benzer durumlarla başa çıkmak üzerine kuruludur (Dreyfus & Dreyfus, 2014: 230). Bu yetenek, insanların karmaşık ve değişken dünya ile etkileşimde bulunabilmesini sağlar. Herhangi bir durumu veya olayı anlamak için, önceki deneyimlerden elde edilen genel kuralları ve örüntüleri kullanırız. Deneyimlerimiz ise aldığımız kararları, çevremizle olan etkileşimlerimizi ve dünyayı algılayışımızı belirler. Ancak deneyimlerin anlamı ve önemi, bağlam tarafından şekillendirilir. Örneğin, bir deneyim bir bağlamda olumlu veya anlamlı olabilirken, başka bir bağlamda farklı duygusal veya bilişsel tepkilere yol açabilir. İnsan zekâsının bu genelleme ve bağlam anlayışı, insanları çevrelerindeki karmaşıklığı anlamalarına ve değişen koşullara uyum sağlamalarına yardımcı olur.

İnsan bilişselliğinin önemli bir yönünü oluşturan tüm bu becerilerin, belirlenen kuralları izlemekten daha çok sezgisel olarak gerçekleştiği düşünüldüğünde, yapay zekâ sistemlerinin bu becerileri yakalaması olanaksız görünüyor. “Dreyfus’un argümanının özü, biçimci arzuların aksine, (1) anlamın doğası gereği bağlama bağımlı olduğu ve (2) bağlama bağımlılığın ilke olarak biçimselleştirilemeyeceğidir, çünkü bağlamlar doğası gereği belirsizdir” (Susser, 2013: 279). Yapay sistemlere bağlam anlayışı kazandırma girişiminin bir gerileme sorununa yol açacağını belirten Dreyfus’un da iddiası bu yöndedir:

Bir yanda bir tezimiz var: Her zaman daha geniş bir bağlam olmalıdır; aksi takdirde, ilgili olguları ilgisiz olgulardan ayırt etmenin bir yolu yoktur. Öte yandan, antitezimiz de var: yorum gerektirmeyen nihai bir bağlam olmalıdır; aksi takdirde, sonsuz bir bağlam gerilemesi olur ve biçimlendirmeye asla başlayamayız. İnsanoğlu bu ikilemden bir çıkış yolu sunabilecek üçüncü bir olasılığı barındırıyor gibi görünmektedir. Bir bağlamlar hiyerarşisi yerine, mevcut durum bir öncekinin devamı veya modifikasyonu olarak kabul edilir. Böylece yakın geçmişten, bir an önce ilgili ve önemli olan şeylere dayanan bir dizi beklentiyi taşıyoruz. Bu aktarım bize neyin fark edilmeye değer olduğuna dair bazı önyargılar verir (Dreyfus, 1972: 134).

“Bazı olguları anlamak için uygun bağlamı belirlemek her zaman başka, daha geniş bir bağlama başvurmayı gerektirir. Dolayısıyla sorunu biçimsel olarak ele almak

kaçınılmaz olarak gerilemeye yol açar” (Susser, 2013: 280). Dreyfus’un ifade ettiği şey, kural tabanlı bir sistemin aksine bizler, dünya ile etkileşimimizde bilgi işlemlere neredeyse hiç ihtiyaç duymayız. Bunun yerine, çevresel ipuçları, fiziksel kısıtlamalar ve deneyimleri kullanarak hareket ederiz. Bu, insan zekâsının genellikle sezgisel ve pratik bilgiye dayalı olarak işlediğini ve karmaşık dünya ile etkileşimde bulunabildiğini gösterir.

Sonuç olarak, mevcut yapay zekâ sistemleri, genellikle belirli veri kümeleri ve algoritmalar üzerine kurulu oldukları için, yeni ve beklenmedik durumlarla karşılaştıklarında uyum sağlamakta zorlanabilirler, çünkü insanların sahip olduğu şekilde bir bağlam anlayışı geliştiremeyeceklerdir. “[Yapay zekâ] gerçek bir insan zekâsının yapacağı gibi diğer bağlamlara uyum sağlamasını mümkün kılacak sağduyuya sahip olmayacaktır” (Dreyfus & Dreyfus, 2014: 230). Sağduyu eksikliği, yapay zekânın insan gibi esnek ve bağlama duyarlı bir şekilde hareket etmesini engeller.

Yapay zekâ sistemleri birçok alanda umut verici ilerlemeler kaydetse de, insan zekâsının tüm yönlerinin güçlü yapay zekâ hedefi açısından önemli bir sınırlama teşkil ettiği söylenebilir. Dolayısıyla en gelişmiş yapay zekâ sisteminin bile insan düzeyinde bir zekâyâ sahip olduğunu iddia edebilmek için aşılması gereken ancak aşılması da mümkün görünmeyen oldukça zor problemler önümüzde durmaktadır.

3.4. Fenomenal Bilinç

Bilincin gizemli doğası, onu tanımlamanın zorluğunu da beraberinde getiriyor. İnsanlık tarihi boyunca, filozoflar, bilim insanları ve düşünürler bilincin mahiyetini anlamaya çalışmışlar; ancak bu çetrefilli ve derin konu, hâlâ tam anlamıyla çözülememiştir. Bilinç, insan deneyiminin merkezinde yer almasına rağmen, keşfedilmeyi bekleyen pek çok yönüyle gizemini korumaktadır.

Bilincin doğası, onun karmaşıklığını ve farklı anlamlarını yansıtan çeşitli tanımlar ve açıklamalarla ele alınabilir. Bilinç; “bilinçli mi” sorusunun ifade etmek istediği zekâ anlamına gelebilir; bireyin uykuda olmadığını, bilinçli olduğunun bir ifadesi olarak uyanıklık anlamına gelebilir; ayrıca, bir şeyin ya da daha genel bir

ifadeyle öznel deneyimlerin farkında olma durumunu ifade eden farkındalık anlamına da gelebilir. (Arıcı, 2019: 113). Öznel deneyimler, “kişinin algılama, duyumsama ve anımsama gibi çeşitli formlarda farkında olduğu deneyimlerdir. Bilinç ise bu farkındalık deneyimlerinin hep birlikte yaşandığı süreçtir” (Alıcı, 2013: 12). Öznel deneyimlerin farkında olmak, bilincin temel unsurlarından biridir; ancak, bu sürecin bir adım ötesine geçtiğimizde, özbilinç kavramı karşımıza çıkar. Özbilinç, bireyin sadece deneyimlerinin farkında olması değil, aynı zamanda bu farkındalığın farkında olması anlamına gelir. “Bilinçli bir varlığın bilinçli bir varlık olduğunun farkında olması” (Arıcı, 2019: 113) anlamında kendilik bilinci şeklinde de ifade edilebilir.

Ned Block, bilinci kategorize ederek bir “fenomenal bilinç” tanımı ortaya koyar: “Fenomenal bilinç deneyimdir; bir durumun fenomenal olarak bilinçli olması, o durumda olmanın nasıl bir şey olduğu ile ilgilidir. Buna karşın, erişim bilincinin işareti; akıl yürütme, rasyonel olarak konuşma ve eyleme rehberlik etmede kullanılabilirliktir” (Block, 1995: 227). Ned Block’un bu iki bilinç türünü tanımlaması, bilincin farklı boyutlarını anlamamıza yardımcı olur ve bilincin karmaşıklığını ortaya koyar. Erişim bilinci ise işlevseldir ve bilgiyi kullanma yeteneğiyle ilgilidir. Fenomenal bilinç ise, tamamen öznel ve kişisel deneyimlerle, bu deneyimlerin nasıl hissettirdiği ile ilgilidir. Block fenomenal bilincin; düşünce, arzu, duygu, duyum ve algıların deneyimsel özelliklerini içerdiğini ileri sürer (Block, 1995: 230).

Deneyim ile ilişkilendirilen “fenomenal bilinç,” “tecrübi karakter” (experiential character), “öznel his” (subjective feel), “ham duyusal his” (raw sensory feel)” (Arıcı, 2014: 4) gibi farklı terimler kullanılarak anlatılmaktadır. Ancak, fenomenal bilincin bilinen en meşhur tasviri Thomas Nagel tarafından kullanılan ve *nasıl bir şey olmak* anlamına gelen “what-it-is-likeness” terimidir. Nagel, *Yarasa Olmak Nasıl Bir Şeydir?* adlı makalesinde, bilinçli durumların öznel olduğunu ileri sürerek fizikalist bir bakış açısıyla bilinemeyeceğine vurgu yapmaktadır. Yarasa olmanın nasıl bir şey olduğu ancak o yarasa olmakla bilinebilir bir durumdur. Dolayısıyla bir organizma için *o olmaklık* fikri anlamlıysa, bilinçli olduğu ileri sürülebilir; “bir organizmanın bilinçli bir deneyim yaşadığı gerçeği, her bakımdan, temelde o organizma olmak gibi bir şeyin varlığını gösterir” (Nagel, 2008: 378).

Makalesinde yarasa olmanın nasıl bir şey olduğunu anlama girişiminde bulunan Nagel, öncelikle herkesin bildiği gibi yarasanın, memeli, görmediği için sonar kullanan bir hayvan olduğunu belirtir. Düşünüldüğü zaman yarasa olmak, çok az görme yetisi olan ve dışsal seslere göre hareket ederek uçabilen bir hayvandır. Ancak ifade ettiğimiz tüm bu deneyimleri biz yarasa olsaydık düşüncesiyle deneyimlemeye çalışıyoruz. Deneyimlerin öznel karakteri, onları algılama alanımızın ötesinde bırakır; “eğer evrenin bir yerinde bilinçli yaşam varsa, bunun bir kısmı bildiğimiz en genel deneysel tanımlarla açıklanamayabilir” (Nagel, 2008: 380). Dolayısıyla anlamaya çalıştığımız şey o öznenin kendisi için öyle olmanın nasıl bir deneyim olduğudur. Nagel, yarasaların sonar kullanarak çevrelerini algılama yeteneklerinin insanlar tarafından tam anlamıyla anlaşılamayacağını savunur. Yarasaların bu deneyimlerinin nasıl bir şey olduğunu anlamak, yarasa olmanın öznel deneyimini gerektirir, ki bu insanlar için imkânsızdır. Bilinci problemlili kılan da bu öznel yönüdür. “Beynin nesnel bir tasvirinin, bu tam bir tasvir olsa bile, beynin izin verdiği öznel deneyimi nasıl olup da izah edebileceğini tahayyül edemeyiz” (Zeman, 2006: 382). Bu durumda, yalnızca birinci şahıs perspektifine açık olan bilincin öznel yönü, bir deneyimleyeni zorunlu kılar. Dolayısıyla fenomenal bilincin deneyimleyeni olarak fenomenal özne deneyimin ön koşuludur.

Fenomenal deneyimimizin iki karakteristik özelliğinden biri olan öznellik, Nagel tarafından da açıklandığı üzere, deneyimimizin yalnızca bize özel olduğu, bizim için o deneyimin nasıl olduğuyla ilgilidir. Fenomenal deneyimin diğer önemli bir özelliği ise nitelikselliktir. “her deneyimimizin kendine özgü nitelikleri vardır” (Arıcı, 2019: 115) ve bu nitelikler “qualia” olarak adlandırılır. Örneğin, çikolata yediğimizde algıladığımız tadı düşünelim; yaşadığımız bu deneyimin bize hissettirdiği ve yalnızca bize özel fenomenal nitelikleri (qualia) vardır.

Birkaç arkadaşınızla yeşilliklerle çevrili yüksek kayalık bir dağın aşağıya doğru akan bir şelaleye karşı oturduğunuzu, bu manzarayı seyrettiğinizi düşünün. Kayalıkların altında berrak bir göl oluşmuş ve kenarlarında rengarenk kır çiçekleri yetişmiş. Bu manzaraya karşı hissettiğiniz şey, sizin fenomenal deneyiminizin nitelikleridir ve yalnızca size özgüdür. Yani yanınızdaki diğer arkadaşlarınızın

manzara karşısındaki deneyimi sizin deneyiminizden farklı olacaktır. Herkes için bu manzara deneyiminin fenomenal nitelikleri farklılık gösterecektir. Hatta sizin o an o manzaraya karşı hissettiğiniz şeyler başka bir manzaraya karşı hissettiklerinizle de aynı olmayacaktır. Fenomenal özne olarak, yaşantımız yalnızca bizim bilinçli deneyimlerimizdir ve deneyimlerin öznel nitelikleri yalnızca bize özgüdür. Acı hissi, susuzluk, kaygılanmak, mutluluk, kırmızı rengi, algılayabilmek, çikolatanın ağzımızda bıraktığı tat, diş ağrısı ve daha birçok zihinsel durum yalnızca onu deneyimleyen özneyle aittir.

Bu bağlamda değerlendirecek olursak, fenomenal bilinç iki karakteristik özelliği ile anlaşılabilir görünür: Öznellik ve niteliksellik. “Öznel bir karaktere sahip olması, benim belli bir zamanda yaşadığım deneyimin aynısına bir başkasının benim deneyimim olarak maruz kalamayacağı anlamına gelir. Niteliksel bir karaktere sahip olmak, deneyimlerimin, örneğin kırmızı bir güle yönelik algımın, gözlerimi kapatıp aynı kırmızı gülü gözümde canlandırdığımda, yani kırmızı gülün yokluğunda bile hala sunulan kırmızılık ve opaklık gibi “qualia” adı verilen belirli ayırt edici nitelikler sunduğu anlamına gelir” (Arıcı, 2018: 16).

Dolayısıyla, birinci şahıs perspektifinden yaşanan zihinsel deneyimlerimiz, bizi diğer insanlardan farklı kılar ve her birimizin kendine özgü bir içsel dünya taşıdığını gösterir. Bizler, kendi zihinsel deneyimlerimizi doğrudan yaşayan bireyler olarak, kendi düşüncelerimizi, duygularımızı ve algılarımızı ilk elden tecrübe ederiz. Bu deneyimlerimiz, bizim iç dünyamızın benzersiz ve öznel karakterini oluşturur. Diğer insanların zihinsel durumları hakkında doğrudan bilgi sahibi olamamamıza rağmen, kendi zihinsel durumlarımızın varlığını inkâr etmek mümkün değildir. Bu durum, zihin felsefesinin temel problemlerinden biri olan kendilik (self) ya da “ben hissi” olarak da ifade edilebilen probleme karşılık gelir. Bu konu hakkında farklı görüşler olsa da en sarsıcı ifadeler David Hume tarafından ortaya konmuştur. Hume’a göre, “ben” bir yanılsamadan ibarettir; ben dediğimiz şeyi bulmak üzere zihnimizin içine yöneldiğimizde, orada bulacağımız şey, bağımsız ve ayrık bir ben değil, yalnızca algı demetleridir (Arıcı, 2014: 66). Hume’un içgözlem yoluyla ulaştığı sonuç, bilincin bağımsız bir kendilikten yoksun olduğunu öne sürer: “Bilinçli bir insanın zihinsel yaşamında, hâlihazırdaki algı parçalarından bağımsız herhangi bir kendilik ya da ben

yoktur. Her bir insanođlu, fenomenal anlamda bir algı demetinden öte bir şey değildir” (Arıcı, 2015: 67). Hume’a göre, zihnimize bağımsız bir şekilde var olan bir kendilik ya da benlik bulamayız; yalnızca sürekli deđişen duyular, algılar ve düşüncelerle karşılaşırız. Ancak, Hume’un bu görüşü, bizim günlük deneyimlerimizde sürekli olarak hissettiğimiz ben hissi ile çelişmektedir. Günlük yaşamda, her birimiz kendimizi süreklilik ve tutarlılık içinde deneyimleriz. Bu deneyim, deđişen algılar ve duygular arasında sürekliliđi sađlayan bir “ben” hissi içerir. Kendiliđin veya benin tözsel varlıđı ile ilgili bu tartışmanın detayları bir tarafa, burada Humecu inkâr haklı olsa bile, insan bilincinin fenomenal gerçekliđi ve bunun yapay bir sistemde modellenemesinin mümkün görünmemesi yadsınamaz bir gerçek olarak karşımızda durmaktadır.

Tamamen fiziksel yapıya sahip olan beynin fizyolojik süreçlerinin, nasıl olup da tüm bu yönleriyle bilinci, daha özeldede fenomenal bilinci ortaya çıkardığını henüz anlayabilmiş de deđiliz. Richardo Manzotti, tahayyül ettiđi bilinç açıklamasıyla bu durumu şöyle ifade ediyor:

Bence aynı derecede güçlü ve sađduyuya dayalı ama birbiriyle çelişen iki sezgiyle karşı karşıyayız. Kendimizin, benliğimizin bedenlerimizle aynı yerde ve büyük ihtimalle de bedenlerimizin içinde olduğunu hissediyoruz. Öte yandan, insan bedeninin içine baktığımızda gördüklerimize benzer bir şeyden oluşuyormuş gibi hissetmiyoruz. Bilinçli deneyimimizin bu hücrelerden, zarlardan, kaslardan, yağ ve bedensel sıvılardan oldukça farklı bir tabiatı var (Manzotti, 2018: 38).

Fenomenal özne olarak sahip olduğumuz niteliksel deneyimler, öznel perspektif ve anlamlandırma kapasitesi şeklinde ifade edebileceğimiz özellikler, insan doğasında bilincin benzersiz yönünü vurgular. Tamamen fiziksel bir sistem olan yapay zekâ sistemleri ise bilincin bu yönünü simüle edebilme şansına dahi sahip görünmemektedir. Oyun oynama alanında satranç programlarının ciddi başarılar elde ettiđini ve Dünya satranç şampiyonlarından ünlü Kasparov’u yendiğinden söz etmiştik. Burada bu başarıya farklı bir açıdan yaklaşmak istiyoruz; Bir insan oyuncunun oyun boyunca veya kazandığı ya da kaybettiđi anda yaşadığı öznel deneyim ile bu deneyimin deneyimsel nitelikleri (qualia’sı) dikkate alındığında, yapay

bir sistemin de aynı şekilde kazanma hissiyatına, bu hissiyatın deneysel niteliklerine sahip olacağını iddia edebilir miyiz? Pekâlâ bu pek olası görünmüyor. Burada önümüzdeki en iyi ihtimal, yapay zekâ sisteminin kazanma veya kaybetme durumunda insan oyuncunun sergileyebileceği davranışı taklit etmesi olacaktır. Ancak gerçek bir başarı veya başarısızlık hissi deneyimlemesi, bu türden bir deneyiminin deneysel nitelikleri (quaila'sı) ile muhatap olması pekâlâ mümkün görünmemektedir. Bu da göstermektedir ki fenomenal bilinç ve bu bilincin nitelikleri, güçlü yapay zekâ veya genel yapay zekânın önündeki en büyük bir engellerden biridir ve ilkece aşılacak bir engel olarak görünmemektedir.

3.5. İçgüdü ve Sezgisel Anlayış

İnsan bilinci, yalnızca bilgi işleme kapasitesiyle sınırlı olmayan, derin ve karmaşık bir yapı taşır. Bu yapı, içgüdü ve sezgisel anlayış gibi tamamen insanî ve öznel deneyimlerle şekillenir. İçgüdü, evrimsel süreçler boyunca geliştirdiğimiz ve hayatta kalmamızı sağlayan temel davranış biçimlerini içerirken, sezgisel anlayış ise bilinçli farkındalığın ötesinde, hızlı ve çoğu zaman bilinçsizce verilen kararların temelini oluşturur. Çevremizdeki uyarılara verdiğimiz otomatik tepkiler içgüdüselidir. Düşünme sürecini içermeyen hızlı ve refleksif tepkilerdir. Örneğin korku anında verdiğimiz tepki içgüdüselidir; öğretilmiş bir tepki değildir. Yalnızca canlılara özgü bir davranış biçimi olduğunu ileri sürebileceğimiz içgüdü yetisinin yapay bir sistemde gerçekten var edilmesi mümkün görünmemektedir. Bizim için bile öğrenilmiş olmayan bu tepkiler yapay zekâ sistemlerinde olası durumlar için olası davranışlar şeklinde programlanacaktır. Ancak daha karmaşık olan ve güçlü yapay zekâ hedefi için daha büyük bir problemi oluşturan şey, sezgisel anlayış biçimimizdir.

İnsan zihni, bilgiye mantık, analiz ve akıl yürütme gibi bilinçli süreçlerin yanı sıra sezgi yoluyla da ulaşır. Sezgisel bilgi ile bilinç halini gerektirmeden ulaşılabilen bir bilgi türünü ifade ederiz. Bazı durumlarda, analitik yaklaşımlar her zaman gerekli bilgileri sağlamayabilir veya zamanında sonuç vermeyebilir. Bu tür durumlarda, sezgisel anlayış, hızlı ve etkili kararlar alınmasına yardımcı olur. Dreyfus'un da önem verdiği sezgisel anlayış, arka planda yer alan bilgilere zahmetsizce ulaşabilmemizi ifade eder. Dolayısıyla Dreyfus, birçok eylemimizi bilinçli bir düşünceye ihtiyaç

duymadan gerçekleştirdiğimizi savunur. Ona göre, çeşitli yollarla edindiğimiz bilgileri bilincimizin saçaklarına kaydeder ve arka planda bir birikim oluştururuz. Bu birikim sayesinde, bilincimiz devreye girmeden, bir tür "aydınlanma" olarak tanımlanabilecek sezgi yoluyla bu bilgilere ulaşırız. Dreyfus insan bilincine özgü gördüğü bu beceriyi “know-how” şeklinde ifade eder; “Sezgi ya da bilgi birikimi dediğimizde, daha önceki deneyimlerle benzerliklerin görülmesiyle zahmetsizce oluşan anlayıştan bahsediyoruz. “Sezgi” ve “know-how” kelimelerini eşanlamlı olarak kullanacağız, ancak bir sözlük bunları birbirinden ayıracak ve “sezgiyi” tamamen bilişsel faaliyetlere, “know-how”ı ise bedensel bir becerinin sıvı performansına atayacaktır” (Dreyfus ve Dreyfus, 1986: 28). Tsai’ye göre, yetenekli veya akıllı eylemleri açıklamak için kullanılan “know-how” kavramı, Dreyfusçu anti-entelektüalizme göre “bunu bilmek” olarak tercüme edebileceğimiz “know-that” terimleriyle açıklanamaz; sadece sezgi terimleriyle açıklanabilir (Tsai. 2022: 1). Dreyfus bu kavramı bisiklete binme örneği ile açıklar:

Bisiklete binebiliyorsunuz çünkü “know-how” denilen, pratik yaparak, bazen de acı veren tecrübelerle edindiğiniz bir şeye sahipsiniz. Öğrendiklerinizi kelimelere dökememeniz, bilgi birikimine olgular ve kurallar biçiminde erişemeyeceğiniz anlamına gelir. ... Hepimiz bisiklete binmek gibi “bilmekle” sınırlı olmayan sayısız şeyin nasıl yapılacağını biliyoruz (Dreyfus ve Dreyfus, 1986: 16).

Belirli bir şeyi bilmek (veya önermesel bilgi) olarak da ifade edilebilecek olan “know-that” kavramı, matematiksel formüller ya da kurallar gibi bilgileri kapsar. Sezgi ile eşdeğer olarak kullanılan “know-how” (yordam bilgisi) kavramı ise, Dreyfus tarafından, pratik ve deneyim yoluyla öğrendiğimiz bilgileri ifade etmek için kullanılmıştır. Dreyfus bu kavram ile bisiklete binmek gibi becerileri nasıl bildiğimizin sözcüklerle ifade edilemeyecek bilgiler olduğunu da açıklar. Bu içgüdüsel ve sezgisel bir anlayış geliştirmeyi gerektirir.

İnsanların içgüdüsel ve sezgisel deneyimleri, onların dünyayı algılama ve anlamlandırma biçimlerini derinlemesine etkiler. Örneğin, bir tehlike anında hissettiğimiz korku ya da birinin samimiyetini sezdiğimizde yaşadığımız güven

duygusu, sadece mantıksal analizlerle açıklanamaz. Bu tür duygusal ve sezgisel tepkiler, bireysel deneyimlerin, kültürel bağlamların ve biyolojik faktörlerin bir birleşimidir.

Dolayısıyla, Dreyfus'un da iddia ettiği gibi, insan zihni ve bilgisayarın çalışma biçimi arasında ciddi farklar bulunmaktadır. Yeniden satranç örneğini ele alacak olursak, “bir insan satranç oyuncusu en iyi hamleye binlerce pozisyonu gözden geçirerek ve olası her hareketin sonuçlarını inceleyerek ve hesaplayarak karar vermez. Bunun yerine, bir büyük ustanın sezgiyi kullanarak ne yapacağına karar vermesi sadece birkaç saniye alabilir” (Fearn, 2005: 51). Usta oyuncu, yapacağı hamlenin yanlış sonuçlarını sezgisel olarak bilir ve hamlelerini bu doğrultuda gerçekleştirir. Buna karşın sezgisel anlayışa sahip olmayan yapay zekâ sistemi, her hamlenin olası her sonucunu zahmetlice işlemek zorundadır. Yine de, yapay bir sistemin sezgiye gerek duymaksızın uyguladığı bu yöntemle insan oyuncularını yendiği, dolayısıyla sezginin burada kritik bir unsur olmadığı ileri sürülebilir. Burada satranç oyunundaki durumun ya da benzer diğer durumların sınırlı bağlama sahip olan durumlar olduğu göz ardı edilmemelidir. Olası tüm sonuçlar bir sistemde modellenilebilir olduğu zaman elbette gelişen teknolojinin de getirisiyle başarı elde edilebilir olacaktır. Ancak, sınırsız bağlama sahip durumları, yani günlük hayatta içinde bulunduğumuz tüm durumları düşündüğümüzde yapay zekâ sistemlerinin kuşatamayacağı bağlamlarla karşı karşıya kalacağı açık bir şekilde ortadadır.

İnsanın sahip olduğu sezgisel anlayış bağlamla ilişkilendirildiğinde, mevcut durum için neyin gerekli, neyin gereksiz olduğunu belirleyebilme imkanını da doğurur. “Sağduyu, genelleme yapma becerisini gerektirir ve bu tür yargılar, bir seferde tek bir bilgiye bir dizi kural uygulamak yerine, bir duruma bütünsel olarak bakılarak yapılır. Bilgi ile “know-how” arasındaki fark budur” (Fearn, 2005: 51). İnsan bilincinin doğal bir sonucu olan sezgisel anlayışa sahip olmayan bir yapay sistemin bu alaka düzeyini belirleme imkânı da yoktur. Dreyfus'un ifade ettiği gibi, bilgiyi temsil etmek ve işlemek için semboller ve kurallar kullanan yapay bir sistemin, pratik beceriler ve deneyim ile edinilen bu bilgi biçimini yakalaması olanaksızdır. Zira “sezgiler sistemleştirilemez ve bu nedenle, gerçekten, herhangi bir algoritmik işlemin dışında kalmaktadırlar!” (Penrose, 2000: 131).

O halde, içgüdü ve sezgisel anlayış gibi derin ve öznel insan deneyimlerinin, bir makine tarafından tamamen modellenmesi ihtimali olası görünmemektedir. Günümüzde geliştirilen yapay zekâ sistemleri belirli görevlerde üstün performans gösteriyor olsalar da, güçlü yapay zekâ görüşünün hedeflediği, insan benzeri bilince sahip bir sisteme ulaşmak, bu açıdan bakıldığında gerçekleştirilebilir bir hedef olarak görünmemektedir. Zira bunun için insan bilincinin doğasında içkin olan içgüdüsel ve sezgisel bilme ve eylemde bulunma yeteneğinin de modellenmesi gerekmektedir.

3.6. Öz Farkındalık ve Anlama

Özbilinç kavramını ile bilinçliliğimizin bilincinde olma durumunu ifade ettiğimize yukarıda işaret etmiştik. Bu tanımlamadan ayrı olarak öz farkındalık; “kişinin dünyayla iç içe geçmiş bir fail olarak kendisinin farkında olmasıdır” (Rousse, 2019: 163) diyebiliriz. Burada önemli olan şey, kişinin kendisinin bir birey olarak farkında olup olmadığıdır (Smith, 2024). Yani, bu farkındalık, kişinin bilinçli bir şekilde kendisini düşüncelerinin ve deneyimlerinin merkezinde görmesini içerir. Bu kavramı daha iyi anlamanın yolu “öndüşünümsel öz farkındalık” ve “yansıtımlı öz farkındalık” olarak adlandırılan farklı özbilinç biçimleri arasındaki ayrımı kavramaktır.

Yansıtımlı öz farkındalık, kişinin kendini bilinçli olarak düşünme süreciyle değerlendirmesidir. Kişinin duygularını, eylemlerini ve deneyimlerini değerlendirmesi ve analiz etmesidir. Tamamen bilinçli bir farkındalık biçimidir. “Bana sağ ayağımda hissettiğim acının veya az önce düşündüğüm şeyin bir tanımını vermeme isterseniz, bunun üzerine düşünür ve böylece acıdan veya düşünceden bir sıra uzakta olan belirli bir bakış açısı edinirim. Dolayısıyla, yansıtımlı öz-bilinç en azından ikinci dereceden bir biliştir” (Gallagher & Zahavi, 2023).

Öndüşünümsel öz farkındalık ya da başka bir ifadeyle yansıtma öncesi öz farkındalık, “kişinin deneyimleyen bir özne olarak kendisinin farkında olmasıdır” (Nida-Rümelin, 2017: 68). Yansıtma öncesi öz farkındalık, kişinin bir deneyimi yaşarken, kendisinin bu deneyimi yaşayan bir özne olduğunun bilinçli olmadan farkında olmasıdır. Bu tür farkındalık, otomatik ve doğrudan bir bilinç halidir. Öz

farkındalık kavramıyla ilgilenen Gallagher ve Zahavi öndüşünümsel öz farkındalığı; “(1) deneyimimiz üzerinde herhangi bir yansıtma yapmadan önce sahip olduğumuz bir farkındalıktır; (2) açık veya daha yüksek dereceli bir özbilinç biçiminden daha çok örtük ve birinci dereceden bir farkındalıktır” şeklinde açıklar (Gallagher & Zahavi, 2023). Dolayısıyla, öndüşünümsel öz farkındalık, deneyimleyen özne olmanın temel bir parçasıdır. Bu farkındalık, kavramsal olmayan, yani düşünce veya dil kullanmadan, doğrudan bir farkındalıktır. Bu, kişinin “ben buradayım ve bu deneyimi yaşıyorum” şeklinde bir hissiyatıdır.

Dolayısıyla, öndüşünümsel öz farkındalık, kişinin deneyimlerinin öznesi olduğunu bilinçli bir şekilde farkında olmadan yaşamasıdır. Bu, doğrudan ve kavramsal olmayan bir bilinç hali sunar. Bu bağlamda, bedensel öz farkındalık önemli bir tamamlayıcı rol oynar. Bedensel öz farkındalık, vücudumuzun içsel durumlarına, hareketlerine ve pozisyonlarına dair örtük bir bilinç hali olarak ortaya çıkar. Bu tür bir farkındalık, vücudumuzun eylemlerini ve durumunu bilinçli olarak düşünmemize gerek kalmadan, doğrudan ve otomatik bir şekilde deneyimlememizi sağlar. Örneğin, bir topu yakalamak için zıpladığımızda, vücudumuzun ne yaptığına dair belirgin bir bilinç hali olmayabilir; ancak yine de vücudumuzun hareketleri ve pozisyonları hakkında örtük bir bilgiye sahip oluruz. Vücudumuzun hareketlerine dair bu örtük farkındalık, deneyimlerimizin ve eylemlerimizin temel bir parçasıdır ve özne olarak deneyimlediğimiz gerçekliği daha derinlemesine anlamamıza yardımcı olur. Bu farkındalık, yalnızca zihinsel bir işlem değil, aynı zamanda bedensel bir durumun bilinçli bir şekilde yaşanmasıdır.

Deneyimleyen bir özne olmayı gerektiren öz farkındalık, özbilinç yetimizin temel bir bileşeni olarak oldukça karmaşık bir süreçtir. Ancak, bilinç sahibi bir varlık olarak insanın önemli bir becerisini oluşturur. İnsanların düşüncelerinin, hislerinin ve bedensel tepkilerinin farkındalığı, onların hem zihinsel hem de fiziksel durumları hakkında derin bir bilgiye sahip olmalarını sağlar. Ayrıca, insanlar bilinçli düşünce olmadan da otomatik tepkiler verebilirler, örneğin bir elin sıcak bir yüzeyden hızla çekilmesi gibi. Ancak, yapay zekâ sistemleri bu tür öznel deneyimlere ve bilinçli farkındalığa sahip olamazlar. Yapay zekâ, belirli görevleri yerine getirme konusunda

elbette etkili olabilir, fakat bu süreçlerde öznel bir deneyim yaşayıp yaşamadığı sorusu oldukça anlamsız görünmektedir. Bu durum, insan deneyiminin derinliğinin ve zenginliğinin, yapay zekânın ulaşabileceği bir özellikten çok daha fazlasını içerdiğini ve güçlü yapay zekânın insanların bilinçli farkındalığını tam anlamıyla modelleyemeyeceğini göstermektedir.

Son tahlilde, yapay zekâ çalışmalarının en başından beri aşılması muhtemel görünmeyen en önemli sınırlılığın, yani semantiğin üretilmemesinin, yukarıda ele aldığımız insan bilincinin farklı yönlerinin doğal bir sonucu olduğu görülmektedir. Bu nedenle, yapay zekânın bu yetiyi tam anlamıyla edinmesi, bilince sahip olma seviyesine ulaşmasıyla ilişkilendirilebilir. Tüm yönleriyle insan bilincinin nihai düzeyine ulaşması mümkün görünmeyen yapay zekâ sistemlerinin anlam probleminin üstesinden gelemeyeceği aşikâr görünmektedir.

Çevremizdeki anlamlı fiziksel olaylar, onları anlamlandıran ve yorumlayan zihinsel süreçlerle doğrudan ilişkilidir; “Hareketleri işarete, sesleri sözcüğe, nesnelere âlete dönüştüren, bunları anlamsız bir şeyken anlamlı bir şeye dönüştüren insanî bilinç durumudur” (İbiş, 2023: 719). Bu insanî bilinç durumu, çevremizdeki dünyayı yalnızca algılamakla kalmayıp, aynı zamanda onu anlamlandırmamızı ve etkileşim içinde olmamızı sağlar. Anlamlandırma süreci ise bağlam anlayışımız ile doğrudan ilintilidir. Dolayısıyla anlam, “pek çok kavramsal boyutun bilinmesiyle ortaya çıkan bir zihin durumudur” (İbiş, 2023: 716).

İnsanlar, bir ifadenin anlamını çevresel bağlam ve deneyimleri ışığında değerlendirir. Ayrıca anlam sosyal ve kültürel bir bağlamı da içerir. Bir nesnenin, olayın veya ifadenin anlamı, onu yorumlayan bireyin kültürel ve sosyal arka planına bağlı olarak şekillenir. Bu bağlamı anlamadan, bir şeyin tam anlamıyla kavranması mümkün değildir. Örneğin, Japon kültürünü hiç tanımayan bir kişi, yemek için bulunduğu bir yerde o kültüre özgü olan yemek çubukları ile ilk karşılaşmasında o çubukların anlamını kavrayamayacaktır. Dolayısıyla, “anlam hiçbir zaman bir varlıkta, o varlığı gösteren süreçte, o varlığı görmemizi sağlayan zihinsel süreçte sabitlenmiş durumda değildir. Anlam bir sözcüğün sosyal ve kültürel dinamiklerin ortaya çıktığı farklı bağlamlar içindeki yeri olarak düşünülmektedir” (Zambak, 2019: 343).

“Bir dil tasavvur etmek, bir yaşam biçimi tasavvur etmek demektir” der (Wittgenstein, 2005: 29). Dil, bir kültür içinde gelişir ve bu bağlam içinde bir anlam kazanır. Wittgenstein’in, “filozoflar bir sözcüğü kullandıklarında—“bilgi,” “varlık,” “nesne,” “ben,” “tümce,” “ad”—ve bir şeyin özünü kavramaya çalıştıklarında kendimize şunu sormalıyız: Bu sözcüğün, anayurdu olan dilde gerçekten böyle kullanıldığı olur mu hiç?” (Wittgenstein, 2005: 68) şeklinde yönelttiği soru, toplum içinde uzlaşım yoluyla belirlenen kuralların dilin kullanımını belirlediğinin ifadesidir. Anlama, cümlenin içinde bulunduğu dil oyunu içinde mümkündür. “Dil-oyunu” deyimini burada, dilin konuşulmasının bir etkinliğin ya da bir yaşam biçiminin bir parçası olduğunu belirgin kılmalıdır” (Wittgenstein, 2005: 32).

Wittgenstein’in dil oyunları teorisi, dilin toplumsal ve kültürel bağlam içindeki işlevine dair önemli bir içgörü sunar. Dil, yalnızca sembolik bir iletişim aracı değil, aynı zamanda sosyal etkileşimlerin ve kültürel normların yapı taşlarını oluşturur. Her dil oyunu, belirli bir yaşam biçimi ve dünya görüşü etrafında şekillenir. Bu bağlamda, dilin anlamı ve kullanımı, yalnızca kelimelerin nesnel anlamlarıyla sınırlı değildir; aynı zamanda dilin konuşulduğu toplumun değerleri, normları ve pratikleriyle de derin bir ilişkiye sahiptir. Dolayısıyla, bir dilin veya terimin anlamını kavrayabilmek için, o dilin konuşulduğu toplumsal bağlamı ve kültürünü anlamak gereklidir. Wittgenstein’in vurguladığı üzere, bir kelimenin anlamını çözümlmek için, o kelimenin kullanıldığı dil oyununu ve bu oyunun kültürel arka planını dikkate almak zorunludur. Bu yaklaşım, dilin sadece soyut bir bilgi değil, aynı zamanda pratik bir yaşam biçimi olduğunu gözler önüne serer.

Bu görüşler, bağlam anlayışı geliştiremeyecek olan yapay bir sistemin hiçbir zaman anlamayı tam olarak gerçekleştiremeyecek olması sonucunu da doğurur. Dolayısıyla dili anlama, özellikle çeviri çalışmalarında alanı oldukça zorlayan bir problem olarak karşımıza çıkar. Bir dilin başka bir dile çevirisini gerçekleştirirken mutlaka bir anlam kayması yaşanır. Zira aynı kelimenin, aynı dil içerisinde dahi bağlama dayalı farklılıkları mevcutken gerçekten mükemmel bir çeviri yapabilen bir program oluşturmak oldukça zor görünmektedir. Daha iyi anlaşılması açısından Williard V. O. Quine’in “çeviri belirsizliği” olarak adlandırılan, “orijinal dili yüzde

yüz doğrulukla yorumlamanın imkânsız” (Zambak, 2019: 355) olduğu tezine yönelik örnek senaryosunu aktaralım: “Ne zaman ortalıkta bir tavşan görünse, bu dilin kullanıcıları tavşanı işaret ederek tek-kelimelelik bir cümle olan “Gavagai” cümlesini sarf etmektedirler. Bu yerel dildeki “Gavagai” cümlesi ile Türkçe “Tavşan” cümlesinin aynı gözlemlenebilir koşullarda sarf edilmesinden hareket eden çevirmen, “gavagai” kelimesinin “tavşan” anlamına geldiğini (ya da Türkçeye bu şekilde tercüme edilebileceğini) makul bir biçimde düşünebilir” (Demircioğlu, 2015: 21). Burada dikkat edilmesi gereken nokta, bu kavramın yalnızca tek bir şeye karşılık gelemeyecek kadar çok fazla anlama karşılık gelebilmesi ihtimali ve bu nedenle çevirinin mümkün olmaması durumudur. Zira bu ifade, bir canlıya, bir hayvana, tavşanın bir parçasına, tavşanın bir hareketine, et yemeğine, ava gitme teklifine, tavşanın sevimliliğine vb. pek çok anlama gönderme yapıyor olabilir.

Bu akıl yürütmede ulaşılan sonuç açıktır; "gavagai" kelimesi belirsizdir, çünkü birçok davranışsal duruma uyum sağladığından, farklı anlamlar ve ifadelerle çevrilebilir. Dolayısıyla, dilin kullanım kuralları geliştiği kültür içinde anlam kazanır. Deneyimlemediğimiz bir kültür içinde dilin anlamını belirlemek oldukça zordur. Aynı cümleler kullanılış yerine ya da tonlamalara göre farklı anlamlara da gelebilir. Bu yüzden Wittgenstein’in ifade ettiği gibi “anlam kullanımdır” (Zambak, 2019: 343).

İnsanlar için bile oldukça büyük bir zorluk oluşturan dil çevirisinin mekanik bir sistemde mükemmel doğrulukta gerçekleştirilebilmesi elbette mümkün değildir. Yapay zekâ sistemleri, kelimeler arasındaki ince anlam farklılıklarını ve bağlamı tam olarak kavrayabilme yeteneğinden yoksundur. Çeviri sürecinde insan çevirmenler, metnin anlamını ve bağlamını yansıtmak için çeşitli yorumlar yapabilirken, yapay zekâ sistemleri ilintili bağlamlarla gerçek bir ilişki kuramamaları nedeniyle insan çevirmenlerin yorumlama gücünden doğal olarak mahrumdurlar. Öte yandan, son yıllarda yapay zekâ temelli çeviri programlarının çok büyük başarılar elde ettikleri ve bağlam inceliklerinin kritik bir öneme sahip olmadığı metin çevirilerinde, neredeyse insan çevirmenler düzeyinde çeviri yapabildikleri de bir gerçektir. Ancak bu çeviri programlarının hiçbir zaman bir insan seviyesinde öz farkındalığa sahip olamayacağı ve bir insanın içinde bulunduğu bağlamlarla kurduğu organik ilişkileri hiçbir zaman gerçekleştiremeyeceği kabul edildiğinde; bir dili anlama, yorumlama ve başka bir

dilden çeviri yapma becerisinin yapay zekâ sistemlerinde hiçbir zaman bir insaninkine yaklaşamayacağı iddia edilebilir.

Güçlü yapay zekâ hedefinden hala epey uzaktayız ve bu hedefe ulaşmak için aşılması gereken engeller, yalnızca teknik zorluklar değil, aynı zamanda ilkesel sınırlamaları da içeriyor görünmektedir. Zira bu engeller, tümüyle insan zihnine ve bilincine özgü olan ve doğasını henüz tam olarak kavrayamadığımız yönleri bünyesinde barındırmaktadır. Bu bağlamda, yapay zekânın insan benzeri bilinçli bir zihinselliğe ulaşabilmesi için sadece teknolojik ilerlemeler değil, aynı zamanda zihinsel süreçlerimizin derinlemesine anlaşılması da kritik bir öneme sahip görünmektedir.

5. SONUÇ: İLKESEL SINIRLILIKLAR BİZE İNSAN BİLİNCİ HAKKINDA TAM OLARAK NE SÖYLER?

Bu çalışmada, Dreyfus'un güçlü yapay zekâya yönelik eleştirileri bağlamında insanın bilinçli zihinselliğinin yapay bir sistemde modellenmesinin önündeki engelleri ortaya koyma amacı güttük. Bu hedef doğrultusunda, önce yapay zekânın kısa bir tarihçesini Giriş bölümünde ana hatlarıyla ele aldık. Ardından Bölüm 1'de Dreyfus'un yaşamı, çalışmaları ve yapay zekâya genel yaklaşımını değerlendirdik. Bölüm 2'de Dreyfus'un güçlü yapay zekâya getirdiği eleştirilerinin geçerliliğini merkeze alan bir sınıflandırma yaparak söz konusu eleştirileri sistematik bir incelemeye tabi tuttuk. Bölüm 3'te ise, Bölüm 2'de yaptığımız sınıflandırma içinden yalnızca insan bilinçliliğine özgü yönleri içeren üçüncü gurubu öne çıkararak bu kapsamdaki eleştirileri derinlemesine analiz etme teşebbüsünde bulunduk. İncelememiz sonucunda, insanın bilinçli zihinselliğini oluşturan fenomenal yönlerin aşılabilir ilkesel sınırlılıklar oluşturduğu ve bu nedenle güçlü yapay zekânın insan benzeri bir bilinç geliştirmesinin mümkün olmadığı sonucuna ulaştık. Kanaatimizce bizi bu sonuca götüren muhakemenin bir miktar daha açılması yerinde olacaktır.

Yapay zekânın, insanın bilinç düzeyine ulaşabileceği ve bu nedenle insan gibi düşünme kapasitesine sahip olabileceği görüşü, zihinsel süreçlerin biyolojik bir yapıya bağlı olmayan; hesaplanabilir, sembolik ve algoritmik süreçlerden oluştuğu savını destekleyen işlevselcilik ve hesaplamacı zihin teorilerine dayanmaktadır. Söz konusu Güçlü (genel) yapay zekâ görüşü, zihinsel durumların fiziksel yapılardan bağımsız olarak işlevsel rollerine göre tanımlanabileceği ve bu nedenle yeterince gelişmiş bir yapay sistemin de insan zihnindeki bilişsel süreçleri taklit ederek, öğrenme, mantık yürütme ve problem çözme gibi yeteneklerde insan benzeri düşünme kapasitesine ulaşabileceği gerekçesiyle savunulmaktadır. Dolayısıyla güçlü yapay zekâ görüşüne göre, makineler veya dijital programlar tümüyle insanlara özgü olduğu düşünülen anlama ve bilinçli düşünme becerilerine de pekâlâ sahip olabilirler.

Ancak, insan bilinçliliğini yalnızca bilişsel süreçlere indirgeyerek açıklayan bu görüşler, zihinsel yetilerin karmaşıklığını ve çok boyutlu doğasını göz ardı etmektedir. Zira insan bilinçliliği, yalnızca bilişsel süreçlerle değil, aynı zamanda derin duygusal

deneyimlerle, sosyal etkileşimlerle ve bedenle olan ilişkiyle şekillenir. Bu bağlamda, insan bilinçliliğinin salt hesaplanabilir süreçlerle açıklanamayacağını ortaya koyan Hubert Dreyfus, insan gibi düşünebilen yapay zekâ fikrine karşı oldukça etkili eleştiriler geliştirmiştir.

Hubert Dreyfus'a göre, insan düşüncesinin özünde, yalnızca sembolik ve algoritmik işlemler değil, bedensel varoluş, bağlamsal anlayış ve sezgisel kavrayış gibi yapay sistemlerin erişemeyeceği derin yapısal unsurlar yatmaktadır. Özellikle Dreyfus'un fenomenolojiye dayanan eleştirileri, insan bilincinin tamamen bağlama özgü, bedensel olarak özümsemiş bir deneyim ve dünyayla ilişkili bir kavrayış içerdiğini savunur. Ona göre, bedensel varoluşa sahip olan insan gömülü, özümsemiş ve bedenlenmiş bir varlıktır. İnsan, bir bedene sahip olarak dünyaya gelir ve zekâsı bu bedensel varoluşuyla birlikte şekillenir. Dolayısıyla insanın bedensel varoluşa sahip bir varlık olarak, sağduyu ve sezgisel anlayışa sahip olması, belirsizlik karşısında neyin alakalı olduğunu bilinçsiz bir şekilde belirleyerek eylemde bulunabilmesi, arka plan bilgisine sahip olması, deneyiminin niteliklerinin yalnızca kendine özgü olması ve kendinin farkındalığına sahip olması, bağlam anlayışı sayesinde dünya ile etkileşimde bulunabilmesi ve bu sayede bir anlama kabiliyeti geliştirebilmesi şeklinde ifade edebileceğimiz yetileri, bilgisayarimsal-hesaplamacı (computational) süreçlere açık olmayan fenomenal yönüyle ilgilidir. Bu durum, yapay zekânın insanın bilinçli zihniyle eşdeğer bir düşünme kapasitesine ulaşmasının önünde önemli ilkesel sınırlılıklar bulunduğunu gösterir.

İlkesel sınırlılıklar, yapay zekânın insan bilincine erişememesinin temel nedenlerini anlamamıza yardımcı olurken, aynı zamanda insan bilincinin kendine özgü yapısını ve derinliğini de ortaya koyar. Hubert Dreyfus'un eleştirilerinden hareketle, insan bilincinin yalnızca hesaplanabilir bir süreç olmadığını ve bu nedenle yapay sistemlerle simüle edilemeyeceği sonucunu çıkarabiliriz. Dreyfus'un vurguladığı bedenlenmiş varoluş, sezgisel anlama ve bağlamsal farkındalık gibi unsurlar, insan bilincinin dünyaya gömülmüş, doğrudan deneyimle iç içe bir varoluş biçimi olduğunu gösterir.

Bu ilkesel sınırlılıklar, insan bilincinin özellikle bedenlenmiş bir deneyimle şekillendiğini ve anlamın soyut hesaplama süreçlerinden değil, bir bedenle dünya içinde yer almanın getirdiği farkındalıktan ve bağlam ilişkilerinden doğduğunu gösterir. Yapay zekâ sistemleri, ne kadar geliştirilmiş olursa olsunlar; bu dünyayla özdeşleşmiş, bedensel farkındalığa dayalı insan bilincini tam olarak taklit etmekte bile başarısızlığa uğramakta, sezgisel ve bağlamsal anlamayı deneyimlemeden anlamlandırma girişiminde bulunduğu için hedeflerine ulaşamamaktadır.

İnsan bilincinin derinliği, sezgi ve içgüdüye dayalı bilinç gibi unsurların varlığıyla daha da belirgin hale gelir. Bu unsurlar, sadece hesaplamalarla yakalanamaz; çünkü insan, dünyayla etkileşim kurarken semboller veya algoritmalar aracılığıyla değil, anlamaya yönelik bütüncül bir deneyimle hareket eder. Bu açıdan ilkesel sınırlılıklar, yapay zekânın insan bilincine yaklaşmasını engelleyen faktörler olarak karşımıza çıkmakla kalmaz, aynı zamanda insan bilincinin ne kadar derin, karmaşık ve öznel olduğunu da bize açıkça gösterir.

Sonuç olarak, Hubert Dreyfus'un yapay zekâyâ yönelik eleştirilerini merkeze alarak, yapay zekânın ilkece insan bilinçliliğine eşdeğer olamayacağını iddia etmekteyiz. Zira Dreyfus'un dikkat çektiği, insan bilinçliliğinin özünü oluşturan ve yalnızca insana özgü olan fenomenal yönler, insan zekâsının yapay bir sistemde modellenmesine imkân vermeyen ilkesel sınırlılıklar olarak karşımıza çıkar. Bu ilkesel sınırlılıklar ilkece aşılamayacağından yapay bir sistemin gerçekten insan gibi anlayarak düşünebilmesi, tüm yönleriyle insan bilinçliliğine eşdeğer olması mümkün görünmemektedir.

Güçlü (veya genel) yapay zekânın hedefi, insan bilinçliliğine eşdeğer, hatta onu aşan bir bilinci modelleyebilme olduğu sürece gerçek bir başarı yakalaması olası görünmemektedir. Ayrıca teknolojik gelişmeler ışığında, bugün halihazırda etkileyici yapay sistemler geliştirilmiş olması, bu hedefin yakın olduğu sonucunu da doğurmamaktadır. Elbette, bundan 10 yıl öncesinde hayal etmesi bile zor olan pek çok teknolojik gelişmenin bugün yaşandığı görmezden gelinebilecek bir durum değildir; ancak yapay bir sisteme gerçekten bilinçli bir şekilde düşünebiliyor diyebilmek için gereken koşulların henüz sağlanmadığı da ortadadır ve gelecekte sağlanabilmesi de bu

çalışmada ele aldığımız ilkesel sınırlılıklar göz önüne alındığında kanaatimizce mümkün görünmemektedir. İnsan bilincinin nasıl ortaya çıktığı ve gerçekten tam olarak neyi kapsadığı hala bilinmezlik içindeyken yapay bir bilinç tasarımı zaten belli sınırlar dahilinde ortaya konabilir. Öyle görünüyor ki, her yönüyle insan bilinçliliğine eşdeğer bir yapay zekâ hedefindeki çalışmalar hayal edildiği şekilde tam başarıyı yakalayamayacaktır.



KAYNAKÇA

- Alıcı, T. (2013). *Gerçek Bir Yanılsama: Bilinç*. İstanbul: Metis Yayıncılık.
- Andler, D. (2000). "Context and background. Dreyfus and cognitive science." Mark Wrathall and Jeff Malpas (Der.), *Heidegger, Coping and Cognitive Science*, içinde, (s. 137-159). Cambridge: MIT Press.
- Arıcı, M. (2014). "Materyalizm, Fenomenal Özne ve Ontolojik Statüsü." *Beytulhikme An International Journal of Philosophy*, 4(2):1-13.
- Arıcı, M. (2015). "Zihinsel Alanda "Ben Hissi" Neye Karşılık Gelir?" *Dört Öge* (7): 65-76.
- Arıcı, M. (2018). "The Problem of Phenomenal Consciousness: A Descriptive and Categorical Analysis." *MetaZihin*, 1(1), 1-19.
- Arıcı, M. (2019). "Zihin Felsefesi: Ben Nerede, Zihin Bedensiz Var Olabilir mi?" Murat Arıcı (Der.), *Felsefeye Giriş: Temel Problemlere Sistematik Yaklaşım*, içinde, (s. 107-152). Ankara: Nobel Yayıncılık.
- Block, Ned (1995). "On a Confusion about a Function of Consciousness." *Behavioral And Brain Sciences*,(18): 227-287.
- Cevizci, A. (2009). *Felsefe tarihi* (5). Say Yayınları. (s. 1123-1140).
- Crevier, D. (1993). *AI The Tumultuous History Of Search For Artificial Intelligence*. USA:BasicBooks.
- Demircioğlu, E. (2015). "Quine ve Semantik Belirsizlik" *Felsefe Tartışmaları*, (51):18-28.
- Dreyfus, H. L. (1965). *Alchemy and Artificial Intelligence*. RAND Paper P-3244. Santa Monica, CA: RAND Corporation.

- Dreyfus, H. L. (1967). "Why Computers Must Have Bodies in Order to Be Intelligent." *The Review of Metaphysics*, 21(1):13-32. Alındığı URL: <https://www.jstor.org/stable/20124494>
- Dreyfus, H. L. (1972). *What Computers Can't Do: A Critique of Artificial Reason*, New York: Harper&Row Pub.
- Dreyfus, H. L. (1974). "Artificial intelligence" *Annals of the American Academy of Political and Social Science, The Information Revolution*, (412): 21–33.
- Dreyfus, H. L., & Dreyfus, S. E. (1986). *Mind Over Machine*, New York: The Free Press.
- Dreyfus, H.L. (1993). "Heidegger's Critique of the Husserl/Searle Account of Intentionality". *Social research*, 60(1): 17-38. Alındığı URL: <https://www.jstor.org/stable/40970727>
- Dreyfus, H. L. (2005). "Overcoming the Myth of the Mental: How Philosophers Can Profit from the Phenomenology of Everyday Expertise," *Proceedings and Addresses of the American Philosophical Association*, 79(2):47-65.
- Dreyfus, H. L., & Dreyfus, S. E. (2014). "Making a Mind versus Modeling the Brain: Artificial Intelligence Back at a Branchpoint." Mark A. Wrathall (Der.), *Skillful Coping: Essays on the Phenomenology of Everyday Perception and Action*, içinde, (s. 205-230). Oxford University Press.
- Dreyfus, H. L. (2014). "Why Heideggerian AI Failed and How Fixing it Would Require Making it More Heideggerian Degenerating". Mark A. Wrathall (Der.), *Skillful Coping: Essays on the Phenomenology of Everyday Perception and Action*. içinde, (s. 249-273). Oxford University Press.
- Eyim, A. (2019). "Alan Turing'in Yapay Zekâ Anlayışına Bir İtiraz Olarak Gödel'in Eksiklik Teoremi." Z. Doğan (Der.), 6. Uluslararası Bilimsel Araştırmalar Kongresi, içinde, (s.217-227). Asos Yayınları.

- Fearn, N. (2005). *Philosophy: The Latest Answers To The Oldest Questions*. London: Atlantic Books.
- Gallagher, S. & Zahavi, D. (2023) “Phenomenological Approaches to Self-Consciousness”, Edward N. Zalta & Uri Nodelman (Der.), *The Stanford Encyclopedia of Philosophy*, (Winter 2023 Edition) içinde. Alındığı URL = <https://plato.stanford.edu/entries/self-consciousness-phenomenological/>
- Gehrman, K., & Schwenkler, J. (2020). “Hubert Dreyfus on Practical and Embodied Intelligence.” *The Routledge Handbook of Philosophy of Skill And Expertise*, içinde, (s. 123-132). Londra: Routledge.
- Gözkân, B. (2010) *Sunuş*. Gödel Kanıtlanması, Boğaziçi Üniversitesi Yayınevi, x-xvii.
<http://jmc.stanford.edu/articles/dartmouth/dartmouth.pdf>
- İbiş, F. (2023). “Anlamın Epistemolojik Ön Koşulu Olarak Bağlam.” *Afyon Kocatepe Üniversitesi Sosyal Bilimler Dergisi*, 25(2):714-722.
- James, J. (1983). *The Principles of Psychology*. Cambridge: Harvard University Press.
- Kaku, M. (2015). *Zihnin Geleceği. Bilimin Zihni Anlamaya ve Geliştirmeye Yönelik Arayışları*, Çev. Emre Kumral. Ankara: ODTÜ Yayıncılık.
- Kaku, M. (2016). *Geleceğin Fiziği: 2100 Yılına Kadar Bilim İnsanlığın Kaderini ve Günlük Yaşamımızı Nasıl Şekillendirecek*, Çev. Yasemin S. Oymak & Hüseyin Oymak. Ankara: ODTÜ Yayıncılık.
- Kaplan, J. (2016). *Artificial Intelligence What Everyone Needs To Know*. Oxford University Press.
- Kenaw, S. (2008). “Hubert L. Dreyfus”s Critique of Classical AI and Its Rationalist Assumptions”. *Minds and Machines*, 18(2):227-238.
- Kutlusoy, Z. (2006). “Dreyfus, Hubert L.” Ahmet Cevizci (Der.), *Felsefe Ansiklopedisi*, (Cilt 4) içinde (s. 729-732). Ankara: Babil Yayınları.

- Learning, J. M. (2021). *Düşünen Makineler*. Çev. Samet Öksüz, Ankara: Say Yayınları.
- Mangan, B. (2007). “Cognition, Fringe Consciousness, and the Legacy of William James”. Max Velmans & Susan Schneider (Der.), *The Blackwell Companion to Consciousness*, içinde, (s. 673-685). USA: Blackwell.
- Manzotti, R. & Parks, T. (2018). *Zihnin Ucu Bucağı: Bilinç ve Dünya Bir midir?*, Çev. Özge Duygu Gürkan. İstanbul: Metis Yayınları.
- Nagel, T. (2008). “Yarasa Olmak Nasıl Bir Şeydir?”. Hofstadter, D. R. & Dennett, D. (Der.), *Aklın G'özü*. Füsün Doruker (Çev.), içinde, (s. 377-389). İstanbul: Boğaziçi Üniversitesi Yayınevi.
- Nagel, E. & Newman, J. R. (2010). *Gödel Kanıtılaması*. Çev. Bülent Gözkân. Boğaziçi Üniversitesi Yayınevi, İstanbul.
- Newell, A., & Simon, H. A. (1963). GPS, A Program That Simulates Human Thought. Edward A. Feigbaum & Julian Feldman (Der.). *Computers and Thought*, McGraw-Hill, 279-293.
- Nida-Rümelin, M. (2017). “Self-awareness.” *Rewiew of Philosophy and Pscyhology*. (8):55–82.
- Nillson, Nils J. (2018). *Yapay Zekâ Geçmişi ve Geleceği*. Çev. Mehmet Doğan, Boğaziçi Üniversitesi Yayınları, İstanbul.
- Penrose, R. (2000). *Bilgisayar ve Zekâ Kralın Yeni Usu I*, Çev. Tekin Dereli, Ankara: Tübitak Yay.
- Rousse, B. S. (2019). “Self-awareness and self-understanding”. *Eur J Philos*. 2019(27): 162–186. <https://doi.org/10.1111/ejop.12377>
- Searle, J. R. (1980). “Minds, Brains and Programs.” *Behavioral and Brain Sciences* 3(3): 417-457.
- Searle, J. R. (1996). *Akıllar Beyinler ve Bilim*. Çev. Kemal Bek, İstanbul: Say Yayınları.

- Smith, J. (2024). "Self-Consciousness," Edward N. Zalta & Uri Nodelman (Der.), *The Stanford Encyclopedia of Philosophy*, (Summer 2024 Edition) içinde. URL = <https://plato.stanford.edu/archives/sum2024/entries/self-consciousness>.
- Susser, D. (2013). "Artificial İntelligence and the Body: Dreyfus, Bickhard, and the Future of AI." Vincent C. Müller (Der.), *Philosophy and theory of artificial intelligence*. içinde, (s. 277-287). Springer Berlin Heidelberg.
- Tsai, C. (2022). "Beyond İntuitive Know-how." *Phenom Cogn Sci*. <https://doi.org/10.1007/s11097-022-09851-5>
- Turing, A. M. (1950). "Computing Machinery and Intelligence." *Mind* (49): 433-460.
- Yeşilkaya, N. (2022). "Felsefi Bir Sorun Olarak Yapay Zekâ [Artificial Intelligence As a Philosophical Problem]." *Bozok Üniversitesi İlahiyat Fakültesi Dergisi*, 22: 97-126.
- Zambak, A. F. (2019). "Dil Felsefesi: Anlam Nedir, Zihin ve Dünya ile İlişkisi Nasıldır?" Murat Arıcı (Der.), *Felsefeye Giriş: Temel Problemlere Sistematik Yaklaşım*, içinde, (s. 331-359). Ankara: Nobel Yayıncılık,
- Zeman, A. (2006). *Bilinç, Kullanım Kılavuzu*. İstanbul: Metis Yayınları.
- Wittgenstein, L. (2005). *Felsefi Soruşturmalar*. Çev. Haluk Barışcan. İstanbul: Metis Yayıncılık.
- Wooldridge, M. (2022). *Bilinçli Makinelere Giden Yol*. Çev. Özge Çelik. İstanbul: Metis Yayınları.
- Wrathall, M. (2014). "Introduction: Hubert Dreyful and the Phenomenology of Human Intelligence", M. Wrathall (Der.), *Skillful Coping: Essays on the Phenomenology of Everyday Perception and Action*, içinde, (s. 1-22). Oxford: Oxford University Press.

Wrathall, M. (2017). "Introduction Background Practices and Understandings of Being", Mark Wrathall (Der.), *Background Practices: Essays on the Understanding of Being*, içinde, (s. 1-15). Oxford: Oxford University Press.

