



OECD

Yapay Zekâ

Yetkinlik Göstergelerinin Tanıtımı

Yapay Zekâ Çeviri Serisi-1



T.C. MİLLÎ EĞİTİM BAKANLIĞI
YENİLİK VE EĞİTİM TEKNOLOJİLERİ
GENEL MÜDÜRLÜĞÜ

2026 | ANKARA



T.C. MİLLÎ EĞİTİM BAKANLIĞI
YENİLİK VE EĞİTİM TEKNOLOJİLERİ
GENEL MÜDÜRLÜĞÜ

OECD

Yapay Zekâ

Yetkinlik Göstergelerinin Tanıtımı

Yapay Zekâ Çeviri Serisi-1

2026 | ANKARA



KÜNYE



**T.C. MİLLÎ EĞİTİM
BAKANLIĞI**

**YENİLİK ve EĞİTİM TEKNOLOJİLERİ
GENEL MÜDÜRLÜĞÜ**

Nisan 2026

Kitap Adı

Bu çalışma, OECD tarafından yayımlanan Introducing the OECD AI Capability Indicators başlıklı eserin Türkçe çevirisidir.

Bu çeviri, T.C. Millî Eğitim Bakanlığı Yenilik ve Eğitim Teknolojileri Genel Müdürlüğü [YEĞİTEK] tarafından hazırlanmıştır.

Yayın Koordinatörü

Mustafa CANLI
YEĞİTEK Genel Müdürü

Editör

Cennet ERKANAT
Yapay Zekâ ve Büyük Veri Uygulamaları Daire Başkanı

Çeviri

Dr. Hüseyin ULUKUZ – Yapay Zekâ Politikaları Koordinatörü / Millî Eğitim Uzmanı
Dr. Çiğdem MERAL – Millî Eğitim Uzmanı

Son Okuma

Banu ÖZDEMİR – Başöğretmen

Grafik & Tasarım

Hüseyin DALKILIÇ – Grafiker

Adres

Emniyet Mahallesi, Milas Sokak, No:8
06560 Yenimahalle / ANKARA

Özgün Eser

OECD [2025], Introducing the OECD AI Capability Indicators, OECD Publishing, Paris.
<https://doi.org/10.1787/be745f04-en>



Çeviri Notu

Bu eserin çevirisinde kaynak metnin anlamı ve yazar[lar]ın niyetinin korunmasına özen gösterilmiş; aynı zamanda hedef dilde doğal ve akıcı bir anlatım sağlamak amacıyla Anlamsal Çeviri [Semantic Translation] yaklaşımı benimsenmiştir.

Özgün metin ile bu çeviri arasında herhangi bir farklılık olması durumunda yalnızca özgün metin geçerli kabul edilir.



EDİTORYAL SUNUŞ

İçinde bulunduğumuz teknoloji çağı, bugün tüm insanlığı kişisel yetkinliklerin ve yeteneklerin her zamankinden daha önemli olduğu yepyeni bir döneme getirmiştir. Bu yeni dönemde yalnızca yapay zekâ araçları ile iş birliği yapabilen bireyler değil; aynı zamanda sosyal ilişki de kurabilen, empati yeteneği yüksek ve zanaat sahibi bireyler bu yarışta öne çıkmaktadır. Bilinmelidir ki inovasyon bir noktada duracaktır, fakat “insan faktörü” hiçbir zaman bitmeyecek, sonunda yine beşerî beceriler önem kazanacaktır. *Türkiye Yüzyılı Maarif Modeli (TYMM)*, insan merkezli yaklaşımı ve beceri temelli yapısıyla bu dönüşümü açıkça ortaya koymaktadır.

Dünya Ekonomik Forumu’nun İşlerin Geleceği 2023 Raporu’nda, mevcut becerilerin yaklaşık %39’unun 2030 yılına kadar değişime uğrayacağı veya geçerliliğini yitirebileceği belirtilmektedir. Bu durum, insan becerileri ile yapay zekâ yetkinliklerinin birlikte ele alınmasını gerekli kılmaktadır. OECD tarafından geliştirilen bu rapor ise yapay zekâ sistemlerinin insan yetenekleriyle ilişkisini ele alan önemli bir referans sunmaktadır. Bu nedenle insan yetenekleri ile yapay zekâ yetkinliklerinin karşılaştırıldığı bu raporu önemli buluyoruz. Söz konusu çalışma, Türkiye’de eğitimde yapay zekâ kullanımının güçlendirilmesine katkı sağlamakta ve “*Eğitimde Yapay Zekâ Politika Belgesi ve Eylem Planı*” [2025–2029] ile uyumlu bir çerçeve sunmaktadır.

Millî Eğitim Bakanlığı olarak temel hedefimiz, öğrencilerimizi yapay zekâ çağının gerektirdiği bilgi, beceri ve yetkinliklerle güçlendirmek, öğrencilerimizin yapay zekâyı kullanan ve aynı zamanda anlayan, sorgulayan ve üreten bireyler hâline gelmelerini sağlamaktır. Bu doğrultuda bir yandan uluslararası gelişmeleri ve politika çerçevelerini takip ederken, bir yandan da eğitim sistemimizi değer temelli vizyonumuz doğrultusunda geliştirmeyi öncelikli bir sorumluluk olarak görmekteyiz.

OECD tarafından hazırlanan bu referans çalışmanın Türkçeye kazandırılmasının, eğitim politikalarımızın daha bilinçli ve stratejik bir zeminde şekillenmesine katkı sağlayacağına inanıyor, eğitim camiasına ve tüm paydaşlara faydalı olmasını temenni ediyorum.

Mustafa CANLI

Yenilik ve Eğitim Teknolojileri
Genel Müdürü

OECD Yapay Zekâ Yetkinlik Göstergelerinin Tanıtımı



Bu çalışma, OECD Genel Sekreterliğinin sorumluluğu altında yayımlanmıştır. Burada ifade edilen görüşler ve ileri sürülen argümanlar, OECD Üye Ülkelerinin resmî görüşlerini zorunlu olarak yansıtmaz.

Bu belge ile burada yer alan veri ve haritalar, herhangi bir ülkenin statüsü veya egemenliği, uluslararası sınırların belirlenmesi ya da herhangi bir ülke, şehir veya bölgenin adı hakkında ön yargı teşkil etmez.

İsrail'e ilişkin istatistiksel veriler, ilgili İsrail makamlarının sorumluluğu altında sağlanmıştır. OECD tarafından bu verilerin kullanımı, uluslararası hukuk çerçevesinde Golan Tepeleri, Doğu Kudüs ve Batı Şeria'daki İsrail yerleşimlerinin statüsü hakkında herhangi bir ön yargı teşkil etmez.

Türkiye Cumhuriyeti'nin Notu

Bu belgede "Kıbrıs" a yapılan atıflar, Ada'nın güney kesimine ilişkindir. Ada üzerinde hem Türk hem de Rum toplumu temsil eden tek bir otorite bulunmamaktadır. Türkiye, Kuzey Kıbrıs Türk Cumhuriyeti'ni (KKTC) tanımaktadır. Birleşmiş Milletler çerçevesinde kalıcı ve adil bir çözüm bulunana kadar Türkiye, "Kıbrıs meselesi" konusundaki tutumunu muhafaza edecektir.

OECD Avrupa Birliği Üyesi Devletlerinin ve Avrupa Birliği'nin Notu

Türkiye hariç olmak üzere Birleşmiş Milletler üyelerinin tamamı Kıbrıs Cumhuriyeti'ni tanımaktadır. Bu belgede yer alan bilgiler, Kıbrıs Cumhuriyeti fiilî kontrolü altındaki bölgeye ilişkindir.

Bu yayına şu şekilde atıfta bulunulmalıdır:

OECD [2025], Introducing the OECD AI Capability Indicators, OECD Publishing, Paris.

<https://doi.org/10.1787/be745f04-en>



Türkçe çeviri:

OECD [2025], OECD Yapay Zekâ Yetkinlik Göstergelerinin Tanıtımı (Türkçe çeviri), OECD Publishing, Paris.

ISBN 978-92-64-53190-1 [print]

ISBN 978-92-64-89309-2 [PDF]

ISBN 978-92-64-83602-0 [HTML]

© OECD 2025



Attribution 4.0 International (CC BY 4.0)

Bu çalışma, Creative Commons Attribution 4.0 International lisansı kapsamında kullanıma sunulmuştur. Bu çalışmayı kullanarak, söz konusu lisansın şartlarına tabi olmayı kabul etmiş olursunuz (<https://creativecommons.org/licenses/by/4.0/>).

Atıf – Çalışmaya atıfta bulunmanız gerekir.

Çeviriler – Özgün çalışmaya atıfta bulunmanız, özgün metinde yapılan değişiklikleri belirtmeniz ve aşağıdaki metni eklemeniz gerekir: Özgün eser ile çeviri arasında herhangi bir farklılık olması durumunda, yalnızca özgün eserin metni geçerli kabul edilir.

Uyarlamalar – Özgün çalışmaya atıfta bulunmanız ve aşağıdaki metni eklemeniz gerekir: Bu çalışma, OECD tarafından yayımlanan özgün bir eserin uyarlamasıdır. Bu uyarlamada ifade edilen görüşler ve kullanılan argümanlar, OECD'nin veya Üye ülkelerinin resmî görüşlerini temsil ettiği şeklinde yorumlanmamalıdır.

Üçüncü taraf materyaller – Lisans, çalışmada yer alan üçüncü taraf materyaller için geçerli değildir. Bu tür materyallerin kullanımı durumunda, ilgili üçüncü taraftan izin almak ve herhangi bir ihlal iddiasına ilişkin sorumluluğu üstlenmek sizin sorumluluğunuzdadır. OECD logosu, görsel kimliğini veya kapak görselini açık izin olmaksızın kullanamaz ve çalışmanın kullanımının OECD tarafından onaylandığını ima edemezsiniz. Bu lisans kapsamında ortaya çıkabilecek herhangi bir uyuşmazlık, Daimi Tahkim Mahkemesi (PCA) Tahkim Kuralları 2012 uyarınca tahkim yoluyla çözümlenecektir. Tahkim yeri Paris [Fransa] olacaktır. Hakem sayısı bir olacaktır.



ÖNSÖZ

Yapay zekâdaki gelişmeler ve bunların toplum üzerindeki etkileri hakkında hâlâ yeterince bilgi sahibi değiliz. Kamu politikaları, piyasaya sürülen her yeni YZ aracına sonradan uyum sağlamak amacıyla öğretim programlarını ve öğretim sistemlerini geriye dönük olarak uyarlamakla yetinmektedir. Kamu politikalarının YZ yetkinliklerinin nasıl gelişeceğini öngörmede daha aktif bir rol üstlenmesi gerekmektedir.

Bu rapor, tam olarak bu duruma yönelik bir metodolojiyi ortaya koymaktadır. Bu metodoloji, insan yeteneklerinin temel boyutlarını kapsayan bir gösterge seti sunmakta ve her bir göstergenin yapay zekânın insanla eşdeğer hâle gelme sürecini tanımlamaktadır. Bu göstergeler şunlardır: Dil; Sosyal Etkileşim; Problem Çözme; Yaratıcılık; Üstbilis ve Eleştirel Düşünme; Bilgi, Öğrenme ve Hafıza; Görme; Manipülasyon; ve Robotik Zekâ. Göstergeler beş düzeyli ölçekler hâlinde sunulmakta olup, YZ sistemleri için en zorlu yetkinlikler üst düzeylerde yer almaktadır. İnsan psikolojisine dayanan bu yaklaşım, YZ gelişimine ilişkin yapılandırılmış ve üst düzey bir bakış açısı sunmaktadır.

YZ yetkinliklerinin insan yetenekleri ile ilişkilendirilmesi, politika yapıcılara YZ'nin eğitimdeki potansiyel rolünü değerlendirme imkânı vermektedir. Örneğin, YZ öğretmenlik mesleği için kritik olan sosyal yetenekleri ne ölçüde taklit edebilir? Bu bağlamda YZ hangi alanlarda öğretmenlerin yerini alabilir ya da onları tamamlayabilir? Ayrıca, YZ yetkinliklerinin bir üst düzeye ulaşmasının olası sonuçları neler olacaktır?

Bu göstergeler, [eğitim] bakanlarının YZ'nin eğitimin geleceğine etkilerini – müfredat tasarımıyla pedagojik yaklaşımlara kadar – tartışmalarını mümkün kılacaktır. Bu durum, öğrencileri bizim geçmişimize göre değil, kendi geleceklerine hazırlayacak bir öğrenme ortamı oluşturmak amacıyla; eğitimde mekân, zaman, insan, teknoloji ve ilişkilerin nasıl yapılandırılması gerektiğini içermektedir.

Eğitimin ötesinde, bu göstergeler bakanların yapay zekânın istihdam, toplumsal katılım, serbest zaman etkinlikleri ve gündelik yaşam gibi alanlara yönelik etkilerini tartışmalarına imkân sağlayan bir çerçeveye sunmaktadır. Tüm bu alanlarda kamu politikalarının geçmişe değil, geleceğe odaklanması gerekmektedir.

Andreas SCHLEICHER
OECD Eğitim ve Beceriler
Direktörü



TEŞEKKÜR

Bu yayın, OECD'nin Yapay Zekâ ve Becerilerin Geleceği Projesi ekibi tarafından planlanmış ve geliştirilmiştir. Proje ekibinde Stuart Elliott [Proje Lideri], Abel Baret, Shivi Chandra, Marc Fuster-Rabella, Brianna Jesme, Margarita Kalamova, Aurelija Masiulyte, Sam Mitchell, Nóra Révai, Gianluca Risi, Mila Staneva ve Chi Sum Tse yer almaktadır.

Bu yayın, projeyi destekleyen saygın bilgisayar bilimciler ve psikologların paha biçilmez katkıları olmadan hazırlanamazdı.

Öncelikle, ilk OECD YZ Yetkinlik Göstergelerini geliştiren uzmanlara [alfabetik sırayla] şükranlarımızı sunmak isteriz: Ana Teresa Antunes, Joana Brito, Regina de Brito Duarte, Kexin-Jiang Chen, Anthony G. Cohn, Henrique Correia da Fonseca, Haohua Dong, Nicholas C. Georgiou, Miguel Faria, António M. Fernandes, Robert B. Fisher, Kenneth D. Forbus, Giorgio Franceschelli, Jonathan Francis, Arthur C. Graesser, Yvette Graham, José Hernández-Orallo, Cherie Ho, Ryota Kanai, Patrick Kyllonen, Christian Lebiere, Inês Lobo, Christopher Lochhead, Rebecca Martin, Elena R. Messina, Mirco Musolesi, Jean Oh, Ana Paiva, Rui Prada, Rebecca Ramnauth, Swen Ribeiro, Brian Scassellati ve Ana Vilaça Carrasco.

İkinci olarak, göstergelerin geliştirilmesine ilişkin değerlendirme veya danışmanlık sağlayan uzmanlara [alfabetik sırayla] teşekkür etmek isteriz: Phillip L. Ackerman, John Anderson, Guillaume Avrin, Eva L. Baker, Chandra Bhagavatula, Oliver Bown, Adrien Doerig, Angelo Cangelosi, Lucy Cheke, Emmanuel Chemla, Henrik I. Christensen, Michael Cohen, Ernest Davis, Charles Fadel, Dan Gutfreund, David Hogg, Jonathan Gratch, Sam Johnson, Zeid Kootbally, Noémie Le Donné, Joel Leibo, Martha Lewis, Matthias Michel, Harold F. O'Neil, Britta Rüschoff, Ute Schmid, Eric Schwitzgebel, Lorrie Shephard, Tadahiro Taniguchi, Marko Tešić, David Traum, Michael Witbrock ve Hiroshi Yamakawa.

Eğitim Araştırmaları ve Yenilik Merkezi [CERI] bünyesindeki meslektaşlarımıza teşekkür ederiz. CERI Başkanı Edmund Misson ve IMEP Bölümü Başkan Yardımcısı Stéphan Vincent-Lancrin, süreç boyunca gözetim, yönlendirme ve değerli tavsiyeler sağlamıştır.

Bu proje ve ekibi, eski Bölüm Başkanı Tia Loukkola'nın özverili desteği, düşünceli liderliği ve rehberliğinden büyük ölçüde faydalanmıştır. Zamansız vefatından önce vizyonu ve bağlılığı projenin gelişimi ve başarısı açısından belirleyici olmuştur. Kendisini derin bir minnettarlık ve hayranlıkla anıyoruz; katkıları bu çalışmanın kalıcı bir parçası olarak yaşamaya devam edecektir.

Eğitim ve Beceriler Direktörlüğü iletişim ekibi ile Halkla İlişkiler ve İletişim Direktörlüğü'ndeki meslektaşlarımız, yayının biçimlendirilmesine ve hazırlanmasına katkıda bulunmuştur.

Yayının bütüncül ve anlaşılır bir metin hâline gelmesini sağlamak üzere kapsamlı ve yapısal düzenlemeler yapan Mark Foss'a teşekkür ederiz.

Projenin geliştirilmesi sürecinde CERI Yönetim Kurulu'nun teşvik ve desteği için minnettarız.

Bu yayın, OECD'nin İş, Yenilik, Verimlilik ve Becerilerde Yapay Zekâ [AI-WIPS] programına katkıda bulunmaktadır. Bu program, politika yapıcılara YZ yetkinlikleri ve yayılışındaki hızlı gelişmeler ile bunların çalışma dünyasına etkileri hakkında güncel kalmalarını sağlayacak yeni kanıt ve analizler sunmaktadır. Program, çalışma dünyasında YZ'nin benimsenmesinin etkili, herkes için faydalı, insan merkezli ve toplum tarafından kabul gören bir şekilde gerçekleşmesini sağlamayı amaçlamaktadır. AI-WIPS, Almanya Federal Çalışma ve Sosyal İşler Bakanlığı [BMAS] tarafından desteklenmektedir ve Bakanlığın Politika Laboratuvarı Dijital, Çalışma ve Toplum bünyesindeki Alman YZ Gözlemevi'nin çalışmalarını tamamlayacaktır. Daha fazla bilgi için: <https://oecd.ai/en/work-innovation-productivity-skills> ve <https://denkfabrik-bmas.de/>



İÇİNDEKİLER

Editorial Sunuş	3
Önsöz	7
Teşekkür	9
Yönetici Özeti	13
Referanslar	14
1. Mevcut Yapay Zekâ Yetkinliklerine Genel Bakış	15
Mevcut Yapay Zekâ Yetkinliklerine İlişkin Karşılaştırmalı Tablo	16
Mevcut Derecelendirmelere İlişkin Yorumlar	18
Referanslar	19
2. Yapay Zekâ Yetkinliklerini Ölçmeye Yönelik Bir Çerçeve Oluşturulması	21
YZ Yetkinliklerine İlişkin Anlayışımızın Ötesinde İlerleme Kaydetmiştir	22
Metodoloji: Özgün ve Yenilikçi Bir Yaklaşım	23
Göstergelerin Oluşturulması ve Geliştirilmesi	25
Sınırlılıklar	26
Sonraki Adımlar	27
Yapay Zekâ Yetkinlik Göstergelerinin Rolü	28
Referanslar	30
3. OECD Yapay Zekâ Yetkinlik Göstergeleri	31
Dil Ölçeği	32
Sosyal Etkileşim Ölçeği	34
Problem Çözme Ölçeği	36
Yaratıcılık Ölçeği	38
Üstbiliş ve Eleştirel Düşünme Ölçeği	40
Bilgi, Öğrenme ve Hafıza Ölçeği	42
Görme Ölçeği	44
Manipülasyon Ölçeği	46
Robotik Zekâ Ölçeği	48
Referanslar	50
4. Yapay Zekâ Yetkinlik Göstergelerinin Politika Amaçlı Kullanım Alanları	51
Göstergelerin, İnsan Becerilerine Yönelik Mesleki İhtiyaçlarla Eşleştirilmesi	52
Eğitimde Dönüştürücü Değişim	55
Sonuç	57
Referanslar	59



İÇİNDEKİLER

ŞEKİLLER

Şekil 2.1: OECD Yapay Zekâ Yetkinlik Göstergelerinin Geliştirilme Süreci	24
Şekil 2.2: OECD Yapay Zekâ Yetkinlik Göstergeleri	24
Şekil 2.3: Yapay Zekânın Beş Düzeyine Genel Bakış	25
Şekil 4.1: Dil, Problem Çözme ve Sosyal Etkileşim Alanlarındaki Yapay Zekâ Yetkinlik Göstergelerinin Öğretmenlik Görevlerinin Gerektirdiği Becerilerle Eşleştirilmesi	53
Şekil 4.2: Eğitimde Artan Yapay Zekâ Yetkinliklerinin Etkilerini Analiz Etmeye Yönelik Bir Çerçeve	55

TABLolar

Tablo 1.1: Mevcut Yapay Zekâ Yetkinlik Düzeylerine Genel Bakış	17
Tablo 3.1: Yapay Zekâ Dil Ölçeği	33
Tablo 3.2: Yapay Zekâ Sosyal Etkileşim Ölçeği	35
Tablo 3.3: Yapay Zekâ Problem Çözme Ölçeği	37
Tablo 3.4: Yapay Zekâ Yaratıcılık Ölçeği	39
Tablo 3.5: Yapay Zekâ Üstbiliş ve Eleştirel Düşünme Ölçeği	41
Tablo 3.6: Yapay Zekâ Bilgi, Öğrenme ve Hafıza Ölçeği	43
Tablo 3.7: Yapay Zekâ Görme Ölçeği	45
Tablo 3.8: Yapay Zekâ Manipülasyon Ölçeği	47
Tablo 3.9: Yapay Zekâ Robotik Zekâ Ölçeği	49

ÖRNEKLER

Örnek 4.1: Problem Çözme Ölçeğinde Orta Düzeydeki Yapay Zekânın Salgınla Mücadeleyi Desteklediği Bir Senaryo	54
--	----



YÖNETİCİ ÖZETİ

Kasım 2022’de ChatGPT’nin kullanıma sunulmasıyla birlikte, yapay zekânın¹ [YZ] insan faaliyetleri üzerindeki olası etkileri geniş kitlelerin dikkatini çekmeye başlamıştır. Bununla birlikte, yapay zekâ hızla gelişmeye devam ederken, bu gelişmelerin doğurabileceği sonuçlara ilişkin toplumsal anlayış aynı hızla ilerlememektedir. Yapay zekânın insan faaliyetlerini nasıl dönüştürebileceğini anlamak için kapsamlı çalışmalara ihtiyaç duyulmaktadır.

Bu rapor, OECD’nin yeni YZ Yetkinlik Göstergelerini tanıtmaktadır. Göstergeler, politika yapımcılarının YZ yetkinliklerini kanıta dayalı bir çerçeve içinde anlamaları ve bunları insan becerileriyle karşılaştırmaları için geliştirilmiştir. Beş yıllık bir çalışmanın ürünü olan göstergeler, geniş bir YZ araştırmacısı, psikolog ve diğer uzman ağının katkılarına dayanmaktadır. Eşlik eden teknik raporun bölümleri [OECD, 2025[1]] 32 uzman tarafından yazılmış ve 25 uzman tarafından gözden geçirilmiştir.

Dokuz gösterge, yapay zekânın insan yetenekleriyle tam eşdeğerliğe doğru gelişimini tanımlayan şu insan becerilerini kapsamaktadır: Dil, Sosyal Etkileşim, Problem Çözme, Yaratıcılık, Üstbiliş ve Eleştirel Düşünme, Bilgi, Öğrenme ve Hafıza, Görme, Manipülasyon ve Robotik Zekâ. Göstergeler beş düzeyli ölçekler hâlinde sunulmaktadır. Ölçeğin üst basamaklarında YZ sistemleri için en zorlayıcı yetkinlikler yer almaktadır. Her düzey, YZ sistemlerinin doğru ve tutarlı biçimde yerine getirebildiği becerilere ilişkin kısa bir açıklama içermektedir. Mevcut YZ performansının her ölçekteki konumu, eldeki kanıtlarla ilişkilendirilmektedir.

Göstergeler burada beta sürüm olarak yayınlanmakta ve iki temel paydaş gruptan geri bildirim talep edilmektedir: YZ araştırmacıları ve politika yapımcılar. Araştırmacıların YZ hakkında çalışmaları göstergeler için kanıt sağlamaktadır. Ölçeklerden elde edilen bulguların doğru yorumlanması ve kullanılması ise bilinçli politika üretimi açısından kritik önemdedir. Diğer paydaş gruplarının geri bildirimleri de memnuniyetle karşılanmaktadır. OECD, paydaşlardan alınan geri bildirimler ve sistemli bir güncelleme sürecinin geliştirilmesinin ardından göstergelerin ilk tam sürümünü yayımlayacaktır.

Sonuçlar

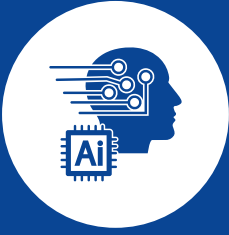
- OECD, kamuya karşı sorumlu bir hükümetler arası kuruluş olarak ve uluslararası karşılaştırmalı beceri değerlendirmeleri konusundaki deneyimine dayanarak, YZ değerlendirmesinde küresel topluluğa güvenilir ve yetkili sonuçlar sunabilecek benzersiz bir konuma sahiptir.
- OECD’nin metodolojisi, mevcut kanıtlardan yararlanarak hem güncel araştırma bulgularını yansıtan hem de geniş bir kitle tarafından anlaşılabilir olan YZ Yetkinlik Göstergeleri üretmektedir. Bu göstergeler, yapay zekâ sistemlerinin dil, problem çözme, sosyal etkileşim gibi alanlarda insanın gösterebildiği beceri düzeylerine ne ölçüde yaklaştığını ortaya koymaktadır.
- Rapor, insan psikolojisinden hareketle geliştirilen dokuz YZ Yetkinlik Göstergesini tanıtmakta ve bu göstergelerin OECD’deki Yapay Zekâ ve Becerilerin Geleceği Ekibi ile 50’den fazla dış uzmanın katkısıyla nasıl oluşturulduğunu açıklamaktadır.

1 OECD, bir yapay zekâ sistemini şu şekilde tanımlar: Açık ya da örtük hedefler doğrultusunda çalışan makine tabanlı bir sistemdir. Bu sistem, aldığı girdilerden hareketle çıkarım yapar ve tahmin, içerik, öneri ya da karar gibi çıktıları nasıl üreteceğini belirler. Ürettiği bu çıktılar fiziksel ya da sanal ortamları etkileyebilir. Farklı yapay zekâ sistemleri, devreye alındıktan sonra sahip oldukları özerklik düzeyi ve yeni durumlara uyum sağlama yetkinliği bakımından birbirinden farklılık gösterebilir [OECD, 2024[2]].

- Göstergeler, YZ sistemlerinin farklı alanlardaki becerilerini değerlendirmek için beş düzeyli ölçekler hâlinde sunulmaktadır. Her düzey, YZ'nin hangi tür görevleri doğru ve tutarlı biçimde yerine getirebildiğini açıklar ve kanıtlarla desteklenir. Mevcut YZ sistemleri, bu ölçeklerde genel olarak 2. ve 3. düzeyler arasında yer almaktadır.
- Göstergeler, YZ'deki ilerlemeyi iş hayatında gerekli olan insan becerileriyle karşılaştırmak için kullanılabilir. Meslekler ve görevler göstergelerle eşleştirildiğinde, YZ'nin hangi görevlerde çalışanlara destek olabileceği ya da bazı görevleri tamamen üstleneceği görülebilir. Bu da mesleklerin zamanla nasıl değişebileceğini analiz etmek için bir temel sağlar. Ayrıca göstergeler, YZ yetkinliklerinin ekonomi genelinde nasıl ve hangi sınırlar içinde kullanılacağına dair değer temelli tartışmaları da destekler.
- Göstergeler, YZ'nin eğitim üzerindeki etkilerini anlamak için de kullanılabilir. Eğitimde hangi öğretmen görevlerinin değişebileceğini, öğrencilerin hangi bilgi ve becerileri öğrenmesi gerektiğinin yeniden düşünülmesi gerekip gerekmediğini değerlendirmek için bir çerçeve sunar. Göstergeler değer temelli kararları belirlemez; ancak eğitimin amacı ve sunuluş biçimi açısından teknik olarak hangi değişimlerin mümkün olduğunu ortaya koyarak müfredat, öğretmen rolleri ve öğrenci yeterliklerine ilişkin gelecekteki tartışmalara katkı sağlar.

Referanslar:

- [1] OECD (2025), AI and the Future of Skills Volume 3: The OECD AI Capability Indicators, OECD Publishing.
- [2] OECD (2024), Explanatory memorandum on the updated OECD definition of an AI system, OECD Artificial Intelligence Papers, No. 8, OECD Publishing, Paris, <https://doi.org/10.1787/623da898-en>. OECD'nin



1. MEVCUT YAPAY ZEKÂ YETKİNLİKLERİNE GENEL BAKIŞ

Eğitim Araştırmaları ve Yenilik Merkezi (CERI) bünyesinde yürütülen Yapay Zekâ ve Geleceğin Becerileri (AIFS) Projesi, yapay zekâ (YZ) ve robotik yetkinlikleri sistematik biçimde ölçmek ve bunları insan becerileriyle karşılaştırmak amacıyla bir çerçeve sunmaktadır. Bu bölümde, OECD'nin Yapay Zekâ Yetkinlik Göstergelerinin her biri açısından YZ performansına ilişkin genel bir değerlendirme yapılmaktadır. Öncelikle, karşılaştırmalı bir tablo tanıtılmakta ve bu tablonun anlaşılması için gerekli bilgiler sunulmaktadır. Tablo, her bir alanda YZ'nin mevcut düzeyini göstermekte ve Kasım 2024 itibarıyla en ileri düzeydeki YZ sistemlerinin sahip olduğu yetkinlik türlerini tanımlamaktadır. Tablonun altında yer alan kısa yorum bölümünde ise OECD uzman grubunun ilgili değerlendirmeyi hangi gerekçelerle yaptığı ve YZ sistemlerinin bir sonraki düzeye ilerleyebilmesi için hangi yetkinlikleri geliştirmesi gerektiği açıklanmaktadır.

Mevcut Yapay Zekâ Yetkinliklerine İlişkin Karşılaştırmalı Tablo

Tablo 1.1, en gelişmiş yapay zekâ (YZ) sistemlerinin mevcut yetkinliklerine genel bir bakış sunmaktadır. Her bir alanda YZ sistemlerine atfedilen mevcut düzey, ilgili düzeydeki sistemlerin sahip olduğu yetkinlik türünü tanımlayan açıklamanın yanında belirtilmiştir. Tablonun altında yer alan yorum bölümünde ise OECD uzman grubunun YZ sistemlerini neden bu düzeyde derecelendirdiği ve bir sonraki düzeye ilerleyebilmeleri için hangi yetkinliklerin geliştirilmesi gerektiği kısaca açıklanmaktadır.

OECD, YZ yetkinliklerinin ilerlemelerini alan dışındaki kişiler tarafından da anlaşılabilir şekilde ifade edebilmek amacıyla beş düzeyli ölçekler geliştirmiştir. Bu ölçekler, tüm YZ sistemi türlerini kapsayacak şekilde tasarlanmıştır. Mevcut derecelendirmeler; dar kapsamlı sembolik YZ sistemlerini, nöro-sembolik sistemleri, büyük dil modellerini (LLM'ler), sosyal ajan tabanlı (agentic) ve ilgili alanlarda en ileri düzeyde bulunan robotik sistemleri içermektedir. Ölçeğin bir ucunda yer alan 1. düzey, günümüz YZ sistemleri için uzun süredir çözülmüş ve tartışmasız biçimde temel nitelikte olan yetkinlikleri temsil etmektedir. Diğer uçta yer alan 5. düzey ise ilgili insan yeteneğinin tüm yönlerini taklit edebilen YZ sistemlerini ifade etmektedir. Aradaki üç düzey, YZ'nin çeşitli boyutlarda insan performans düzeyine eşdeğer bir seviyeye doğru nasıl geliştiğini göstermektedir.

OECD, ölçeklerin geliştirilme yaklaşımını 2. bölümde ve eşlik eden teknik raporda [OECD, 2025[1]] daha ayrıntılı biçimde açıklamaktadır. YZ sistemlerine ilişkin değerlendirmeler, Kasım 2024 itibarıyla alanın en ileri YZ performansını yansıtmaktadır.

Bir YZ sisteminin belirli bir düzeyde değerlendirilebilmesi için, o düzeyde tanımlanan yetkinliklerin büyük bölümünü tutarlı ve güvenilir biçimde sergilemesi gerekmektedir. Örneğin, uzmanlarımız büyük dil modellerinin (LLM) Dil Ölçeğini 2. ve 3. düzey arasında değerlendirmiştir. LLM'ler 3. düzeyde tanımlanan dil yetkinliklerinin birçok yönüne sahiptir. Bununla birlikte, sağlam analitik akıl yürütme yapamama, yanlış bilgi üretme (halüsinasyon) eğilimi ve dinamik biçimde öğrenememe gibi sınırlılıklar, bu sistemlerin daha üst düzeye yerleştirilmesini engellemektedir. Ancak 3. düzeydeki diğer birçok dil yetkinliğini büyük ölçüde yerine getirdikleri için 3. düzeyde değerlendirilmiştir.

Günümüz YZ sistemlerinin öne çıkan bir zayıflığı olan LLM'lerdeki kalıcı halüsinasyon sorunu, ölçekler boyunca farklı biçimlerde karşımıza çıkmaktadır. Bilgi, Öğrenme ve Hafıza Ölçeğinde halüsinasyon sorununun 5. düzeyde giderileceği belirtilirken, Dil Ölçeğinde eleştirel düşünmenin 5. düzeyde ortaya çıkacağı ifade edilmektedir; Üstbilis ve Eleştirel Düşünme Ölçeğinde ise bilginin eleştirel değerlendirilmesinin 3. düzeyde yer alacağı belirtilmektedir. Ölçekler arasındaki bu farklılık, söz konusu sorunun çözümüne ilişkin zorluk derecesine dair farklı öngörülerini yansıtmaktadır ve gelecekteki sürümlerde uyumlaştırılması gerekebilir. Bununla birlikte, ölçeklerin önemli işlevlerinden biri, halüsinasyonun YZ'nin insan düzeyine ulaşmasının önündeki pek çok zorluktan yalnızca biri olduğu göz önünde bulundurulmalıdır.

Uzmanlarımızın mevcut tüm YZ sistemlerini 2. ve 3. düzeylerde değerlendirmiş olması, ölçeklerin ilk kez oluşturulma yaklaşımının dolaylı bir sonucudur. Ölçekler, her bir yetkinlik alanında YZ'nin ilgili insan yeteneklerinin tüm yönlerine ulaşabildiği varsayımsal bir geleceğe doğru yaşanan temel gelişim aşamalarını ifade etmeyi amaçlamaktadır. Her bir ölçek, alandaki temel gelişim basamaklarını tanımlamaktadır. Hâlihazırda gerçekleştirilmiş olan gelişmeler alt düzeylerde; henüz tamamlanmamış olanlar ise üst düzeylerde yer almaktadır. Genel olarak 4. ve 5. düzeyler, YZ sistemlerinin henüz tam anlamıyla tutarlı ve güvenilir biçimde yerine getiremediği yetkinlikleri tanımlamaktadır.

Alandaki birçok araştırmacı, 2024 itibarıyla YZ'nin mevcut durumuna ilişkin değerlendirmelerimiz veya beş düzeyli ölçekler arasındaki yetkinlik dağılımıyla aynı fikirde olmayabilir. OECD, güncelleme sürecine katkı sunmak ve ölçekleri en güncel gelişmelerle daha iyi uyumlaştırmak amacıyla YZ araştırmacılarını organizasyon ile iletişime geçmeye davet etmektedir. Bu bölümde yer alan düzey tanımları kısaltılmıştır; her bir düzeyin tam açıklaması ve ilgili ölçekler 3. bölümde sunulmaktadır.

Tablo 1.1: Mevcut Yapay Zekâ Yetkinlik Düzeylerine Genel Bakış

Alan	Düzy [1-5]	Yetkinlik Açıklaması ²
Dil	3	Bu düzeydeki YZ sistemleri, doğal dil işleme yöntemlerini ve çoklu veri kaynaklarını kullanarak içeriği güvenilir biçimde anlar ve anlamlı çıktılar üretir. Gelişmiş mantıksal ve sosyal akıl yürütme yeteneği gösterir. Metin, konuşma ve görselleri işleyebilir. Birden fazla dili destekler ve yinelemeli öğrenme teknikleri sayesinde yeni durumlara uyum sağlayabilir.
Sosyal Etkileşim	2	YZ sistemleri, duyguları ifade etmek için basit hareketleri bir araya getirebilir ve bu etkileşimlerden yararlanarak gelecekte karşılaşacağı yeni durumlara hazırlanabilir. Olayları hatırlayabilir ve deneyime dayalı olarak sınırlı ölçüde uyum sağlayabilir. Temel sinyalleri tanır; tonlama ve bağlamdan hareketle insan duygularını algılar. Ayrıca bireysel farklılıkları ayırt edebilir ve geçmiş deneyimlerini benzer yeni durumlara uygulayabilir.
Problem Çözme	2	YZ sistemleri, mekânsal ve zamansal ilişkiler gibi nitel akıl yürütmeyi nicel analizle birleştirerek karmaşık mesleki sorunları ele alır. Birden fazla durumu ve bu durumlar arasındaki geçişleri yönetebilir. Ayrıca sistemlerin zaman içinde nasıl değişebileceğini ya da gelişebileceğini öngörebilir.
Yaratıcılık	3	YZ sistemleri, öğrendiği verilerden belirgin biçimde farklılaşan ve alışılmış kalıpların ötesine geçen değerli çıktılar üretebilir. Sahip olduğu becerileri yeni görevlere aktarabilir ve farklı alanlardaki fikirleri bir araya getirerek bir bütün olarak sunabilir.
Üstbiliş ve Eleştirel Düşünme	2	YZ sistemleri, verdiği yanıtların ne kadar doğru ya da yeterli olduğunu kontrol edebilir ve gerektiğinde yaklaşımını değiştirebilir. Tam olarak net olmayan ya da eksik bilgi içeren verilerle çalışabilir. Böyle durumlarda temkinli davranır ve tahminlerini dikkatli biçimde oluşturur. Bilgi eksik olduğunda, hangi konularda yeterli bilgiye sahip olduğunu ve hangi konularda bilgi eksikliği bulunduğunu ayırt edebilir.
Bilgi, Öğrenme ve Hafıza	3	YZ sistemleri, bilgilerin anlam yapısını (semantik yapıyı) çok sayıda örnekten öğrendikleri örüntüler üzerinden temsil eder ve bu bilgiyi yeni durumlara aktarabilir. Başka bir deyişle, kelimelerin ve kavramların anlamını çok sayıda veri içindeki tekrar eden ilişkilerden çıkarırlar. Büyük veri kümelerini işleyerek bağlama duyarlı sonuçlar üretebilirler. Ancak gerçek zamanlı, yani anlık öğrenme yeteneğine sahip değildirler.
Görme	3	YZ sistemleri, hedef nesnenin görünümünde meydana gelen değişikliklere uyum sağlayabilir. Birden fazla alt görevi yerine getirebilir ve verilerde ya da durumlarda önceden bilinen değişikliklerle başa çıkabilir.
Manipülasyon	2	YZ sistemleri, farklı nesne şekilleriyle ve orta derecede esnek malzemelerle çalışabilir. Düşük ya da orta düzeyde dağınıklık bulunan kontrollü ortamlarda görev yapar. Açık alanlarda küçük engellerin etrafından dolaşabilir. Belirli bir alan içinde rastgele yerleştirilmiş nesnelerle çalışabilir. Görevleri zaman kısıtlaması olmadan yerine getirir.
Robotik Zekâ	2	Robotik sistemler, kısmen bilinen, büyük ölçüde sabit ve yarı yapılandırılmış ortamlarda çalışır. Bu ortamlarda belirli ve önceden tanımlanmış bir değişkenlik bulunur. Kısa vadeli, basit ve birden fazla işlev içeren görevleri yerine getirebilir. Bu görevler iyi tanımlanmış olsa da doğası gereği bir miktar belirsizlik içerir. Sınırlı düzeyde insan etkileşimine girebilir; örneğin basit arayüzler üzerinden iletişim kurabilir. Tanıdık görev ortamlarında ortaya çıkan bazı beklenmedik sonuçları yönetebilir. Etik açıdan ise çok az ya da hiç sorun içermeyen durumlarla karşılaşır.

² Karşılaştırmalı tabloda yer alan açıklamalar, 3. bölümde sunulan ilgili ölçek düzeylerinin kısaltılmış versiyonlarıdır.

Mevcut Derecelendirmelere İlişkin Yorumlar

Dil

Yukarıda belirtildiği gibi, günümüzdeki en gelişmiş büyük dil modelleri [LLM'ler] — örneğin ChatGPT'de kullanılan GPT-4o [OpenAI, 2024a[2]] — 3. düzeyin alt sınırında değerlendirilmektedir. LLM'ler, geniş bilgi birikimini kullanma, birden fazla dilde çalışma ve modelin sonradan yeniden eğitilmesiyle performansını artırma konusunda başarılıdır. Ancak adım adım ve tutarlı biçimde akıl yürütme gereken durumlarda zorlanırlar. Yapılandırılmış analiz yapmakta sınırlıdırlar ve zaman zaman yanlış bilgiler üretebilirler. Bu durum, ilerlemenin önündeki başlıca engellerden biridir.

Sosyal Etkileşim

GPT-4o ve benzeri LLM'ler, güçlü sosyal bellek yetkinlikleri nedeniyle Sosyal Etkileşim Ölçeğinde 2. düzeyde derecelendirilmiştir. Bununla birlikte, bedensel [embodied] değildirler, kimlik algısına sahip değildirler ve sosyal algı yetkinlikleri sınırlıdır. Sony'nin AIBO gibi sosyal robotları da 2. düzey sistemlerdir; ancak bu sistemler LLM sistemlerinden farklı yetkinliklere sahiptir. Bu robotlar bedensel varlığa ve temel algı ile kimlik özelliklerine sahip olmakla birlikte, problem çözme yetkinlikleri LLM sistemlerine kıyasla daha sınırlıdır.

Problem Çözme

Sembolik yapay zekâ sistemleri, lojistik planlama ve model doğrulama gibi dar alanlarda insanüstü performans gösterebilir. Bu nedenle 2. düzey sistemler olarak değerlendirilmektedir. Büyük dil modelleri [LLM'ler], doğal dilde tanımlanan problemleri çözme gibi 3. düzeydeki bazı gereklilikleri yerine getirse de yanlış bilgi üretme eğilimleri nedeniyle kırılan bir yapı sergilerler. Uzmanlarımıza göre bu durum, 2024'ün sonlarında kullanıma sunulan GPTo1'in ön sürümü gibi erken dönem "akıl yürütme" modelleri için de geçerlidir [OpenAI, 2024b[3]]. Daha gelişmiş "akıl yürütme" modelleri olan GPTo3 [OpenAI, 2025[4]] ve DeepSeek R1 V3 [DeepSeek-AI, 2025[5]] için bu durumun devam edip etmediği ise OECD YZ Yetkinlik Göstergelerinin tam sürümünde incelenmektedir.

Yaratıcılık

Mevcut YZ sistemleri, insanlar için değerli, kısmen yeni ve zaman zaman şaşırtıcı çıktılar üretebilmektedir. 3. düzey bir sisteme örnek olarak, nöro-sembolik altyapı kullanarak problemler için verimli ve beklenmedik stratejiler geliştiren Google'ın AlphaZero sistemi [Silver et al., 2017[6]] verilebilir. LLM'lerin olasılıksal bir mimariye dayanması ve insan tarafından üretilmiş verilerle eğitilmesi [yani önceki insan içeriklerine dayanması], bu sistemlerin mevcut insan bilgisinden önemli ölçüde farklı çıktılar üretmesini mümkün kılmamaktadır. Bununla birlikte, bu çıktılar çoğu zaman yararlı ve zaman zaman yenilikçi olduğundan, LLM'ler tipik olarak 2. düzey sistemler olarak değerlendirilmektedir.

Üstbilgi ve Eleştirel Düşünme

En gelişmiş LLM'ler genellikle Üstbilgi ve Eleştirel Düşünme Ölçeğinde 2. düzeyde performans göstermektedir. Kendi anlayışlarını izleyebilir ve mevcut probleme kendi yaklaşımını uygulayabilirler. Ancak, tanıdık olmayan bilgileri bütünleştirme veya kendi bilgi durumlarını değerlendirme konularında zorlanmaktadır; bu yetkinlikler 3. düzey sistemler için gereklidir. Değerlendirmenin yapıldığı dönemde, etken sistemler de genellikle 2. düzeyde performans göstermiştir; bu durum, YZ'nin kendi akıl yürütmesini izleme ve uyarlamalı biçimde düzenleme yeteneğindeki süregelen sınırlılıkları yansıtmaktadır. 2025 yılında yayımlanan ajan sistemleri, OECD'nin bir sonraki baskısında değerlendirilecektir.

Bilgi, Öğrenme ve Hafıza

LLM'ler ve üretken yapay zekâ sistemleri bu alandaki en gelişmiş sistemlerdir. Depolanmış bilgiden genelleme yapabildikleri için 3. düzeyde değerlendirilmektedirler. Ancak henüz dünya ile etkileşim içinde sürekli öğrenme ya da kendi bilgi eksikliklerinin farkında olma gibi 4. düzey özelliklerine sahip değildirler.

Görme

En ileri düzey YZ görme sistemleri 3. düzeydedir. Uzmanlarımız sınırlı 4. düzey yetkinlikler sergileyen az sayıda sistem tespit etmiştir; ancak bu performans henüz herhangi bir sistemin 4. düzey olarak derecelendirilmesini sağlayacak kadar güvenilir değildir. 3. düzey sistemler, sınırlı sayıda veri türünü güvenilir biçimde işleyebilir ve (farklı) aydınlatma koşullarındaki, hedef nesnelerin şekil ve görünümündeki makul değişimlerle başa çıkabilir. 4. düzey sistemlerin aksine, mevcut YZ görme sistemleri kendi geri bildirimlerinden öğrenerek performanslarını iyileştirmekte ve büyük ölçekli değişiklikler yapmakta da yetersiz kalmaktadır.

Manipülasyon

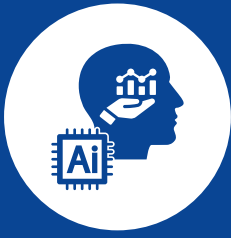
Manipülasyon sistemleri 2. düzeyde değerlendirilmiştir. Bu düzeydeki tipik ve en gelişmiş sistemlerden biri, yüksek derecede kontrollü üretim ortamlarında kullanılan robotik kollarıdır. Buna karşılık, 3. düzey sistemler orta derecede dağınık ve dinamik ortamlarda; değişken şekil, boyut ve ağırlıktaki nesnelerle çalışabilmektedir. Manipülasyon sistemleri hâlâ insan düzeyine ulaşmaktan uzaktır. Bununla birlikte, nesne ve ortamların standartlaştırılabildiği alanlarda -örneğin fabrikalarda olduğu gibi- bu sistemler yine de insan istihdamını ve beceri talebini etkileyecektir.

Robotik Zekâ

En gelişmiş robotik sistemler; otonom teslimat robotları ve endüstriyel otomasyon sistemleridir; uzmanlarımız bu sistemleri 2. düzeyde derecelendirmiştir. Bu sistemler yapılandırılmış ortamlarda ve önceden tanımlanmış görevlerde başarılı performans göstermektedir. Robotik sistemler, çok adımlı görevleri yerine getirme veya insanlarla güvenilir biçimde iş birliği yapma konusunda henüz yeterli değildir; bu yetkinlikler 3. düzeye ulaşmak için gereklidir.

Referanslar:

- [5] DeepSeek-AI [2025], "DeepSeek-R1: Incentivizing Reasoning Capability in LLMs via Reinforcement Learning", arXiv, Vol. 2501.12948, <https://doi.org/10.48550/arXiv.2501.12948>.
- [1] OECD [2025], AI and the Future of Skills Volume 3: The OECD AI Capability Indicators, OECD Publishing.
- [4] OpenAI [2025], OpenAI o3 and o4-mini System Card, 16 April, OpenAI, <https://openai.com/index/o3-o4-mini-system-card>.
- [2] OpenAI [2024a], "GPT-4o System Card, OpenAI", arXiv, Vol. 2410.21276, <https://arxiv.org/abs/2410.21276>.
- [3] OpenAI [2024b], "OpenAI o1 System Card, OpenAI", arXiv, Vol. 2412.16720, <https://arxiv.org/abs/2412.16720>.
- [6] Silver, D. et al. [2017], "Mastering Chess and Shogi by Self-Play with a General Reinforcement Learning Algorithm", arXiv, Vol. 1712.01815, <https://arxiv.org/abs/1712.01815>.



2. YAPAY ZEKÂ YETKİNLİKLERİNİ ÖLÇMEYE YÖNELİK BİR ÇERÇEVE OLUŞTURULMASI

OECD'nin Eğitim Araştırmaları ve Yenilik Merkezi (CERI) bünyesinde yürütülen Yapay Zekâ ve Geleceğin Becerileri (AIFS) Projesi, yapay zekâ (YZ) ve robotik yetkinlikleri sistematik biçimde ölçmek ve bunları insan becerileriyle karşılaştırmak amacıyla bir çerçeve sunmaktadır. Bu bölüm, bu yayında sunulan beta göstergelerin geliştirilmesine yönelik OECD metodolojisini açıklamaktadır. Göstergelerin, politika yapıcılara YZ gelişmeleri ile bunların toplum, çalışma hayatı ve eğitim üzerindeki etkileri hakkında açık ve kanıta dayalı bilgiler sunulabileceğini ortaya koymaktadır. OECD ayrıca, göstergelerin yapay genel zekâyâ (AGI) doğru ilerlemeyi ölçmek için ilkeli ve ihtiyatlı bir çerçeve sunduğunu ileri sürmektedir. Bu çerçeve, politika yapıcılara teknoloji liderleri ve araştırmacılar tarafından yapılan YZ ilerlemesine ilişkin iddiaları bağımsız biçimde doğrulayabilmesi için gerekli araçları sağlamaktadır.

YZ, Yetkinliklerine İlişkin Anlayışımızın Ötesinde İlerleme Kaydetmiştir

Kaynağa bağlı olarak yapay zekâ ya dünyayı kurtaracak ya da yok edecektir. Abartı ve korku söylemlerinin hâkim olduğu bir ortamda, YZ'nin gerçek yetkinliklerine ilişkin açık, güvenilir ve dengeli bilgiler dikkat çekici derecede yetersiz kalmıştır. Hatta YZ geliştiricileri dahi mevcut sistemlerin gerçek yetkinliğini ya da bu yetkinliğin ne kadar hızlı ilerlediğini tam olarak anlamakta zorlanmaktadır. Bağımsız ve yetkili bir uluslararası kuruluş olarak OECD, bu bilgi boşluğunu doldurmak için uygun bir konumdadır. Karşılaştırmalı değerlendirme konusundaki deneyimi, önde gelen bilgisayar bilimcileri ve mühendislerle yürüttüğü kapsamlı iş birlikleri ve uluslararası perspektifi sayesinde OECD, YZ'nin gerçek dünya performansına ilişkin titiz, kanıta dayalı ve açık tespitler sunan özgün bir metodoloji geliştirmiştir. Bu çerçeve, politika yapıcılara giderek karmaşılaşan teknolojik ortamda yol almaları ve geleceğe yönelik stratejiler geliştirmeleri için ihtiyaç duydukları açıklığı sağlamaktadır.

2022'de ChatGPT'nin kullanıma sunulmasıyla birlikte YZ ve robotik sistemler hızla gelişmiş; dünya genelindeki politika yapıcılar bu teknolojilerin yetkinliğini değerlendirme gerekliliğini fark etmiştir. Örneğin, Avrupa Birliği Yapay Zekâ Tüzüğü [EUR-Lex, 2024[1]] düzenli izleme yükümlülüğü getirmektedir. Benzer şekilde, OECD Konseyi'nin Yapay Zekâ Tavsiye Kararı [OECD Legal Instruments, 2024[2]] ve 2025 Paris Yapay Zekâ Zirvesi [La Maison Élysée, 2025[3]], YZ'nin iş gücü piyasası üzerindeki etkilerinin anlaşılmasının önemini vurgulamaktadır.

Artan ilgiye rağmen kalıcı bir boşluk bulunmaktadır: YZ yetkinliklerini hem anlaşılır hem de politika açısından anlamlı bir biçimde kapsamlı şekilde ölçen sistematik bir çerçeve mevcut değildir. Bu boşluğu gidermek amacıyla OECD, YZ yetkinliklerini değerlendirmeye yönelik bir çerçeve geliştirmiş ve beta Yapay Zekâ Yetkinlik Göstergelerini sunmuştur [geliştirme sürecinin genel görünümü için Şekil 2.1'e bakınız]. OECD göstergelerin şu özellikleri taşımasını amaçlamaktadır:

1. **Anlaşılır** – Yapay zekânın güçlü yönlerini ve sınırlarını açık ve net bir şekilde aktarmak.
2. **Politika odaklı** – Yapay zekânın eğitim, istihdam ve ekonomi üzerindeki etkisine dair derinlemesine bilgiler sunmak.
3. **Kapsamlı** – Yapay zekâ yeteneklerinin tüm önemli yönlerini ele almak.
4. **Güncel** – Sistematik güncellemelerle zaman içindeki gelişmeleri takip etmek.

YZ yetkinliklerinin insan becerileri ile ilişkilendirilmesi, politika yapıcılarının YZ'nin eğitim, çalışma hayatı ve gündelik hayat içindeki potansiyel rolünü değerlendirmelerine imkân tanımaktadır. MLCommons [MLCommons, 2025[4]] ve Stanford AI Index [Maslej et al., 2025[5]] gibi mevcut çerçeveler YZ yetkinliklerini yalnızca kıyas ölçütü [benchmark] performansı üzerinden ele almakta, insan becerileri ile karşılaştırma yapılmamaktadır. Bu kıyas sonuçları, alan dışındaki kişiler için belirsizdir; hatta YZ araştırmacıları açısından bile bu sonuçların gerçek dünya görevleriyle nasıl ilişkilendirildiği yeterince açık değildir.

Mevcut yapay zekâ ölçütlerinin [benchmark] sınırlılıklarına ilişkin bir tartışmada, 2025 Yapay Zekâ Endeksi [2025 AI Index] raporunun yazarları, "YZ sistemlerinin yetkinliklerini gerçekten değerlendirebilmek için daha titiz ve kapsamlı değerlendirmelere ihtiyaç vardır." ifadesini kullanmıştır [Maslej vd., 2025[5]]. OECD'nin çerçevesi, yapay zekâ yetkinliklerini işgücü piyasasında kullanılan insan yeteneklerinin tamamıyla karşılaştırma çabası bakımından özgündür. Bu yaklaşım, özellikle ileri düzeyler için yeterli yapay zekâ ölçütlerinin çoğu durumda hâlâ eksik olduğunu kabul etmektedir.

Burada sunulan göstergeler, yapay zekâ yetkinliklerini hayat ve istihdam açısından ilgili tüm insan beceri alanlarında sistematik biçimde değerlendirebilecek testlerin belirlenmesi ya da geliştirilmesi için kavramsal bir çerçeve sunmaktadır. İnsan performansını temel almadan yalnızca ölçüt sonuçlarına dayanan çerçeveler, yapay zekâdaki hızlı gelişmeler karşısında kısa sürede geçerliliğini yitirme riski taşımaktadır. Yapay zekâ yetkinliklerini insan yetenekleriyle karşılaştırarak değerlendiren bir çerçevenin ek bir avantajı ise insan yeteneklerinin zaman içinde görece sabit kalmasıdır. Başka bir ifadeyle, yapay zekâ insan performansının tüm kapsamını gerçekten aşmadığı sürece bu çerçeve hızlı ilerlemeler karşısında istikrarlı ve bilgilendirici olmaya devam edecektir.

Anthropic'in yayımladığı Yapay Zekâ Ekonomik Endeksi [Anthropic AI Economic Index] çalışması [Handa vd., 2025[6]], yapay zekâ gelişmelerinin ekonomi üzerindeki etkisini analiz etmek amacıyla Claude.ai'nin yetkinliklerini bazı insan görevleriyle ilişkilendiren yenilikçi bir yaklaşım benimsemiştir. Bu analiz, Claude Büyük Dil Modelleri [LLM'ler] ile kullanıcılar arasındaki milyonlarca etkileşimin incelenmesi yoluyla gerçekleştirilmiş ve O*NET'te yer alan görevlerle ilişkilendirilebilen faaliyetler değerlendirilmiştir. Ancak bu yaklaşım, LLM'lerle yapılan sohbet temelli etkileşimlerin analizi ile sınırlı kalmıştır. Dolayısıyla, yapay zekâ performansını mesleklerde kullanılan insan yeteneklerinin tamamıyla karşılaştırmayı amaçlamamıştır. OECD'nin Yapay Zekâ Yeterlilik Göstergelerini kullanarak yapay zekâ gelişmelerinin meslekler üzerindeki etkisini analiz etmeye yönelik kendi yaklaşımı bu raporun 4. bölümünde açıklanmaktadır.

Mevcut çerçevenin sınırlılıklarını kabul eden OECD, bu göstergeleri beta sürüm olarak sunmakta ve yapay zekâ ile insan psikolojisi alanındaki uzmanlarla iş birliğine davet etmektedir. Aşağıda sunulan yapay zekâ sistemlerine ilişkin derecelendirmeler Kasım 2024'te kesinleştirilmiş olup, o tarihteki en ileri durumu yansıtmaktadır. Gelecekte yapılacak iyileştirmeler, göstergeleri yapay zekâ gelişmelerini izlemek için daha hassas ve duyarlı bir araç hâline getirecektir.

Bu bölüm, göstergelerin ortaya çıkış gerekçesini ve oluşturulma yöntemini tanıtmaktadır [göstergelerin geliştirilme sürecine genel bakış için Şekil 2.1'e bakınız]. Ayrıca, akran değerlendirme sürecinde belirlenen beta ölçeklerin sınırlılıklarını da ele almaktadır. Yönteme ilişkin daha ayrıntılı bir tartışma, bu raporla birlikte yayımlanan teknik ciltte yer almaktadır [OECD, 2025[7]]. Ayrıca bu bölüm, politika yapıcılarının hızlı yapay zekâ gelişmelerinden doğan soruları yanıtlamalarına yardımcı olmak üzere göstergeleri nasıl kullanabileceklerine ilişkin bir değerlendirme ile sona ermektedir.

2. bölüm, dokuz ölçeğin tamamını, her düzeydeki performans tanımlayıcılarını ve en ileri düzey yapay zekâ sistemlerine ilişkin derecelendirmeleri sunmaktadır. Her bir gösterge ayrıca ilgili alanda neyin ölçülmesinin önemli olduğuna ve yapay zekâ performansına ilişkin mevcut kanıtlara dair bir tartışma içermektedir.

3. bölüm, araştırmacıların ve politika yapıcılarının bu ölçekleri, yapay zekânın işgücü ve eğitim üzerindeki etkilerini kanıta dayalı olarak analiz etmek için nasıl kullanabileceklerine dair somut örnekler sunmaktadır.

Metodoloji: Özgün ve yenilikçi bir yaklaşım

Görevlerden yetkinliklere: Yapay zekâyı değerlendirmede yeni bir yaklaşım

Yapay zekânın toplum üzerindeki etkisini anlamak, yalnızca belirli iş veya görevleri yerine getirip getirmediğini değerlendirmekten ibaret değildir. Görev temelli analizler işgücü ekonomisinde merkezi bir yer tutmakla birlikte, YZ'ye uygulandığında çoğu zaman netlik ve tutarlılık sağlamamaktadır; zira insan görevlerinin büyük bölümü birden fazladır ve birbirine bağlı yetkinlik gerektirmektedir.

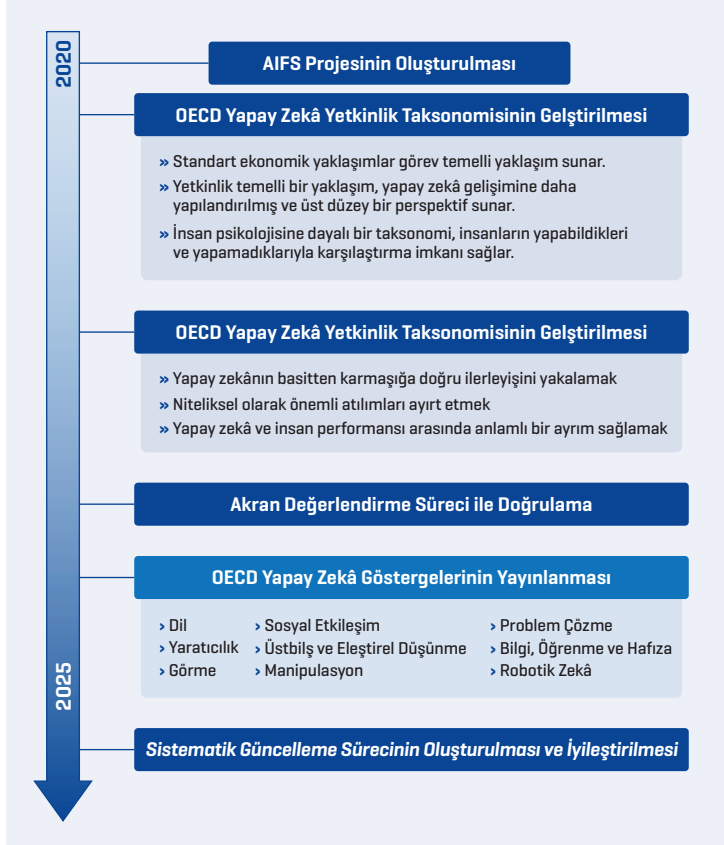
Son yıllarda ekonomi literatürü, YZ'nin işgücü üzerindeki etkisini meslek temelli analizlerden görev temelli analizlere kaydırmıştır. Bu yaklaşım, mesleklerin farklı görevlerin birleşiminden oluştuğunu ve herhangi bir YZ tekniğinin ya da teknolojinin bu görevleri farklı biçimlerde etkileyebileceğini kabul etmektedir [Autor, Levy and Murnane, 2023[8]]³. Ancak görev taksonomileri, modern ekonomide binlerce hatta on binlerce farklı görevi içerdiğinden, bu yaklaşım YZ gelişmelerini açıklamak için yeterince açık ve anlaşılır bir çerçeve sunmamaktadır. Buna karşılık, temel yetkinlik türlerinin sayısı görece sınırlıdır.

3 Büyük dil modellerinin [LLM'ler] meslekler üzerindeki etkisini görev temelli olarak inceleyen çalışmalara örnek olarak bkz. Eloundou et al. [2023[12]] veya Handa et al. [2025[6]].

Ayrıca tek bir görev genellikle birden fazla beceri gerektirir ve bu beceriler farklı YZ tekniklerinden farklı biçimlerde etkilenebilir. Bu nedenle görev temelli analizler, YZ'nin temel ilerleme alanlarını ve sınırlılıklarını açık ve sistematik biçimde ortaya koymakta yetersiz kalmaktadır.

OECD bunun yerine yetkinlik temelli bir çerçeve benimsemektedir. Bu yaklaşım, parçalı görevlerden ziyade muhakeme, dil, sosyal etkileşim ve psikomotor beceriler gibi temel insan yeteneklerine⁴ odaklanmaktadır. İnsan psikolojisine dayanan bu yaklaşım, YZ gelişimine ilişkin yapılandırılmış ve üst düzey bir bakış açısı sunmaktadır. YZ gelişimini insan yeteneklerinin tamamı üzerinden gösterebilmek amacıyla OECD, Şekil 2.2'de sunulan dokuz Yapay Zekâ Yetkinlik Göstergesini geliştirmiştir.

Şekil 2.1: OECD Yapay Zekâ Yetkinlik Göstergelerinin Geliştirilme Süreci



Kaynak:

OECD [2025], Introducing the OECD AI Capability Indicators, OECD Publishing, Paris, <https://doi.org/10.1787/be745f04-en>.

Bu görsel OECD'nin orijinal çalışmasının Türkçe çeviri ve uyarlamasıdır. Orijinal çalışma ile çeviri arasında herhangi bir farklılık olması durumunda yalnızca orijinal metin geçerlidir.

Şekil 2.2: OECD Yapay Zekâ Yetkinlik Göstergeleri



Kaynak:

OECD [2025], Introducing the OECD AI Capability Indicators, OECD Publishing, Paris, <https://doi.org/10.1787/be745f04-en>.

Bu görsel OECD'nin orijinal çalışmasının Türkçe çeviri ve uyarlamasıdır. Orijinal çalışma ile çeviri arasında herhangi bir farklılık olması durumunda yalnızca orijinal metin geçerlidir.

4 Bu rapor boyunca "yetkinlik" terimi YZ'nin yapabildiği temel şeyleri; "yetenek" terimi ise insanların yapabildiği temel şeyleri ifade etmek üzere kullanılmaktadır.

Göstergelerin Oluşturulması ve Geliştirilmesi

Beş Düzeyli Ölçeklerin Oluşturulması

OECD, basit görevlerden karmaşık görevlere doğru yapay zekâ ilerlemesini yansıtan; niteliksel açıdan önemli atılımları ayırt eden, ve yapay zekâ yetkinliklerinin insan yetenekleriyle karşılaştırmalı olarak anlamlı biçimde anlaşılmasını sağlayan kapsamlı göstergeler geliştirmek amacıyla 30 bilgisayar bilimci, psikolog ve ölçme-değerlendirme uzmanından oluşan çekirdek bir grupta iş birliği yapmıştır.

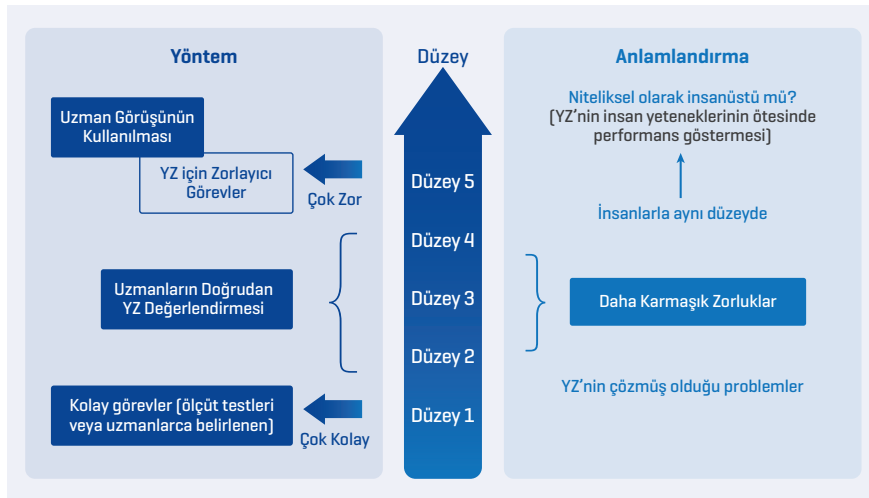
OECD, yapay zekâ sistemleri için görevlerin artan zorluk derecesini temsil etmek üzere beş düzeyli ölçekler oluşturmuştur [bkz. Şekil 2.3]. Ölçekler, tüm yapay zekâ ve robotik sistem türlerini kapsayacak şekilde tasarlanmıştır. Ölçeklerin mevcut sürümünde; dar kapsamlı sembolik yapay zekâ sistemleri, nöro-sembolik sistemler, büyük dil modelleri [LLM'ler], sosyal etkenler ve robotik sistemler, farklı alt alanlarda ortaya çıktıkları biçimiyle dikkate alınmaktadır. Her bir ölçekte 1. düzey, çözülmüş yapay zekâ zorluklarını [örneğin Google Search'ün bilgi erişim yetkinliği] yansıtırken; 5. düzey, ilgili insan yeteneğinin tüm yönlerini simüle eden performansı temsil etmektedir.

OECD'nin beş düzeyli ölçekler geliştirmesinin temel motivasyonu, yapay zekâ yetkinliklerindeki ilerlemeyi alan dışındaki kişiler için de anlaşılabilir bir şekilde aktarmaktır. Her bir ölçek genellikle, yapay zekâ açısından farklı zorluk derecelerini yansıtan birden fazla boyut içermektedir. Bu düzeyler yalnızca kademeli performans artışlarını değil, açık niteliksel farklılıkları temsil etmektedir. Her bir gösterge, yapay zekâ sistemlerinin mevcut performans düzeyini beş düzeyli ölçek üzerinde tanımlamaktadır.

İdeal olarak, her bir ölçeğin hem yapay zekâ sistemleri hem de insanlar için karşılaştırılabilir sonuçlar üreten standartlaştırılmış testlerle desteklenmesi beklenirdi; ancak birçok alanda bu tür testler henüz mevcut değildir. Bu nedenle OECD, mevcut tüm kanıtlardan yararlanmakta ve ortaya çıkan boşlukları uzman yargısıyla tamamlamaktadır. Kanıtlar; standartlaştırılmış testlerin yanı sıra yarışmalar ile bireysel yapay zekâ sistemlerinin performans analizlerini de içerebilir. Uzman katkısı ayrıca, mevcut testlerin uygunluğunu ve yorumlanmasını eleştirel biçimde değerlendirmek amacıyla kullanılmakta; böylece performans iddialarının sağlam ve kanıta dayalı temellere dayanması sağlanmaktadır.

Uluslararası Öğrenci Değerlendirme Programı [PISA] ve Yetişkin Becerilerinin Uluslararası Değerlendirilmesi Programı [PIAAC] gibi standartlaştırılmış insan değerlendirme araçları zaman zaman yapay zekâyı değerlendirmek amacıyla kullanılmıştır. Bununla birlikte, bu araçların temel değeri, yapay zekâ sistemleri açısından insan yeteneklerinin tam kapsamını daha iyi yansıtabilecek yeni YZ testlerinin tasarımına ilham vermelerindendir.

Şekil 2.3: Yapay zekânın beş düzeyine genel bakış



Kaynak:

OECD [2025], Introducing the OECD AI Capability Indicators, OECD Publishing, Paris, <https://doi.org/10.1787/be745f04-en>. Bu görsel OECD'nin orijinal çalışmasının Türkçe çeviri ve uyarlamasıdır. Orijinal çalışma ile çeviri arasında herhangi bir farklılık olması durumunda yalnızca orijinal metin geçerlidir.

Gelecekte ölçekler, insan yeteneklerinin kapsamının ötesine geçen yapay zekâ yetkinliklerini de içerecek şekilde genişletilebilir; bu durum, yapay zekâ performansının niteliksel olarak insanüstü yönleri olarak tanımlanabilir. Bu yaklaşım, niceliksel insanüstü performanstan farklıdır. Niceliksel insanüstü durumda yapay zekâ, hız, doğruluk veya veri kapsamı açısından insanı aşmaktadır; ancak temelde insanların da yapabildiği şeyleri yapmaktadır. İnsanüstü yetkinlikler kavramı yapay zekâyâ özgü değildir. Tarihsel olarak mikroskoplar, hesap makineleri ve elektrikli aletler gibi yenilikler insanlara doğal sınırlarının ötesinde performans sergileme imkânı sağlamıştır. Bununla birlikte, yapay zekâ yalnızca mekanik veya hesaplama üstünlüğü değil, zekâ gerektiren davranışları içeren yeni yetkinlik sorularını gündeme getirmektedir.

Göstergelerin Doğrulanması

Süreç beş yıl boyunca gelişmiş ve 2024 yılının sonlarında yapılandırılmış bir akran değerlendirme süreci ile mevcut halini almıştır. Bu değerlendirme; yapay zekâ, psikoloji ve eğitim alanlarından 25 araştırmacıyı kapsamış ve araştırmacılar iki kategoriye ayrılmıştır:

1. Genel değerlendiriciler – Çerçevenin kapsamlılığını değerlendirenler.
2. Alan-özel değerlendiriciler – Kendi uzmanlık alanları içindeki bireysel ölçekleri değerlendirenler.

Uzman geri bildirimleri iyileştirmelere yön vermiş ve mevcut beta sürümün şekillenmesini sağlamıştır. Gelecekteki sürümler; ek kanıtların entegrasyonunu, ölçüm metodolojilerinin standartlaştırılmasını ve politika açısından anlamlılığın güçlendirilmesini içerecektir. OECD, beta göstergeleri daha da geliştirmek amacıyla yapay zekâ araştırmacıları, psikologlar, eğitim uzmanlarından ve ekonomistlerden daha fazla geri bildirim almayı amaçlamaktadır.

Yapay Zekâ Performans Ölçütlerinin Ölçeklerle İlişkilendirilmesi

Yapay zekâ performans ölçütlerini ölçeklerle ilişkilendirmenin amacı, belirli bir zamanda YZ'nin sağlam ve güvenilir performans gösterdiği her bir ölçek üzerindeki en yüksek düzeyi belirlemektir. Ölçek düzeyleri zorluk derecesine göre sıralanmıştır. Buna bağlı olarak, YZ belirli bir düzeyde değerlendirildiğinde, o düzeye kadar olan yetkinliğinin tüm yönlerini sağlam ve güvenilir biçimde simüle edebilmektedir. Bununla birlikte, yapısal sınırlılıklar nedeniyle belirli bir YZ sistemi yetkinliğinin bazı yönlerinden yoksun olabilir. Ölçekler, belirli sistem sınırlamalarını değil, genel olarak alanın en ileri durumunu yansıtmaktadır.

Bazı ölçekler, dar uygulamalardaki alt düzey yetkinliklerin genellenmiş performansla sahip bir sisteme kolayca entegre edilemeyebileceğini vurgulamaktadır. Alt düzeyler, YZ'nin genellikle niceliksel olarak insanüstü olduğu – görevleri insanlardan daha hızlı ve daha doğru biçimde yerine getirdiği – çözülmüş olan problemleri yansıtmaktadır. Orta düzeylerde çoğunlukla standart kıyas ölçütü [benchmark] testleri bulunmaktadır. Üst düzeyler ise hâlen mevcut YZ sistemleri için zorluk teşkil eden, ortaya çıkan veya henüz değerlendirilmemiş yetkinlik boyutlarını kapsamaktadır. Bu üst düzeylere ilişkin olarak, güvenilir performans ölçütleri mevcut olmasa dahi uzman değerlendirmelerine dayalı tahminler yapılmıştır.

Sınırlılıklar

Yapay zekânın değerlendirilmesindeki başlıca zorluklardan biri, farklı yetkinlikler için ölçüm araçlarının eşit ölçüde bulunmamasıdır. Dil ve görme gibi alanlar onlarca yıllık kıyas ölçütlerinden [benchmark] yararlanırken, sosyal etkileşim ve yaratıcılık gibi alanlarda standartlaştırılmış değerlendirme araçları bulunmamaktadır. Bu durumlarda boşluklar uzman yargısı ile doldurulmaktadır.

Tam anlamıyla sistematik bir yapay zekâ değerlendirme çerçevesi henüz mevcut olmamakla birlikte, bu yaklaşım mevcut kanıtları bir araya getirmekte ve gelecekteki gelişim alanlarını vurgulamaktadır. Göstergeler hâlen beta sürüm aşamasındadır ve araştırmacılar ile uzmanlarla iş birliği içerisinde geliştirilmesi planlanmaktadır.

Yapay Zekâ Yetkinlik Göstergelerinin beta sürümü, YZ'nin neyi yapabildiğini ve neyi yapmakta yetersiz kaldığını anlamak için önemli bir temel sunmaktadır. Bununla birlikte, bu çerçevenin daha da geliştirilmesi, daha geniş bir uzman topluluğuyla derinleştirilmiş iş birliği gerektirmektedir. Önümüzdeki dönemde atılacak temel adımlar şunlardır:

- **Kapsamın genişletilmesi:** YZ'nin yetkinlik alanlarını daha kapsamlı biçimde ele almak için yeni yetkinlik ve göstergelerin eklenmesi.
- **Yapının iyileştirilmesi:** Karşılaştırma ölçütlerinde tutarlılığın sağlanması, çok etkenli sistemlerin dikkate alınması ve örtüşen yetkinliklerin yönetilmesi.
- **İnsan performans temel çizgilerinin netleştirilmesi:** Ölçeklerin YZ'yi kimi zaman ortalama, kimi zaman uzman insan performansı ile karşılaştırması nedeniyle, zorluk düzeylerinin anlamlı biçimde yansıtılabilmesi için daha fazla tutarlılık sağlanması ya da gerektiğinde bu ayrımın bilinçli olarak korunması.
- **Dinamik güncellemelerin sağlanması:** Hızlı YZ ilerlemeleri karşısında güncel kalabilmek amacıyla düzenli gözden geçirme süreçlerinin oluşturulması ve performans ölçütlerine ilişkin bir ölçüt havuzunun⁵ sürdürülmesi.

YZ yetkinliklerinin temel insan yetenekleriyle hizalanması yoluyla OECD göstergeleri, özellikle çalışma hayatı ve eğitim gibi alanlarda YZ'nin toplumsal etkisini değerlendirmek için şeffaf ve politika açısından anlamlı bir araç sunmaktadır. Henüz gelişim sürecinde olmakla birlikte, bu çerçeve YZ ilerlemesine ilişkin daha sistematik ve kanıta dayalı bir anlayışa doğru önemli bir adımı temsil etmektedir. Göstergelerin geliştirilmesinde kullanılan metodolojiye ve sınırlılıklarına ilişkin daha ayrıntılı tartışma, bu raporla birlikte yayımlanan teknik ciltte yer almaktadır [OECD, 2025[7]].

Sonraki Adımlar

Beta göstergeler gerekli iyileştirmeler sonrasında olgunlaştırıldığında, OECD ilave uygulama ve geliştirme faaliyetleri yürütmeyi planlamaktadır. Bu faaliyetler, göstergelerin YZ alanındaki hızlı gelişmelere sistematik ve sürdürülebilir biçimde yanıt verilmesini sağlamayı amaçlamaktadır. Ayrıca, YZ araştırmacılarının YZ yetkinliklerini ölçmeye yönelik geçerli ve bilgilendirici testler tasarlayıp uygulamalarını desteklemeyi amaçlamaktadır.

1. Düzenli Güncellemeler

OECD, sistematik bir güncelleme süreci uygulamayı planlamaktadır. Bu kapsamda ölçeklerle ilişkilendirilen mevcut YZ kıyas ölçütü testlerinde elde edilen gelişmiş sonuçlar izlenecektir. Ayrıca bilimsel literatür taranarak ölçeklerle ilişkilendirilmesi gereken yeni kıyas ölçütleri belirlenecektir. Son olarak, bu sonuçlar ölçekler bağlamında YZ'nin mevcut performansına ilişkin açıklamalar hâline getirilecektir. Güncellemeler, OECD'nin YZ araştırmacıları ve psikologlardan oluşan ağı tarafından gözden geçirilecektir.

5 OECD, bu raporla birlikte YZ yetkinliklerini test eden kıyas ölçütlerinden kanıtları sistematik biçimde toplamak amacıyla çevrimiçi bir veri tabanı başlatmıştır. <https://aicapabilityindicators.oecd.org> adresinde YZ araştırmacıları, ölçeklerde yer alan herhangi bir yetkinliği değerlendiren kıyas ölçütlerini ve diğer değerlendirme araçlarını sunabilecektir. Sunulan değerlendirmeler OECD ve uzman grubu tarafından incelenerek ölçeklerin gelecekteki güncellemelerinde kullanılmaya uygunlukları değerlendirilecektir.

OECD, güncelleme metodolojisini 2025 yılı boyunca geliştirmeyi ve ilk güncellemeyi 2026 yılının başında gerçekleştirmeyi planlamaktadır.

2. Yapay Zekâ Atılımlarını Öngörmek

OECD bugüne kadar literatürde ortaya konmuş mevcut YZ yetkinliklerinin tanımlanmasına odaklanmıştır. Göstergelerin YZ'nin olası etkilerini anlamadaki katkısını genişletmek amacıyla OECD, potansiyel atılımları ön görmeye yönelik bir yaklaşım geliştirmeyi planlamaktadır. Bu kapsamda uzman grupları, alandaki temel araştırma gündemlerini analiz ederek göstergeler çerçevesinde potansiyel atılımların kapsamını tanımlamaya davet edilecektir. Amaç, YZ'nin henüz sahip olmadığı yetkinliklere ilişkin araştırmaları inceleyerek hangi alanlarda performans sıçramalarının muhtemel olduğunu belirlemektir. OECD, göstergeleri YZ atılımlarıyla ilişkilendirmeye yönelik ilk çalışmayı 2026 yılında yürütmeyi planlamaktadır.

3. Uzman Anketi

Uzman görüşlerini tamamlayıcı nitelikte, OECD, Chicago Üniversitesi Ekonomi Uzmanları Paneli'ni örnek alarak [Clark Center for Global Markets, 2025[9]], yapay zekânın yeteneklerine ilişkin temel ifadeleri gözden geçirmek ve bu konuda katkı sağlamak amacıyla resmi ve döngüsel bir uzman anketi geliştirecektir. Bu anket, kamuya açık kıyas ölçütlerinin henüz değerlendirmede YZ yetkinlik boyutlarına ilişkin düzenli güncellemeler sağlayacaktır. Böylece, daha önce alan için aşırı zor kabul edilen YZ yetkinliklerine yönelik yeni çalışmaların ortaya çıkmaya başladığı durumlarda erken göstergeler sunacaktır. OECD, panel için uzmanları 2025 yılında davet etmeyi ve paneli 2026 yılında başlatmayı planlamaktadır. Chicago modeline uygun olarak aylık anketler gerçekleştirilecektir.

4. Yeni Ölçüt Testleri ve Yarışmalar

Gelişen YZ yetkinliklerine ilişkin daha kapsamlı bilgi sağlamak amacıyla OECD, mevcut kıyas ölçütlerinin ve yarışmaların yetersiz kaldığı ölçek düzeylerini belirleyecektir. Bu düzeyler, yeni YZ yetkinliklerinin önemli bir gelişim eşliğine yaklaşmakta olduğu alanları ortaya koyacaktır. Uzun vadeli hedef, eksik yetkinlik düzeyleri için bağımsız ve yetkin kıyas ölçütü sonuçları üretebilecek yeni bir test programı geliştirmektir. Böylece kamuoyuna YZ gelişmeleri hakkında yeterli ve güvenilir bilgi sunulması amaçlanmaktadır. OECD, yeni test veya yarışmaların faydalı olabileceği düzeyleri tartışmak ve mevcut değerlendirme yaklaşımlarını belirlemek amacıyla 2026 yılında bu alanda bir başlangıç çalışmayı düzenlemeyi planlamaktadır. Bu sürecin, 2027 yılında ilk değerlendirme geliştirme çalışmalarının başlatılmasına zemin hazırlaması öngörülmektedir.

Yapay Zekâ Yetkinlik Göstergelerinin Rolü

OECD'nin göstergeleri, politika yapıcılara yapay zekâ ilerlemesini teknik olmayan bir kitle tarafından anlaşılabilir terimlerle değerlendirebilecek kanıta dayalı bir araç sunmaktadır. Bu bilgi, özellikle Avrupa Birliği Yapay Zekâ Tüzüğü ve OECD Yapay Zekâ Tavsiye Kararı gibi önemli politika girişimlerinin uygulanması açısından büyük önem arz etmektedir.

Göstergelerin Eğitim, Çalışma Hayatı ve Toplum Üzerindeki Etkileri Öngörmede Kullanımı

Yapay zekâ sistemleri geliştikçe artan yetkinlikleri, bu sistemlere ne zaman ve hangi koşullarda güvenilmesi, devreye alınması ya da sınırlandırılması gerektiğine ilişkin yeni soruları gündeme getirmektedir. Göstergeler, YZ performansının hangi düzeylerinin etik veya güvenlik kaygılarını gündeme getirebileceğini açıklığa kavuşturmaya yardımcı olmaktadır. Bu durumlara, hesap verebilirlik olmaksızın karar alma süreçleri ile savaş veya sağlık hizmetleri gibi yüksek riskli alanlarda özerk kullanım örnek olarak verilebilir.

Bu göstergeler, ahlaki boyutun ötesinde YZ yetkinliklerinin insan emeğinin gereklilikleri ile nasıl örtüştüğünü değerlendirmeye yönelik işlevsel bir araç sunmaktadır. Bu sayede otomasyona daha açık olan mesleklerin belirlenmesi ve daha geniş ekonomik etkilerin öngörülmesi mümkün hâle gelmektedir. Göstergeler, YZ'nin insan çalışanların yerini alıp almayacağına dair doğrudan bir tahminde bulunmasa da, teknik olarak hangi temel görevleri yerine getirebileceğini ortaya koymaktadır. Bu durum, benimsenme sürecine ilişkin ekonomik, düzenleyici veya etik engellerin daha derinlemesine analiz edilmesini teşvik etmektedir.

Eğitim alanında göstergeler hem öğretimde yapay zekâ kullanımına hem de öğrencilerin ihtiyaç duyacağı değişen beceri setlerine ilişkin içgörüler sunmaktadır. Yapay zekâ daha karmaşık görevleri yerine getirebilir hâle geldikçe, eğitim sistemleri hangi yeteneklerin insanlar için temel olmaya devam ettiğini – ister pratik ister bilişsel, isterse içsel nedenlerle – ve gelecek nesillerin giderek daha güçlü YZ sistemleriyle birlikte başarılı olabilmesi için nasıl hazırlanması gerektiğini göz önünde bulundurmalıdır.

Yapay Genel Zekânın Tanımlanması ve Ölçülmesine Yönelik Bir Çerçeve

OECD Yapay Zekâ Yetkinlik Göstergeleri, genellikle insanın bilişsel ve sosyal yeteneklerinin tam kapsamıyla eşleşen yapay zekâ olarak anlaşılan yapay genel zekânın (AGI) tanımlanması ve ölçülmesi için potansiyel bir çerçeve sunmaktadır. Birçok kişi AGI'nin henüz uzak olduğunu düşünse de bazı teknoloji liderleri ve araştırmacılar yakın zamanda ortaya çıkabileceği ve varoluşsal riskler taşıyabileceği konusunda uyarıda bulunmaktadır. Bu bağlamda, açık ve kanıta dayalı izleme araçlarına duyulan ihtiyaç daha da acil hâle gelmektedir.

İnsanüstü AGI'yi tanımlamaya yönelik mevcut girişimler, insan yeteneklerine ilişkin/insan yeteneklerini betimleyen soyut ve uygulamada ölçülmesi güç nitelendirmelere dayanmaktadır. Örneğin Google DeepMind, insanüstü bir AGI sistemini “tüm görevlerde insanların %100'ünden daha iyi performans gösteren” bir sistem olarak tanımlamıştır [Morris et al., 2024[10]]. Buna karşılık OECD'nin Yapay Zekâ Yetkinlik Göstergeleri, YZ gelişmelerini insan yetenek alanlarının tamamı boyunca sistematik biçimde insan performansı ile karşılaştırmaya imkân tanıyan bir çerçeve sunmaktadır. Bu sayede politika yapıcılarının YZ test sonuçlarını doğrudan yorumlamasını gerektirmeden, YZ'nin yetkinlikleri ve sınırlılıklarına ilişkin önemli açıklamalar sunmaktadır.

Bu durum özellikle önemlidir; çünkü YZ yetkinliklerinin farklı alanlarda farklı hızlarda artması muhtemeldir. Büyük dil modeli (LLM) yetkinliklerinin gelişimi, bazı alanlarda [örneğin olgusal bilgi genişliği] görece ileri; diğer alanlarda [örneğin biçimsel akıl yürütme] sınırlı performans nedeniyle “alanlara göre farklılaşan yetkinlik eşiği” olarak tanımlanmıştır [Dell'Acqua et al., 2023[11]]. Bu durumun gelecekteki YZ ilerlemelerinde, varsayımsal AGI sistemleri dâhil olmak üzere, devam etmesi muhtemeldir. İnsan yetenek alanları açısından ilgili alanlardaki güçlü ve zayıf yönlerin izlenebilmesi, YZ ilerlemelerinin sosyal, ekonomik ve politik sonuçlarını haritalamak açısından önem taşımaktadır.

Bu göstergeler tüm ölçeklerde 5. düzey performansın insan becerileri düzeyinde genel yapay zekâ için olası bir kıyas noktası oluşturduğu varsayımından hareketle, politika yapıcılarının YZ ilerlemesini sistematik biçimde izlemesine imkân tanımaktadır. Yaratıcılık, muhakeme ve üstbilgi gibi ileri YZ yetkinliklerini kapsayan bu çerçeve, kamuoyundaki kaygılar ile teknik gerçeklik arasındaki boşluğun daraltılmasına katkı sağlamaktadır. 5. düzey performansın henüz erişilebilir olmadığı durumlarda ise potansiyel “AGI düzeyi” riskler, varsayımlar yerine ampirik kanıtlar temelinde değerlendirilmektedir. Bu yaklaşım, daha sağlam temellere dayanan ve ileriye dönük siyasi önlemlerin alınmasını sağlar.

Paydaşların Göstergelerle Etkileşiminin Desteklenmesi

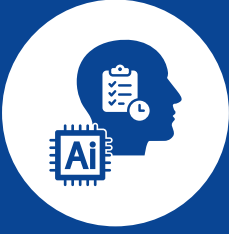
OECD zaman içinde bu göstergelerin hükümetler, akademik araştırmacılar ve sanayi için küresel bir referans noktası hâline gelmesini amaçlamaktadır. Bu hedefe ulaşmak için birkaç temel strateji uygulamaya konulacaktır:

- **Kullanılabilirlik ve yaygınlaştırma:** Göstergelerin geniş bir paydaş kitlesi tarafından kolayca kullanılabilmesi için etkileşimli formatların ve aranabilir veri tabanlarının geliştirilmesi.
- **Uygulama araçları:** Ortaya çıkan YZ yetkinliklerinin çalışma hayatı ve eğitim ortamları üzerindeki etkilerini örnekleyen nicel ve nitel araç ve kaynakların [örneğin mesleki örnekler] sağlanması.
- **Sistematiik izleme:** Politika, iş dünyası ve araştırma topluluklarından çeşitli kullanım örneklerinin toplanarak gerçek dünya uygulamalarına ilişkin erişilebilir bir bilgi tabanı oluşturulması.
- **Sürekli paydaş katılımı:** Uzmanların, uygulayıcıların ve politika yapıcılarının göstergelerin geliştirilme sürecine dâhil edilmesi; göstergelerin geniş ölçekte benimsenmesinin ve iş birliğine dayalı iyileştirilmesinin teşvik edilmesi.

Nihayetinde bu stratejiler, OECD Yapay Zekâ Yetkinlik Göstergelerinin teknik değerlendirmeler ile uygulanabilir politika arasında köprü kurmasını sağlamayı amaçlamaktadır. Yapay zekâ yetkinliklerinin kapsamlı biçimde ölçülmesi sayesinde hükümetler ve paydaşlar, ortaya çıkan YZ teknolojilerinin faydalarını ve risklerini daha doğru biçimde öngörebilecektir. Bu da sürdürülebilir ve yenilik odaklı büyüme doğrultusunda sorumlu ve bilinçli adımlar atılmasına imkân sağlayacaktır [MLCommons, 2025[4]].

Referanslar:

- [8] Autor, D., F. Levy and R. Murnane [2023], "The skill content of recent technological change: Anempirical exploration", The Quarterly Journal of Economics, Vol. 118/4, pp. 1279-1333, <https://academic.oup.com/qje/article/118/4/1279/1925105>.
- [9] Clark Center for Global Markets [2025], Kent A. Clark Center for Global Markets, <https://www.kentclarkcenter.org/us-economic-experts-panel/>.
- [11] Dell'Acqua, F. et al. [2023], "Navigating the Jagged Technological Frontier: Field Experimental Evidence of the Effects of AI on Knowledge Worker Productivity and Quality", Harvard Business School Working Paper, No. 24-013, https://www.hbs.edu/ris/Publication%20Files/24-013_d9b45b68-9e74-42d6-a1c6-c72fb70c7282.pdf.
- [12] Eloundou, T. et al. [2023], "GPTs are GPTs: An Early Look at the Labor Market Impact Potential of Large Language Models", arXiv, <https://arxiv.org/abs/2303.10130>.
- [1] EUR-Lex [2024], Regulation [EU] 2024/1689 of the European Parliament and of the Council of 13 June 2024, EUR-Lex, <https://eur-lex.europa.eu/eli/reg/2024/1689/oj/eng>.
- [6] Handa, K. et al. [2025], "Which Economic Tasks are Performed with AI? Evidence from Millions of Claude Conversations", arXiv, <https://arxiv.org/abs/2503.04761>.
- [3] La Maison Élysée [2025], Statement on Inclusive and Sustainable Artificial Intelligence for People and the Planet, 11 February, La Maison Élysée, Paris, <https://www.elysee.fr/en/emmanuel-macron/2025/02/11/statement-on-inclusive-and-sustainable-artificial-intelligence-for-people-and-the-planet>.
- [5] Maslej, N. et al. [2025], Artificial Intelligence Index Report 2025, <https://arxiv.org/abs/2504.07139>.
- [4] MLCommons [2025], Better AI for Everyone, <https://mlcommons.org/> [accessed on 27 May 2025].
- [10] Morris, M. et al. [2024], Levels of AGI for Operationalizing Progress on the Path to AGI, DeepMind, <https://deepmind.google/research/publications/66938>.
- [7] OECD [2025], AI and the Future of Skills Volume 3: The OECD AI Capability Indicators, OECD Publishing.
- [2] OECD Legal Instruments [2024], Recommendation of the Council on Artificial Intelligence, OECD, Paris, <https://legalinstruments.oecd.org/en/instruments/%20OECD-LEGAL-0449>.



3. OECD YAPAY ZEKÂ YETKİNLİK GÖSTERGELERİ

[AIFS] OECD'nin Eğitim Araştırmaları ve Yenilik Merkezi [CERI] bünyesinde yürütülen Yapay Zekâ ve Geleceğin Becerileri [AIFS] Projesi, yapay zekâ [YZ] ve robotik yetkinlikleri sistematik biçimde ölçmek ve bunları insan becerileriyle karşılaştırmak amacıyla bir çerçeve sunmaktadır. Bu bölümde, OECD'nin Yapay Zekâ Yetkinlik Göstergelerini yer almaktadır. Bu göstergeler hâlihazırda dokuz alanda yapay zekâ yetkinlik düzeylerine ilişkin tanımlamalar ve insan becerileriyle karşılaştırmalar içermektedir: Bunlar: Dil, sosyal etkileşim, problem çözme, yaratıcılık, üstbilis ve eleştirel düşünme, bilgi, öğrenme ve hafıza, görme, manipölasyon; ve robotik zekâ olarak belirlenmiştir. OECD, bu göstergeleri beta sürüm olarak yayınlamaktadır. Bu yaklaşım, yapay zekâ araştırmacıları ve insan psikolojisi uzmanlarıyla sürdürülecek etkileşimin, daha güçlü bir uzlaş zeminini oluşturmak ve yapay zekâ alanındaki hızlı gelişmelere duyarlılığı korumak açısından gerekli olduğuna dair anlayışı yansıtmaktadır.

Dil, birçok bilişsel görevin temelini oluşturan temel bir insan yeteneğidir. Yapay zekâ (YZ) alanında dil ile ilgili yürütülen kapsamlı çalışmalar, bilgisayarların insan dilini anlamasını, yorumlamasını ve üretmesini mümkün kılmaktadır. Bu durum, son dönemde öne çıkan büyük dil modellerinde (LLM'ler) açık biçimde gözlemlenmektedir. Dil, insan etkinliklerinin çok geniş bir alanında merkezi bir rol oynadığından, dil performansının sınırlarını tanımlamak güçleştirmektedir. YZ araştırmacılarının ele aldığı uygulamaların kapsamı doğrultusunda, Dil Ölçeği YZ'nin dil alanındaki yetkinliğini tanımlarken geniş ve kapsayıcı bir yaklaşım benimsemektedir. Bu nedenle, dil ile ilgili görevlerin çeşitliliğini kapsayan birden fazla kritik boyutu içermektedir.

Ölçülmesi önemli olan nedir?

Yazarlar, bir YZ sisteminin genel dil yetkinliğini değerlendirmek için altı kritik boyut belirlemiştir. Bunlardan ilki, kelimelerde kodlanan anlam, dil bilgisi, anlam bilimi (semantik), söylem yapıları ve dil üslubu ile ilgilidir. İkinci ve üçüncü boyutlar, dil kullanımının diğer temel özelliklerine ilişkindir: iletişim kanalları (sözlü ya da yazılı, anlama ya da üretme) ve kapsanan dil sayısı. Kalan üç boyut ise bir dil sisteminin dil ile ilişkili görevler yelpazesine dair yetkinliğini ele almaktadır: bilgiye erişme yeteneği, mevcut bilgi üzerinde akıl yürütme yeteneği ve öğrenme yeteneği. Her bir boyut, ölçeğin genel düzeyi arttıkça ilkel düzeyden gelişmiş YZ dil yetkinliklerine doğru ilerleyen bir gelişim çizgisi göstermektedir.

Mevcut kanıtlar

YZ'nin dil performansını değerlendiren binlerce test bulunmaktadır. İnceleme kapsamında, soru-cevap sistemleri, çeviri ve diyalog sistemleri dâhil olmak üzere alanı yapılandırmak için tipik olarak kullanılan yaklaşık 40 görev türünün yarısı örneklem olarak ele alınmıştır. Her bir alanda performansı ölçmek, ilerlemeyi takip etmek ve teşvik etmek amacıyla araştırmacılar tarafından ortaklaşa geliştirilen bir ya da daha fazla temel kıyaslama (benchmark) veya "ortak görev" incelenmiştir. Bu görevlerdeki mevcut performans çoğu zaman ölçeğin belirli düzeylerinden biri ile örtüşmektedir. Ancak YZ performansı geliştikçe ortak görevlerin zorlaştırılması durumunda, ilgili düzey değişebilmektedir. Ölçek, her bir düzeyle kabaca örtüşen tipik görev türlerini göstermektedir. Teknik rapor, örneklenen görev türlerinin her biri için belirli testlere dair dikkat çekici örnekler sunmaktadır.

Mevcut YZ düzeyi

ChatGPT'de kullanılanlar gibi günümüzün en gelişmiş büyük dil modelleri yaklaşık olarak 3. düzeydedir. LLM'ler geniş kapsamlı bilgiye erişimde başarılıdır; ancak önceden eğitilmiş ve uyarlanabilir olmayan modeller olduklarından akıl yürütme, öğrenme ve dilin incelikli yönlerini ya da iletişimin fiziksel iletişim kanallarını ele alma konusunda sınırlılıklar göstermektedirler. Birden fazla görev için genel amaçlı sistemler olmayı hedefleyen üretken YZ sistemlerinden farklı olarak, üretken olmayan YZ sistemleri daha düşük düzeyde yer almaktadır. Bu sistemlerin belirli bir metin külliyatı üzerinde eğitilmesi ve belirli görevler için makine öğrenmesi teknikleriyle optimize edilmesi gerekmektedir. Örneğin Apple'ın Siri asistanı düşük düzeyli (2. düzey) bir YZ sistemi olarak değerlendirilmektedir. Siri, bilgiye erişim ve akıl yürütme boyutlarında belirgin sınırlılıklar sergilemekte; ChatGPT ile karşılaştırıldığında dil ve çok dillilik yetkinliği daha dar kalmaktadır.

Devam eden zorluklar

YZ'nin dil performansını sınırlayan mevcut zorluklar arasında, yapılandırılmış bilginin sisteme etkili biçimde entegre edilmesindeki güçlükler ve ileri düzey akıl yürütme yeteneklerinin eksikliği yer almaktadır. Bu durum, YZ'nin doğruluğu değerlendirme, mantıksal çıkarımları bütünleştirme ve alana özgül muhakeme yapma yetkinliğini sınırlandırmaktadır. Kıyaslama testlerindeki dilsel ve kültürel önyargılar ise özellikle yeterince temsil edilmeyen diller açısından adil değerlendirmeyi engellemektedir. Ayrıca mevcut sistemler, ölçeklenebilir ve sürekli gelişimi destekleyen öğrenme sistemlerine henüz sahip değildir.

Tablo 3.1: YZ Dil Ölçeği

Performans Düzeyi	Düzy Tanımı
5	Üslup, ton ve mizahı yakalayabilen incelikli dil yetenekleri sergiler; bu yetenekleri gerçek zamanlı genel bilgiye erişim ve eleştirel düşünmeyle bütünleştirerek gerçek yaşam bağlamlarında uygular. Sınırlı veriyle bile yeni bir dili hızlı biçimde öğrenebilir. Herhangi bir dili anlık olarak işleyebilir veya öğrenebilir. Hayat boyu öğrenme yoluyla sürekli evrilir ve öğrenme döngülerini yeniden yapılandırmaya gerek duymadan dinamik biçimde uyum sağlar. Tipik görevler arasında otomatik video betimleme (ör. video altyazıları oluşturma), gerçek zamanlı bilgi, eleştirel düşünme ve çoklu iletişim kanalları girdilerini işleme yeteneği gerektiren yapılandırılmış akıl yürütme görevleri yer alır.
4	İletişimde bağlamı doğru şekilde yorumlar ve karmaşık konuları analiz ederken web ölçekli genel bilgi birikiminden yararlanır. Tüm iletişim biçimlerini kullanarak, kaynakları kısıtlı diller de dahil olmak üzere çok çeşitli dilleri destekler. Sürekli öğrenme yeteneği sayesinde, büyük yapısal değişikliklere gerek kalmadan yeni sürümlere uyum sağlayabilir. Tipik görevler arasında, bağlamı anlayabilen, geniş ölçekli bilgiye erişebilen ve farklı dil girdilerini birlikte işleyebilen gelişmiş diyalog sistemleri yer alır.
3	Farklı metin kaynaklarından edinilen geniş bilgi birikimine dayanarak anlamı doğru ve güvenilir biçimde yorumlar ve üretir. Belirli bir düzeyde problem çözme, mantık yürütme ve sosyal akıl yürütme becerisi gösterir. Çoğu veri türünü etkin bir şekilde işler ve sınırlı miktarda öğrenme verisiyle bile çeşitli dilleri destekleyebilir. Yetkinliklerini yinelenmeli öğrenme, ince ayar (fine-tuning) ve sonradan işleme teknikleriyle geliştirebilir. Tipik görevler arasında, geniş metin bilgisine dayanan ve gelişmiş anlamsal ile sözdizimsel becerileri yansıtan kompozisyon puanlama ve metin sınıflandırma uygulamaları yer alır.
2	Tek bir metin kaynağına dayalı olarak dilbilgisi açısından doğru ifadeler üretir; temel düzeyde problem çözme ve analitik becerilere sahiptir. En iyi kaynaklara sahip dillerde iki farklı iletişim biçimini işleyebilir. Model güncellemeleri ise önemli yapısal değişiklikler ve yeniden eğitim gerektirebilir. Tipik görevler arasında sözdizimsel çözümleme yer alır.
1	Dil yorumlama ve üretiminde anahtar kelime eşleştirmesine dayanır; genel bilgi birikimi ya da akıl yürütme yetkinliği bulunmaz. Yalnızca metin girdisini işler ve tek dillidir. Öğrenme yetkinliği, insan tarafından yazılmış kurallarla sınırlıdır; başlangıçta yapılan programlamanın ötesine geçemez ve kendini uyarlayamaz. Bu düzeyde tipik bir görev, anahtar kelime temelli web aramasıdır.

Sosyal Etkileşim Ölçeği

Brian Scassellati, António M. Fernandes, Ana Teresa Antunes, Rebecca Ramnauth, Nicholas C. Georgiou, Miguel Faria, Haohua Dong, Regina de Brito Duarte, Joana Brito, Henrique Correia da Fonseca, Ana Vilaça Carrasco, Inês Lobo, Rui Prada, Ana Paiva

Sosyal zekâ, bir bireyin dinamik kişilerarası ortamlarda sosyal ipuçlarını algılama, yorumlama ve bunlara uygun biçimde yanıt verme becerisini ifade eder. Yapay zekânın sosyal zekâsını ölçmek ise kendine özgü zorluklar içerir; çünkü insanlar, YZ sistemleri gerçekte sosyal olarak duyarlı olmasalar bile onları sosyal açıdan duyarlı olduğunu algılama eğilimindedir.

Sosyal etkileşim ölçeği, insan sosyal etkileşiminin yalnızca anlık tepkilerden ibaret olmadığını; zaman içinde, diğer bağımsız ve fiziksel varlıklarla gerçekleşen süreğen etkileşimleri kapsadığını kabul eder. Bu nedenle sosyal etkileşimin tüm karmaşıklığını tanımlamak, diğer YZ yetkinlik göstergelerinde yer alan dil, problem çözme ve fiziksel etkileşim boyutlarının birlikte ele alınmasını gerektirir.

Neyin Ölçülmesi Önemlidir?

İnsan düzeyindeki sosyal etkileşimin tüm karmaşıklığını yansıtabilmek için Sosyal Etkileşim Ölçeği, sosyal bağlamın zorluk düzeyini tanımlayan üç temel boyuttan oluşur: Bedenleşme (embodiment), sosyal hafıza ve kimlik. Bir beden olmaksızın, kısa süreli ve belirgin bir kimlikten yoksun etkileşimler mümkün olabilir. Ancak insan düzeyinde tam bir sosyal etkileşim, zaman içinde diğer bedensel bireylerle gerçekleşen süreğen ve fiziksel varlığı içeren etkileşimleri kapsar. Sosyal bağlamı tanımlayan bu üç boyut, dört temel sosyal beceri alanı için kavramsal bir zemin oluşturur: Sosyal iletişim, duyuşsal beceriler, sosyal algı ve sosyal problem çözme.

Mevcut Kanıtlar

Sosyal etkileşimin tüm karmaşıklığını kapsayan karşılaştırmalı ölçütler (benchmark) sınırlıdır. Mevcut değerlendirmelerin incelenmesi, yedi boyutun her biri için farklı düzeylerdeki çalışmaları temsil eden YZ sistemlerine odaklanmıştır. Ölçeğin, farklı YZ sistemlerinin sosyal performansını nasıl tanımlayabileceğini göstermek amacıyla, iyi bilinen çeşitli YZ sistemleri tüm boyutlar açısından analiz edilmiştir.

Mevcut YZ Düzeyi

ChatGPT 4o, OECD'nin Sosyal Etkileşim Ölçeğinde 2. düzeyde yer almaktadır. Güçlü bir sosyal hafıza yetkinliğine sahip olmakla birlikte, bedenlenmiş değildir; belirgin bir kimlik algısına sahip değildir ve sosyal algı becerileri sınırlıdır. Sony'nin AIBO sosyal robotu da 2. düzey bir sosyal YZ sistemi olarak değerlendirilmektedir. Ancak güçlü ve zayıf yönleri, LLM tabanlı sosyal etkileşim sistemlerinden farklıdır. AIBO bedenlenmiş bir sistemdir ve temel düzeyde sosyal algı ile kimlik özelliklerine sahiptir; buna karşın sosyal problem çözme becerileri ChatGPT'ye kıyasla daha sınırlıdır.

Devam eden zorluklar

YZ sistemleri zihin kuramına (theory of mind) sahip değildir; bu nedenle sosyal niyetleri çıkarım yoluyla anlayamaz. Zayıf sosyal algı ve muhakeme ise ipuçlarının yanlış yorumlanmasına ve zamanlaması hatalı müdahalelere yol açmaktadır. Sosyal hafıza sınırlı olduğundan konuşmalar kopuk hâle gelebilmekte, toplumsal normlara uyum sağlama yetkinliğinin zayıf olması ise yazılı olmayan sosyal kuralların öğrenilmesini engellemektedir.

Belirsiz durumlarda esnek yargı yerine katı karar verme süreçleri devreye girmekte, bu da YZ sistemlerinin muğlak sosyal ikilemlerle başa çıkmasını zorlaştırmaktadır. Yetersiz duygu öz-düzenleme yetkinliği nedeniyle sistemler, bir durumun duygusal ağırlığına göre ayarlama yapmak yerine genelleştirilmiş teskin edici ifadeler sunma eğilimindedir.

Tablo 3.2: Yapay Zekâ Sosyal Etkileşim Ölçeği

Performans Düzeyi	Düzye Açıklaması
5	YZ, herhangi bir sosyal ortama kusursuz biçimde uyum sağlar; rolleri doğal biçimde üstlenir ve gerçek zamanlı olarak davranışlarını ayarlayabilir. Sınırsız ve uyarlanabilir bir sosyal hafızaya sahiptir; bağlama duyarlı ve tam uyumlu bir kimlik sergiler. İletişimi derin ve inceliklidir; güçlü bir duygusal anlayış içerir. Sosyal algısı; grup davranışları ve niyetleri hakkında hassas çıkarımlar yapılmasına olanak sağlar. Sosyal problem çözme becerisi ustalık düzeyindedir; bu sayede en karmaşık sosyal durumlarda bile zorlukları öngörebilir ve çözümlerini anında uyarlayabilir. Bu düzeydeki YZ, bir durumu başkasının bakış açısından betimlemek, yeni sosyal normları hızla öğrenmek ya da karmaşık sosyal ilişkileri değerlendirmek gibi görevlerde üstün performans gösterir. Sınırsız uyarlanabilirlik, derin duygusal kavrayış ve kusursuz bağlamsal uyum bu performansın temelini oluşturur.
4	YZ, son derece doğal sosyal davranış sergiler; jestlerini farklı senaryolara göre uyarlar ve yapılandırılmış sosyal hafızayı yönetir. Gruplar içinde net bir rol sürdürür, belirsizlik ve nüanslı iletişimle başa çıkar ve duygusal yoğunluğu ile bunun davranışsal etkilerini anlar. Sosyal algı, güdülerin anlaşılmasına ve grup içindeki rollerin tanınmasına olanak tanır. Sosyal problem çözme yüksek derecede çok yönlü hâle gelir; belirsizlikleri çözmek ve sonuçları öngörmek için sosyal bilgiyi kullanır; bu da karmaşık sosyal ortamlarda akıcı biçimde hareket etmeyi mümkün kılar. Bu düzeydeki YZ, bir garsonun dikkatini çekmek, öğrencinin ilgisizliğini belirlemek veya bir grubu ne zaman ayırmaya karar vermek gibi nüanslı görevleri yönetir. Jestleri uyarlama, duygusal yoğunluğu anlama ve güdülerini yorumlama gibi ileri yetkinlikleri kullanır.
3	YZ, beden dilini yorumlar, grup etkileşimlerini taklit eder ve geçmiş deneyimlere dayanarak yanıtlarını günceller. Tutarlı fakat gelişen bir kişilik sürdürür ve temel duygusal alışverişlere katılabilir. YZ'nin sosyal algısı, sosyal niyetleri çıkarım yapmasına ve davranışsal ipuçlarını yorumlamasına olanak tanır. Sosyal problem çözme daha karmaşık hâle gelir; karmaşık sosyal zorluklara yönelik birden fazla çözümün değerlendirmesine ve uygulamasına izin verir; bu da çeşitli bağlamlarda daha derin farkındalık ve uygulanabilirliği yansıtır. YZ, kavşaklarda sıra alma düzenini koordine etmek veya basit grup dinamiklerini yönetmek gibi görevleri yerine getirebilir; bunu beden dilini yorumlama, niyet çıkarımı yapma ve orta derecede karmaşık sosyal senaryolara dinamik biçimde yanıt verme yetkinliğine dayanarak gerçekleştirir.
2	YZ bu düzeyde sosyal olarak uyum sağlamaya başlar; basit hareketleri birleştirerek duyguları ifade edebilir ve etkileşimlerden öğrenerek gelecekteki karşılaşmalara uyum gösterir. Sınırlı bir sosyal hafıza geliştirir, geçmiş olayları hatırlar ve deneyimlerine dayalı olarak davranışlarını kısmen ayarlar. İletişimi, temel sinyal tanıma becerileriyle gelişir; ton ve bağlam üzerinden duyguları algılayabilir. Sosyal algı, bireyler arasında basit ayrımlar yapmayı içerir; sosyal problem çözme ise geçmiş deneyimleri benzer durumlara uygulayarak temel düzeyde esneklik sağlar. Bu düzeydeki YZ, bireyleri tanıma ve tekrarlayan sorunlarda geçmiş deneyimlerden yararlanma gibi temel görevleri yerine getirebilir. Ancak grup içi etkileşimleri yönetme ya da incelikli duygusal durumları değerlendirme gibi karmaşık eşgüdüm gerektiren görevlerde zorlanır.
1	YZ bu düzeyde basit ve katı sosyal davranışlar sergiler; temel hareketlere ve duygusal ipuçlarına dayanır. Hafızası sabittir ve değişmez; kimliği statiktir. İletişimde önceden tanımlanmış ve senaryolaştırılmış yanıtlar kullanır. Sosyal algısı oldukça sınırlıdır ve yalnızca temel girdiler aracılığıyla birinin varlığını tespit edebilir. Sosyal problem çözme becerisi basit ve önceden belirlenmiş görevlerle sınırlıdır. Bu nedenle yalnızca kısıtlı ve temel düzeyde sosyal etkileşim kurabilir. Bu düzeydeki YZ, başkalarının varlığını algılayabilir ve basit, değişmeyen görevleri yerine getirebilir. Ancak uyum sağlama ve bağlamı anlama yetkinliği sınırlı olduğundan, örneğin bir garsonun dikkatini çekmek ya da konuşmada sırayla söz almayı eşgüdümlemek gibi daha karmaşık sosyal görevleri yerine getiremez.

Problem çözme; analiz yapma, sonuçları öngörme, açıklama üretme ve “şöyle olsaydı ne olurdu?” diye düşünme gibi çok adımlı akıl yürütme süreçlerini içerir. Bu süreçte nitel, nicel ve mantıksal bilgiler birlikte kullanılır. Yapay zekâ ile insanın problem çözme becerilerini karşılaştırmak zordur; çünkü insanlar günlük hayattaki belirsiz durumları yorumlarken örtük bilgiden ve deneyimlerden yararlanılır. Ancak bu tür zorluklar hem insanlara hem de YZ sistemlerine uygulanan problem çözme testlerinde genellikle göz ardı edilmektedir.

Neyin Ölçülmesi Önemlidir?

YZ’de problem çözme zorluğu dört temel boyutla tanımlanır. İlk iki boyut, gerekli çözüm türünü ve değerlendirilen alternatiflerin kapsamını içerir. Bu boyutlar, YZ’nin erken gelişim aşamalarında problem çözme görevlerinin zorluk düzeyini ayırt etmede önemli bir rol oynamıştır.

Ancak günümüzde problem çözmedeki asıl güçlükler daha çok son iki boyutla ilgilidir. Bunlar, mesleki ya da uzman bilgisinin karmaşıklığı ile model kurma ve modeli yorumlama süreçlerinin karmaşıklığıdır. En zorlayıcı durumlar ise genellikle gündelik hayat bağlamında ortaya çıkar: Problemi doğru tanımlamak ve onu çözüme ilerlemeyi mümkün kılacak yapılandırılmış bir biçime dönüştürmek, sağduyu ve sosyal muhakeme gerektirir.

Mevcut Kanıtlar

Hem YZ hem de insan tarafında bu alana ilişkin çeşitli testler bulunmaktadır. Ölçeğin her bir düzeyi için beş ila on arasında YZ kıyas ölçütü belirlenmiş; YZ ölçütlerine dönüştürülebilecek insan değerlendirmeleri tespit edilmiş; ayrıca mevcutsa örnek YZ sistemleri belirtilmiştir.

Mevcut YZ Düzeyi

Düzy 2 YZ sembolik sistemleri - STRIPS/PDDL planlayıcıları, Tatmin Edilebilirlik [Satisfiability] çözücüleri ve model denetleyicileri - lojistik planlama ve model denetimi gibi açıkça tanımlanmış alanlarda insanüstü performans gösterebilmektedir. LLM’ler doğal dilde ifade edilen problemleri ele alabilir; bu yönüyle düzey 3 yetkinliğine sahiptir. Ancak bu yetenek kırılgandır ve çözebildikleri problem türleri bakımından çoğunlukla düzey 1 sınırlarında kalmaktadır. Benzer şekilde, sosyal etkileşim sistemleri temel sosyal muhakeme gerektiren problemleri çözebilir. Bu nedenle iletişim becerileri açısından düzey 3 özellikleri gösterirken, ele alabildikleri problem türleri açısından düzey 1 düzeyinde kalmaktadırlar.

Devam Eden Zorluklar

Başlıca zorluklar arasında nitel muhakemenin otomatikleştirilmesi, sağduyu ve örtük bilgideki boşlukların giderilmesi ile YZ sistemlerinin yeni ya da açık uçlu durumlara uyum sağlama konusundaki esnekliğinin aşılması yer almaktadır. Sosyal zekâ henüz yeterince gelişmemiştir. YZ, ilişkiler, etik meseleler ve incelikli psikolojik etkileşimler hakkında muhakeme yapmakta zorlanmaktadır. YZ matematiksel muhakeme alanında önemli ilerleme kaydetmiştir. Bununla birlikte, nesnelerin uzay içindeki hareketlerine ilişkin, fiziksel sağduyu gerektiren muhakeme hâlâ güçlük oluşturmaktadır. Bu yetkinlikleri ölçen testler de genelleme ve sağlamlık açısından devam eden boşlukları ortaya koymaktadır.

Tablo 3.3: Yapay Zekâ Problem Çözme Ölçeği

Performans Düzeyi	Düzey Açıklaması
5	Bu arzu edilen düzeydeki YZ sistemleri, bilim, hukuk, eğitim ve tıp gibi alanlarda karmaşık, disiplinler arası problemleri çözer; örtük, sosyal ve teknik bilgi boyutlarını bütünlleştirir. Canlı etkileşimler sırasında duyguları ve perspektifleri derinlemesine anlayarak uzun vadeli ilişkiler kurar. Bu sistemler etik zorluklarda yol alır, konuşma ve ikna görevlerinde üstün performans gösterir, çatışmaları çözer, zorbalık gibi nüanslı sorunları tespit eder ve mesleki bilgiyi erişilebilir biçimde etkili şekilde iletir. Bu yetkinliğe ulaşmak mevcut teknolojik sınırların ötesindedir. Bu düzeydeki bir YZ sistemi; sosyal karmaşıklık içeren, birden fazla alandan çözüm yaklaşımları gerektiren ve diğer problemlerle etkileşim hâlinde olan yapılandırılmamış, gerçek dünya problemlerini tanımlayabilir ve çözebilir.
4	Bu düzeydeki YZ sistemlerinin tıp, hukuk ve gazetecilik gibi alanlarda gündelik sağduyu problemlerini ve bazı mesleki problemleri çözmesi beklenmektedir. Sosyal, psikolojik ve fiziksel bilgiyi kullanarak kullanıcılarla ilişki kurar. Bu sistemler geçmiş deneyimlerden öğrenir, gelecekteki performansı ve uygulanabilirliği geliştirir. Daha geniş yapılandırılmamış problem çözmeye doğru bir adımı temsil eder; etkili etkileşimi, alana özgü muhakemeyi ve sürekli öz-gelişimi birleştiren yetkinlikler sunar. Bu düzeydeki YZ sistemleri, karmaşık bir sosyal ortamda etkileşimleri yorumlayabilir, çözülmesi gereken problemleri belirleyebilir ve bu problemleri çözmeye yönelik bir yaklaşım geliştirebilir.
3	Bu düzeydeki YZ sistemleri, gündelik dilde tanımlanan problemleri ele alabilir; doğal ve yapılandırılmamış anlatımları yapılandırılmış modellere dönüştürebilir. Sosyal biliş ve zihin kuramı muhakemesini kullanabilir; insan zihinsel durumlarını simüle edebilir ve niyetleri öngörebilir. Canlı ve cansız varlıklar arasındaki etkileşimleri analiz eder; konuşmalarda duyguları ve niyetleri belirleme ya da etik kararlar verme gibi görevlerde güçlü performans gösterir. Bu sistemler gelişmiş bağlamsal anlayış sergiler ve ahlaki muhakeme, duygusal tanıma ile sosyal etkileşim analizi gibi incelikli görevleri yerine getirebilir. Bu düzeydeki YZ sistemleri, matematik, doğa bilimleri, tıp veya mühendislik gibi alanlarda problemi gündelik terimlerle ifade edilmiş olsa bile çözebilir. Bu tür problemler, söz konusu alanlardaki standartlaştırılmış insan testlerinde yer alan, özellikle sözel problem türlerine benzer. Ayrıca bazı YZ sistemleri, açıkça tanımlanmış sosyal ve etik muhakeme problemlerini de çözebilmektedir.
2	Bu düzeydeki YZ sistemleri alan veya zaman ilişkileri gibi niteliksel muhakemeyi niceliksel analizle bütünlleştirerek karmaşık zorlukları ele alır. Birden fazla niteliksel durumu ve geçişi tasavvur edebilir; sistemlerin zaman içinde nasıl değişebileceğini öngörebilir, bu da düzey 1'e kıyasla daha dinamik ve nüanslı problemleri çözmelerini sağlar. Bu düzeydeki YZ sistemleri; matematik, doğa bilimleri, tıp veya mühendislik gibi alanlarda, problemin geleneksel alan soyutlamaları kullanılarak tanımlandığı durumlarda problem çözebilir.
1	Bu düzeydeki YZ sistemleri yapılandırılmış alanlarda çalışır; bağlayıcı kurallar, matematiksel denklemler veya simülasyonlar gibi kesin ve alana özgü terimler kullanarak problemleri çözer. Veri içindeki tutarsızlıkları, eksik değerleri veya çelişkileri analiz eder; planlama ve zamanlama gibi görevleri yerine getirir. Tıpta, görüşme yanıtları veya test sonuçları gibi yapılandırılmış verilere dayanarak basit sorunları teşhis eder; önceden tanımlanmış parametreler ve dar uygulamalar içinde kalır. Bu düzeydeki YZ sistemleri; matematik, doğa bilimleri, tıp veya mühendislik gibi alanlarda, problemin açıkça tanımlandığı yapılandırılmış problemleri çözebilir. Bu problemler, bu alanlardaki tipik standartlaştırılmış insan testlerindeki sorulara benzer.

Yaratıcılık hem problem çözme hem de sanatsal ifade ile ilişkilendirilen önemli bir insan yeteneği olarak kabul edilir. Yaratıcılığın yalnızca insana özgü olduğu ve YZ'nin sınırlarının ötesinde kaldığı sıkça varsayılmaktadır. Ancak YZ'nin yaratıcı yetkinliğini ampirik olarak değerlendirmek önemlidir. İnsan yaratıcılığı yüzü aşkın farklı biçimde tanımlandığından, bunu insanlar açısından bile tartışmasız şekilde ölçmek zordur. Ayrıca YZ sistemleri genellikle, insan yaratıcılığının temel unsurlarından biri olan özerklikten yoksundur.

Boden [2003[1]] ve Rhodes [1961[2]] tarafından geliştirilen insan yaratıcılığına ilişkin iyi bilinen çerçevelerden yararlanılarak YZ için bir Yaratıcılık Ölçeği oluşturulmuştur. Bununla birlikte, makine yaratıcılığı, insan yaratıcılığını tanımlamakta kullanılan boyutlardan farklı ölçütler gerektirebilir.

Neyin Ölçülmesi Önemlidir?

Önerilen ölçek, alt düzeylerde YZ'nin yaratıcılığını; ürettiği çıktıların değeri, yeniliği, dönüştürücü niteliği ve şaşırtıcılığı üzerinden değerlendirmektedir. Ölçeğin üst düzeylerinde ise odak, YZ sisteminin niyetliliği, öz değerlendirme kapasitesi ve uyum sağlama yeteneğine kaymaktadır.

Mevcut Kanıtlar

YZ yaratıcılığını değerlendirmeye yönelik kapsamlı kıyas ölçütleri bulunmamaktadır. Son dönemde alana özgü çeşitli ölçütler ve kıyas ölçütleri önerilmiştir. Ancak bunlar esas olarak etkililik ve çeşitlilik üzerine odaklanmakta, yalnızca ölçeğin alt düzeylerini ele almaktadır. Ölçeğin geliştirilmesine yönelik ilk çalışmalar, mevcut YZ sistemlerinin üretebildiği yaratıcılık türünü gösteren bir örnekler kümesini belirlemiştir.

Mevcut YZ Düzeyi

Mevcut YZ sistemleri, insan kullanıcılar için değerli [düzey 1] ürünler oluşturabilmektedir. Bu çıktılar aynı zamanda yenilikçi [düzey 2] ve şaşırtıcı [düzey 3] olabilir; bu nitelikler son dönemdeki temel modellerde ve difüzyon modellerinde açıkça görülmektedir. Nitekim yenilik ve şaşırtıcılık, AlphaZero[3] gibi karar verme sistemlerinde de görülmektedir; bu sistemler çok çeşitli problemler için beklenmedik derecede verimli stratejiler üretmektedir.

Devam Eden Zorluklar

Olasılıksal yapıları ve eğitim verilerinin [önceden üretilmiş insan eserlerinden oluşan bir koleksiyon] niteliği nedeniyle, çoğu üretken YZ sistemi — örneğin LLM'ler — gerçekten şaşırtıcı çıktılar üretmekte zorlanmaktadır. İnsan tarafından üretilmiş metinlere dayanması, LLM'lerin insan düşüncesini dönüştüren ya da ileriye taşıyan içerikler üretme yetkinliğini de sınırlamaktadır. Mevcut YZ sistemleri ayrıca niyetlilik, öz değerlendirme ve değişen ortamlara esnek biçimde uyum sağlama gibi daha üst düzey insan yeteneklerini henüz taklit edememektedir.

Çıkarımlar

Yakın zamana kadar yaratıcılık, yalnızca insana özgü bir yetenek olarak görülmüştür. Nitekim YZ yaratıcılığına ilişkin sistematik değerlendirmeler hâlâ sınırlıdır. Bu nedenle yazarlar, politika yapımcıların bu alanda kuramsal çerçeveler ve kıyas ölçütleri geliştirilmesini desteklemesini önermektedir.

Politika yapıcılar ayrıca yaratıcı YZ sistemleri üzerinde insan gözetimini güçlendirmeye odaklanmalıdır. Özellikle insan sanatçılar ya da diğer YZ sistemleri tarafından geliştirilen tarz ve ürünlerden yararlanan çıktılar söz konusu olduğunda, fikri mülkiyetle ilgili olası uyuşmazlıkları ele almaları gerekmektedir.

Tablo 3.4: Yapay Zekâ Yaratıcılık Ölçeği

Performans Düzeyi	Düzye Açıklaması
5	YZ, niyetlilik, özgünlük ve tam etken tabanlı düzyeye ulaşır; dünya çapında insan çıktılarıyla eşdeğer dönüştürücü çıktılar üretir. İçsel hedefleri doğrultusunda neyi-ne zaman üreteceğini özerk biçimde belirler; ayrıca kendini eleştirme, yeniden tasarlama ve kültürel bir bağlam içinde kendini konumlandırma yeteneğine sahiptir. Çıktılar mevcut kombinasyonların ötesine geçer; insanlar ve hatta diğer YZ sistemleri tarafından takdir edilen tamamen yeni estetikler veya paradigmlar ortaya koyar. Görev örnekleri arasında moda pazarına hâkim olan yeni bir moda stili tasarlamak; eleştirmenler tarafından övgüyle karşılanan uluslararası çok satan bir otobiyografi yazmak veya mevcut pazarları sarsan ve yeni sektör standartları belirleyen yenilikçi bir teknoloji tasarlamak yer alabilir.
4	YZ bu düzyede süreç odaklı yaratıcılığı içselleştirir ve ürettiği çıktıları gelişen alanlara uyarlayabilir. Yinelemeli ve keşif temelli arama yoluyla sonuçlarını bağlama uygunluk ve kalite açısından sürekli iyileştirir. Alana ilişkin bilgi ve yaratıcılığa özgü becerileri bir arada sergileyerek, yenilik ile bağlamsal uygunluğu dengeleyen, genel nüfus düzyesindeki yaratıcılığı yansıtır. Örnek görevler arasında özel bir gün için konuşma yazmak yer alabilir. Örneğin bir düşün konuşmasında, yeni evlenen çiftin hayatındaki önemli olayları mizahi, kişisel ancak uygun bir dille seçip bir araya getirebilir; üzücü bir olayın yaşandığı bir ülkedeki genel ruh halini yansıtan bir gazete makalesi yazılabilir; ya da güncel gelişmeleri düşünceli biçimde değerlendiren günlük yazıları yazabilir.
3	YZ bu düzyede, eğitim verilerinin kalıplarını ve yerleşik beklentileri aşarak değerli, yenilikçi ve şaşırtıcı çıktılar üretir. Becerilerini yeni görevlere aktarabilir, farklı alanlardan fikirleri bir araya getirebilir ve geleneksel sınırları zorlayan çözümler geliştirebilir. Böylece yaratıcılığın üç temel unsurunu — değer, yenilik ve şaşırtıcılık — tam anlamıyla karşılar. Örnek görevler arasında, beklenmedik stratejiler geliştirerek video oyunlarında kazanmak; bir siyasi tartışmaya katılıp bir görüşü etkili biçimde savunmak; ya da karmaşık bir olayı aktarmak amacıyla görsel sanat, müzik ve etkileşimli öğeleri birleştiren bir enstalasyon tasarlamak yer alabilir.
2	Bu düzyede YZ, yalnızca taklit etmekle kalmaz; yeni ve işe yarar fikirler üretebilir. Ortaya çıkan sonuçlar, eğitim verilerinin basit bir tekrarıdır ibaret değildir. Sistem, verilen sınırlar içinde farklı olasılıkları dener ve hem yeni hem de anlamlı sonuçlara ulaşır. Bu, tahmin edilmesi kolay olmayan ama işe yarayan buluşlar üretmesi anlamına gelir. Örneğin, bir devlet başkanını klasik Hollandalı ressam tarzında resmetmek; bilim kurgu ile tarihsel romanı birleştiren kısa bir hikâye yazmak; ya da her seferinde farklı şehirler üreten video oyunu seviyeleri tasarlamak bu kapsama girebilir.
1	YZ, basit olmayan görevleri yerine getirmek için insan üretimlerini taklit eder. Ürettiği sonuçlar kabul edilebilir ve işe yarar; insan çalışmalarına benzer, ancak gerçek anlamda özgün değildir. Bu aşama, başkalarının eserlerini yeniden icra eden gruplarda olduğu gibi, yaratıcılığa giden yolda taklit aşamasını temsil eder.

Üstbiliş, bir sistemin kendi muhakeme süreçlerini değerlendirme, güven düzeyini ayarlama ve karmaşık görevlerde hangi bilginin ilgili olduğunu belirleme yetkinliğini ifade eder. Bu yetkinliğin ölçülmesi kendine özgü güçlükler barındırır. Hem insanlar hem de YZ sistemleri açısından, gerçek üstbilişsel süreçler ile sezgisel akıl yürütmeler arasındaki farkı net biçimde ayırt etmek zordur. Mevcut değerlendirme çerçeveleri çoğu zaman görev karmaşıklığını üstbilişsel gereksinimlerle karıştırmakta, bu da ölçümlerin etkililiğini sınırlamaktadır. Yazarlar, yapay zekâ için karşılık gelen bir ölçek geliştirmek amacıyla insanlarda üstbiliş ve eleştirel düşünme üzerine yapılan araştırmalardan yararlanmaktadır.

Neyin Ölçülmesi Önemlidir?

Önerilen model üç temel boyuttan oluşmaktadır: bilişsel bir görevi yerine getirirken stratejiyi değerlendirmek ve ilerlemeyi izlemek için eleştirel düşünme süreçlerine duyulan ihtiyaç; sistemin belirli bir bilgiyi bilme ya da belirli bir problemi çözme olasılığını doğru biçimde değerlendirme yetkinliği ve belirli bir problemi çözmek için hangi bilginin verildiğini ve hangi bilginin gerekli olduğunu ayırt edebilme yeteneği. Bu boyutlar, YZ'nin kendini izleme, belirsizlik karşısında muhakemesini uyarlama ve temel bilgi ile gereksiz bilgiyi ayırt etme yetkinliğini değerlendirmek için bir temel oluşturur. Model, açık akıl yürütme stratejilerinin yanı sıra örtük öz değerlendirme mekanizmalarını da kapsayarak geleneksel YZ kıyas ölçütlerinin önemli bir sınırlılığını gidermeyi amaçlamaktadır.

Mevcut Kanıtlar

Üstbiliş ve Eleştirel Düşünme Ölçeği, görevlerin gerektirdiği bilişsel zorluk düzeylerini nicel bir yaklaşımla tanımlamak üzere geliştirilmiştir. Ölçeğin tanımlarında YZ'ye ya da insana doğrudan atıfta bulunulmamıştır. Modelde yer alan boyutları ele almak amacıyla BIG-bench'ten [Srivastava vd., 2022 [3]] üç kıyas ölçütü kullanılarak bir prototip oluşturulmuştur. Bilginin Gerekliliğinin Değerlendirilmesi [Evaluating Information Essentiality] kıyas ölçütü [Papers with Code, tarih yok [4]], YZ'nin bir soruyu yanıtlamak için gerekli bilgiyi ne ölçüde doğru kabul ettiğini ölçmektedir. Bilinen Bilinmezler [Known Unknowns] kıyas ölçütü, YZ'nin belirli bir bilginin gerçekten bilinebilir olup olmadığını tahmin etme yetkinliğini değerlendirmektedir. Vitamin C Gerçeklik Doğrulama [VitaminC Fact Verification] kıyas ölçütü [Schuster, Fisch ve Barzilay, 2021[5]] ise YZ'nin çelişkili kanıtlar üzerinde muhakeme yapma yetkinliğini ölçmektedir. Bu yaklaşım kapsamında, kıyas ölçütlerindeki her soru için üstbiliş ve eleştirel düşünme gereksinimi tahmin edilmiş ve mevcut büyük dil modelleri performansı bu tahmini düzeyle karşılaştırılmıştır. Ayrıca Dil Modellerinin Bütüncül Değerlendirmesi [Holistic Evaluation of Language Models] [Liang vd., 2022[6]] çerçevesindeki genel kıyas ölçütleri, üstbilişsel performansı genel görev zorluğuyla karşılaştırmak amacıyla kullanılmıştır. Böylece kıyas ölçütü sorularının üstbiliş ve eleştirel düşünmeye ne ölçüde duyarlı olduğu belirlenmiştir.

Mevcut YZ Düzeyi

GPT-3.5 ve GPT-4 gibi en ileri modeller, Üstbiliş ve Eleştirel Düşünme Ölçeğinde genel olarak 2 ile 3. düzey arasında performans göstermektedir. Temel düzeyde güven ayarlaması ve eleştirel düşünme sergilemeler de 4. ve 5. düzeyler için gerekli olan daha ileri üstbilişsel ve eleştirel muhakeme becerilerinde zorlanmaktadır. Ajan tabanlı [agentic] sistemler ise çoğunlukla 3. düzeyin altında performans göstermektedir. Bu durum, YZ'nin kendi muhakeme süreçlerini izleme ve bunları esnek biçimde düzenleme yetkinliğindeki önemli sınırlamalara işaret etmektedir.

Devam Eden Zorluklar

YZ'nin üstbiliş ve eleştirel düşünme yetkinliklerini geliştirmesinin önünde çeşitli engeller bulunmaktadır. Başlıca zorluklardan biri, özellikle aşına olunmayan alanlarda güven düzeyinin doğru biçimde kalibre edilmesidir; bu durum yanıtların aşırı güvenle ya da yetersiz güvenle cevaplanmasına yol açabilmektedir.

Kıyas ölçütlerinin yeterince geliştirilmemiş olması da, üstbilişsel becerilerin doğru biçimde değerlendirilmesini engellemektedir. Ayrıca bilişsel süreçlerin iç içe geçmiş yapısı, üstbilişi diğer akıl yürütme işlevlerinden net biçimde ayırmayı zorlaştırmaktadır.

Tablo 3.5: Yapay Zekâ Üstbiliş ve Eleştirel Düşünme Ölçeği

Performans Düzeyi	Düzye Açıklaması
5	Bu görev, hedefler, kaynaklar ve gerekli beceriler arasında karmaşık denge ve ödünleştirmeleri yönetmeyi gerektiren ileri düzeyde üstbiliş ve eleştirel düşünme içerir. Uzun vadeli görevler başka görevlerle kesişebilir; bu da yetki devri yapma, kendini geliştirme ya da görevi sonlandırma gibi kararlar almayı gerektirebilir. Bu düzeydeki zorlukların başarıyla yönetilebilmesi için doğru öz değerlendirme ve kullanılan yöntemleri ayarlayabilme yetkinliği kritik öneme sahiptir. Örnek: Bir asistanın, bilgisayarda adı “tutulma” (eclipse) ya da benzer bir ifadeye atıfta bulunan bir dosyayı bulup Jason’a e-posta yoluyla göndermesi gerekmektedir. Asistan, dosyayı bulamazsa nasıl yapacağına karar vermeli; kabul edilebilir benzerlik düzeyini belirlemede ve e-posta sistemine erişim sağlayıp dosyayı gönderme yetkisine sahip olup olmadığını değerlendirmelidir.
4	Bu düzeyde görev, düşünme süreçlerinin etkin biçimde düzenlenmesini de içeren yüksek düzeyde üstbiliş ve eleştirel düşünme gerektirir. YZ, alışık olmadığı alanlarda karmaşık ve belirsiz problemlerle karşılaşır; bu da mevcut bilgisini dikkatle değerlendirmeyi ve güven düzeyini doğru ayarlamayı gerektirir. Gerekli bilgiler eksik ya da belirsiz olabilir; bu nedenle bilgiyi etkili biçimde değerlendirmek ve uygulamak için yoğun bir üstbilişsel çaba gerekir. Örnek: Bir asistanın bazı resmî evrak işlemlerini yürütmesi gerekmektedir. Bu süreçte, gerekli tüm eklerin mevcut olup olmadığını belirlemeli; eksik belgeler varsa ilgili kişilerle iletişime geçmesi gerekıp gerekmediğine karar vermelidir.
3	Yüksek düzeyde üstbiliş ve eleştirel düşünme gerektirir hem tanıdık hem de tanıdık olmayan kavramların analiz edilmesini ve bir araya getirilmesini içerir. Kişi, bilgisini eleştirel biçimde değerlendirir, gerekçeli çıkarımlarda bulunur ve karmaşık ya da incelikli bilgileri bütünleştirir. İlgili ayrıntıları belirlemek, ince bağlantılar ve sonuçlar arasında ilişki kurmayı gerektirir; bu da daha güçlü bilişsel esneklik ve stratejik problem çözme becerisi ister. Örnek: Bir robot, daha önce hiç görmediği türden bir kapı koluyla karşılaşır. Bu kolun nasıl kullanılacağını anlamak için bilgi araması yapması ya da farklı denemeler yaparak nasıl çalıştığını keşfetmesi gerekir.
2	Bu düzeydeki görevler, orta düzeyde üstbiliş ve eleştirel düşünme gerektirir; anlama sürecini izlemeyi ve gerektiğinde yaklaşımı uyarlamayı içerir. Konu kısmen tanıdık, ancak ölçülü bir güven düzeyi ve bilinçli tahminler gerektiren belirsizlikler barındırır. Gerekli bilgiler eksiktir; bu nedenle önemli ayrıntıları ayırt edip etkili biçimde kullanmak için üstbilişsel çaba gerekir. Örnek: Bir asistanın, bir müşteri için haftalık alışveriş yapması gerekmektedir. Asistana bir alışveriş listesi, tercih edilen süpermarketler ve sınırlı bir bütçe verilir. Asistan, kalite ile fiyat arasındaki ödünleştirmeleri belirleyip yönetmeli; indirimlere ya da stokta olmayan ürünlere eleştirel biçimde tepki vererek bunları benzer ürünlerle değiştirmelidir. Bu süreçte müşterinin tercihleri hakkındaki bilgisini dikkate alır ve yalnızca tereddüt durumunda müşteriye danışır.
1	Bu düzeydeki görevler, en az düzeyde üstbiliş ve eleştirel düşünme gerektirir; daha çok bilgiyi temel düzeyde yorumlama ya da tanıma üzerine odaklanır. Konu genellikle tanıdık, açık ve nettir ya da oldukça uzmanlaşmıştır; bu da güvenle yanıt vermeyi veya sınırları hızla fark etmeyi mümkün kılar. Gerekli bilgi kolayca belirlenebilir; çoğu ayrıntı verilmiştir ve yalnızca küçük bir eleme ya da basit mantıksal bağlantılar kurmak yeterlidir. Örnek: Bir robotun laktoz intoleransı olan misafirler için Vichyssoise çorbası (Fransız mutfağına ait, genellikle krema içeren pırasa çorbası) hazırlaması ve bunun ne kadar süreceğini kullanıcıya bildirmesi gerekir. Robot, tarifte yer alan kremayı laktozsuz bir alternatifle değiştirip değiştirmeyeceğini belirlemeli, gerekli malzemeleri temin etmeli ve mutfaktaki araçları kullanarak pişirme sürecini tamamlamalıdır.

Bilgi, öğrenme ve hafıza hem insan hem de yapay zekâ için geçerli olan bilişsel sistemler içindeki kritik süreçleri kapsamaktadır. Temel kavramlar birbiriyle ilişkilidir: bilgi yapılandırılmış bilgiyi temsil eder, öğrenme bu bilginin edinimini içerir ve hafıza depolama ile geri çağırmaı sağlar. Bu süreçler insan bilişinin temelini oluşturur ve birçok başka yeteneğin dayanağını teşkil eder. Bu alandaki insan yeteneklerinin tam kapsamını simüle etmek, onlarca yıldır YZ geliştirmesinin temel hedeflerinden biri olmuştur. Ölçek, insanlarda bilgi, öğrenme ve hafıza modellerine dayanmaktadır ve insan yeteneğinin temel yönlerini tanımlamaktadır.

Neyin Ölçülmesi Önemlidir?

En temel düzeyde, bir YZ sisteminin insanlarda görülen türde bilgi, öğrenme ve hafıza yetkinliğine sahip olup olmadığını belirlemek önemlidir. Bilişsel bilim [Cognitive science], kolayca ifade edilip aktarılabilen açık bildirim dayalı bilgiyi [explicit declarative knowledge], farklı becerilerin temelini oluşturan örtük süreçsel bilgiden [implicit procedural knowledge] ayırt etmektedir. İnsanlar bilgiyi doğrudan deneyim yoluyla, başkalarını gözlemleyerek ya da kitaplar ve videolar aracılığıyla sunulan öğretim sayesinde edinebilir. Öğrenme süreci pasif biçimde gerçekleşebileceği gibi belirli bir hedef doğrultusunda aktif olarak da yönlendirilebilir. Deneyimlerin genellenmesi, daha bilinçdışı ve istatistiksel süreçler aracılığıyla ya da daha sembolik ve mantıksal analizlere dayalı süreçler yoluyla gerçekleşebilir. İnsanlar farklı hafıza sistemlerine sahiptir ve hafıza performansı zaman içinde hem güç hem de erişilebilirlik açısından değişkenlik gösterir. İnsan bilgisinin, öğrenmesinin ve hafızasının bu çeşitli yönlerinin YZ sistemlerinde de karşılıkları bulunmaktadır.

Mevcut Kanıtlar

Farklı YZ sistemlerinin performansı, tasarımları analiz edilerek ölçekle ilişkilendirilmektedir. Bu analiz, sistemlerin hangi bilgi, öğrenme ve hafıza işlevlerini mümkün kıldığını; yani hangi tür bilgilerin depolanabildiğini, geri çağırabildiğini ve öğrenilebildiğini ortaya koymayı amaçlamaktadır. Yazarlar ayrıca, nitel açıklamaları desteklemek üzere geliştirilebilecek bazı nicel ölçütleri de tanımlamaktadır. Bu kapsamda; bilgilerin ne kadar verimli biçimde depolanıp geri çağırabildiği; sistemin belirli bir bağlamda potansiyel olarak ilgili bilgileri ne ölçüde başarılı biçimde bulup kullanabildiği; hangi tür bilgileri öğrenebildiği ve bunları ne kadar doğru biçimde genelleyeabildiği; hedeflerini desteklemek için aktif öğrenmeyi ne ölçüde gerçekleştirebildiği ve sahip olduğu bilgiyi kullanarak yerine getirebildiği görevlerin kapsamı değerlendirilmektedir.

Mevcut YZ Düzeyi

Mevcut YZ sistemleri büyük ölçüde 3. düzeyde faaliyet göstermektedir. Bu sistemler genellikle statik biçimde eğitilmiş modellerdir ve istatistiksel genelleme ile geniş veri kümelerine dayanma gibi sınırlılıklara sahiptir. LLM'ler ve ilgili üretken YZ sistemleri bu düzeyi temsil etmektedir. Daha ileri düzeyde ise, belirli alanlarda kendi bilgisini üretebilen ajan tabanlı [agentic] [4. düzey] sistemler geliştirmeye ve bilgi, öğrenme ile hafızanın farklı biçimlerini daha bütünleşik sistem yapıları içinde bir araya getirmeye [5. düzey] yönelik sınırlı çalışmalar yürütülmektedir.

Devam eden zorluklar

Bilgi, öğrenme ve hafıza alanındaki temel zorluklar arasında; bisiklete binmek gibi “nasıl yapılır” becerileri ile tarihleri hatırlamak gibi olgusal bilgi türleri arasında denge kurmak ve otomatik işleyen bilgiyi sistemli muhakeme süreçleriyle birleştirmek yer almaktadır. Bir diğer önemli zorluk ise hızlı ve etkili öğrenebilen sistemler geliştirmek ve bu sistemlerin öğrendiklerini tamamen yeni ve alışılmadık durumlara uyarlamaktır. Ayrıca YZ sistemleri için anlık hatırlama, uzun vadeli depolama, kişisel deneyimlere dayalı bilgi ve genel olgusal bilgi gibi farklı hafıza türlerini sorunsuz biçimde birlikte çalışacak şekilde bütünleştirmek hâlen güçtür.

Tablo 3.6: Yapay Zekâ Bilgi, Öğrenme ve Hafıza Ölçeği

Performans Düzeyi	Düzey Açıklaması
5	Bu düzeyde sistemler, sağlam gerçek zamanlı uyum ve muhakeme için çeşitli bilgi türlerini, öğrenme yöntemlerini ve hafıza sistemlerini bütünleştirir. Halüsinasyon gibi sınırlılıkları ele alırken insan becerileri benzeri bilişsel esneklik ve verimlilik elde ederler. Gelecekteki ilerlemeler, önyargıları ve sınırlılıkları aşarak insan bilişini geride bırakabilir. Bu düzeydeki YZ, bilimsel araştırma yürütmek, kamu politikası kararları almak ve hukuki davaları savunmak gibi açık uçlu bilişsel esneklik gerektiren görevleri yerine getirebilir.
4	Bu düzeyde YZ sistemleri, dünya ve diğer ajan tabanlı sistemlerle etkileşim yoluyla artımlı [kademeli] öğrenme gerçekleştirir. Bilgi boşluklarına odaklanmak ve keşif ile kullanım arasında denge kurmak için üstbilişsel farkındalığı dâhil ederler. Bu paradigmanın açık uçlu ve dinamik alanlara genişletilmesi hâlen bir zorluktur. Bu düzeydeki YZ, ev işleri yapmak, yaşlıları desteklemek veya açık planlı bir endüstriyel ortamda faaliyet göstermek gibi bilinmeyen, belirsiz veya değişen ortamlarda çalışmayı içeren görevleri yerine getirebilir.
3	Bu düzeyde sistemler, bilginin anlam bilimini öğrenmek için dağıtılmış temsilleri kullanır; anlam çıkarır ve yeni durumlara genelleme yapar. Gelişmiş algoritmalar, bağlama duyarlı anlayış için büyük veri kümelerini işler. Önceki düzeylere göre daha uyarlanabilir olmakla birlikte, bu sistemler kapsamlı kaynaklar gerektirir ve gerçek zamanlı öğrenme yeteneklerinden yoksundur. Bu düzeydeki YZ, hikâye yazma, görsel oluşturma, bilgiyi özetleme ve bilgisayar programlama gibi içerik üretmeyi içeren görevleri yerine getirebilir.
2	Bu düzey, katı yapılandırmadan daha az yapılandırılmış bilgi içinde arama yapmaya geçişi temsil eder. Sistem, istatistiksel çıkarım yoluyla arama terimlerini ilgili sonuçlarla eşleştirir ve doğal dil ile diğer yapılandırılmamış veri türlerini esnek biçimde işleyebilir. Ancak veri eksik ya da yetersiz olduğunda doğru genelleme yapmakta zorlanır.
1	Bu temel düzey, yapılandırılmış bilginin kesin ve kurallı hesaplama yöntemleriyle depolanmasını ve geri çağırılmasını içerir. Bilgi; tablolar ve kurallar gibi biçimsel formatlarda temsil edilir ve mantıksal sorgular doğru ve tam geri getirmeyi mümkün kılar. Yapılandırılmış veriler için oldukça etkilidir; ancak örtük ya da açıkça tanımlanmamış bilgilerle çalışmakta zorlanır ve önemli ölçüde teknik geliştirme çabası gerektirir. Bu düzeydeki YZ sistemleri, finansal muhasebe yapmak, istatistiksel hesaplamalar gerçekleştirmek veya takvimleri yönetmek gibi kesin kayıt tutmayı gerektiren görevleri yerine getirebilir.

Görme, insan algısının temel unsurlarından biridir ve birçok bilişsel ile fiziksel görev için önemli bir girdi sağlar. İnsanın görme sistemi, farklı ışık koşullarında ve çeşitli ortamlarda karşılaşılan sahneleri tüm karmaşıklığı ile yorumlayabilir. Hem tanıdık hem de yeni nesne ve sahneleri anlayabilir. Görme ölçeği, bilgisayarların görüntüleri analiz etmesine yönelik alanda pek çok farklı görevin başarıyla ele alındığını göstermektedir. Bununla birlikte, mevcut YZ görme sistemleri genellik ve esneklik açısından insanın görme sistemine göre geride kalmaktadır. Bu alan; nesne tanıma, hareketli sahneleri analiz etme ve otonom yön bulma gibi geniş bir görev yelpazesini kapsar.

Neyin Ölçülmesi Önemlidir?

Belirli görsel analiz uygulamalarının performansını tanımlayabilmek için, sistemlerin hangi tür nesne ve sahnelere ve nasıl yorum yapabildiklerini ve farklı görsel koşullara ne kadar dayanıklı olduklarını belirlemek önemlidir. Ölçekte ayrıca iki ek boyut dikkate alınmaktadır: Gerçekleştirilen görevlerin çeşitliliği ve sistemin geri bildirim yoluyla ne kadar öğrenebildiği olarak yorumlanır. Yazarlar, farklı görsel analiz uygulamalarının performansının temelini oluşturan 32 ayrı görsel yetkinlik alanı belirlemiştir. Bu yetkinlik alanları; nesne tespiti, konum belirleme, özellik tanımlama, hareket analizi, geometrik analiz, örüntü tanıma ve görsel öğrenme gibi becerileri kapsamaktadır.

Mevcut Kanıtlar

Yazarlar iki tür kanıt toplamıştır. İlk olarak, 600'den fazla görsel analiz uygulamasını içeren bir veri tabanından 120 uygulama seçilmiş ve bu uygulamaların performansı ölçeğe göre incelenmiştir. Örneklem, seçilen görevlerde görece güçlü performans gösteren uygulamalara odaklanacak şekilde belirlenmiştir. İkinci olarak, temel 32 görsel yetkinliğinin performans düzeyine ilişkin değerlendirmeler üç kaynaktan elde edilmiştir: yazarların literatür incelemesi, bilgisayar bilimi topluluğuna yönelik bir anket ve ChatGPT 4o'dan alınan yanıtlar. Bu kaynaklardan elde edilen bulgular birlikte değerlendirilmiştir.

Mevcut YZ Düzeyi

120 uygulamadan oluşan örneklem, uygulamaların yarısının düzey 2'de performans gösterdiğini, ancak düzey 1 ve 3'te de önemli sayıda uygulama bulunduğunu göstermiştir. Düzey 4'te yalnızca üç uygulama bulunmuş ve düzey 5'te hiçbir uygulama yer almamıştır. Benzer şekilde, 32 bileşen yetkinliğinin değerlendirilmesi, üçte birinin düzey 2'de performans gösterdiğini; önemli bir kısmının düzey 1 ve 3'te; küçük bir kısmının ise düzey 4'te veya düzey 1'in altında performans gösterdiğini ortaya koymuştur. Bu iki kaynak, güvenilir performans gösteren YZ sistemlerinin bulunduğu en yüksek düzeyin ölçeklerde düzey 3 olduğuna dair örtüşen kanıt sunmaktadır.

Devam Eden Zorluklar

Bilgisayarların görüntüleri analiz etmesi alanındaki temel zorluklardan biri, değişken ve karmaşık gerçek dünya koşullarına uyum sağlamaktır. Mevcut sistemler gerçek zamanlı düşünme ve duruma göre uyarlanma konusunda sınırlıdır. En yüksek performansa ulaşabilmek için görme sistemlerinin yalnızca sabit modellere dayanması yeterli değildir; sürekli öğrenen ve kendini geliştiren sistemler olmaları gerekir.

Tablo 3.7: Yapay Zekâ Görme Ölçeği

Performans Düzeyi	Düzey Açıklaması
5	Bu en üst düzeyde, sistemler görevleri insan gözüyle aynı başarı düzeyinde yerine getirir. Bir insanın karşılaşılabileceği tüm durumlara uyum sağlayabilirler. Buna farklı ışık koşulları, bakış açıları, şekiller, görünüm, konumlar ve sahne değişiklikleri dâhildir hem alışılmış hem de yeni durumlarla başa çıkabilirler. Bu sistemler, kendi performanslarını geri bildirim olarak geliştirir. Nesneleri bulabilir, sınırlarını çizebilir hem genel hem de ayrıntılı şekilde tanımlayabilir, bir nesnenin nasıl tutulacağını veya hareket ettirileceğini tahmin edebilir ve nesnelerin birbiriyle nasıl etkileştiğini anlayabilirler. Ayrıca çevredeki değişikliklere uyum sağlayarak yeni özellikleri, nesneleri ve davranışları öğrenebilirler. Bu düzeye örnek olarak, otonom araçların yoğun ve sürekli değişen trafikte nesneleri tanınması, hareketleri takip etmesi ve çevreyi anlık olarak anlaması verilebilir.
4	Düzey 4 sistemler, çok farklı veri türleriyle çalışabilir. Mikroskop görüntüleri, kırmızı-yeşil-mavi (RGB) görüntüler, insanlara ait görüntüler, mekanik parçalar ve doğal sahneler gibi geniş bir içerik yelpazesine uygulanabilirler. Hedef nesnelerdeki ışık, şekil ve görünüm değişikliklerine uyum sağlayabilir; birbirine çok benzeyen nesneler arasındaki küçük farkları ayırt edebilirler. Bu sistemler, kendi yaptıkları değerlendirmelerden ya da dışarıdan gelen geri bildirimlerden öğrenerek performanslarını geliştirebilirler. İnsanların yapabildiği her görevi yerine getiremeseler de birçok farklı görevi başarıyla yapabilirler. Yaptıkları görevlerde insan düzeyine yakın sonuçlar üretirler. Ayrıca bir görevin çıktısını başka bir görevin girdisi olarak kullanarak farklı görevleri bir arada yürütebilirler. Örneğin bir mutfak asistanı robotun; şekilleri tanıması, nesnelerin yerini belirlemesi, nasıl tutulacağını tespit etmesi, hareketleri takip etmesi ve ortaya çıkan sonucun kalitesini değerlendirmesi gerekebilir. Bu düzeydeki sistemler genellikle dinamik ortamlarda çalışır. Örneğin robotların mutfakta çeşitli görevleri yerine getirmesi, montaj hatlarını izlemesi ya da üretimde karmaşık kalite kontrol işlemleri yapması bu kapsamdadır.
3	Bu düzeydeki sistemler; mikroskop görüntüleri, RGB görüntüler ve doğal sahneler gibi farklı veri türleriyle çalışabilir. Işık koşullarındaki ve hedef nesnenin görünümündeki bazı değişikliklere uyum sağlayabilirler. Birden fazla alt görevi yerine getirebilir ve daha önce bilinen veri ve durum değişiklikleriyle başa çıkabilirler. Bazı alanlarda insan performansına yaklaşırsalar da insan yetkinliğine tam olarak ulaşamazlar. Örneğin gelişmiş bir otonom araç görme sistemi; araçları, engelleri ve yayaları tespit edebilir, hareketlerini takip edebilir ve aynı zamanda rota, yol durumu, hava koşulları ve aracın hareketi gibi bilgileri bir araya getirebilir. Ancak bu sistemler, kendi uzmanlık alanlarının dışındaki görevlerde zorlanabilir. Bu düzeyde tipik görevler arasında otonom araçların yön bulması, yüz tanıma sistemleri ve robotlar için çevre haritalama yer alır.
2	Düzey 2 sistemler; ışık koşullarındaki değişikliklere, sensörün sahneye göre konumunun farklılaşmasına ve gözlenen alandaki bazı değişimlere uyum sağlayabilir. Düzey 1'e göre daha esneklerdir. Hareketlerin hız ve zamanlamasındaki değişikliklerle ve sahnedeki nesnelerdeki farklılıklarla başa çıkabilirler. Örneğin şerit takibi ve otonom sürüşte engel tespiti ya da güvenlik sistemlerinde yüz tespiti ve tanıma gibi, belirli ölçüde değişkenlik içeren ortamlarda özel görevleri başarıyla yerine getirebilirler. Ancak hâlâ belirli alanlara ve görevlere odaklıdır. En iyi performansı gösterebilmeleri için dikkatle tasarlanmış ve kontrol edilmiş koşullara ihtiyaç duyarlar. Bu düzeyde tipik görevler arasında yüz tespiti, kontrollü sürüş ortamlarında engelden kaçınma ve üretimde belirli alanlara yönelik görsel denetimler yer alır.
1	Düzey 1'de sistemler, değişkenliğin çok az olduğu ve sıkı biçimde kontrol edilen ortamlarda çalışır. Genellikle yalnızca tek bir görevi yerine getirirler. Bu görevi neredeyse hatasız yapabilirler, ancak sadece belirli ve sınırlı koşullar altında. Üretimde yapılan görsel denetimler ya da posta kodu tanıma gibi birçok endüstriyel uygulama bu gruba girer. Sistem, sabit bir sahne ve belirli nesnelerle iyi çalışır. Ancak ortamda ya da nesnelerde bir değişiklik olduğunda zorlanır. Bu sistemler esnek değildir ve düzenli, değişmeyen koşullara büyük ölçüde bağlıdır. Bu düzeyde tipik görevler arasında sabit ortamlarda basit nesne tanıma, barkod okuma ve düzenli üretim ortamlarında kalite kontrol yer alır.

Manipülasyon, insanın temel fiziksel becerilerinden biridir. Çevremizdeki nesnelerle etkileşim kurma yeteneğini ifade eder. Bu süreç; fiziksel hareketi, dokunma ve görme gibi duyularla geri bildirim almayı ve hareketleri planlayıp gerektiğinde düzeltmeyi kapsar. Robotik manipülasyon ise robotların nesnelerle etkileşim kurmasını sağlar. Bu, basit al-ve-yerleştir işlemlerinden daha karmaşık görevlere kadar uzanır. Örneğin çamaşır katlamak gibi şekli değişebilen nesnelerle çalışmak ya da dağınık bir ortamda parçaları birleştirerek montaj yapmak bu kapsamdadır.

Neyin Ölçülmesi Önemlidir?

Bir manipülasyon görevinin ne kadar zor olduğu birçok faktöre bağlıdır. Görev; kavrama, sabitleme ya da nesneyi el içinde yönlendirme gibi farklı temel hareketleri içerebilir. Ayrıca nesnenin özellikleri (örneğin sert ya da yumuşak olması), görevin yapıldığı ortam ve bazı kısıtlamalar da önemlidir. Bu kısıtlamalar; zaman sınırı ya da diğer nesnelere ne kadar yakın çalışabileceği gibi şartlar olabilir. Bir manipülasyon görevinin zorluk düzeyini belirlemek için tüm bu özellikler dikkate almak gerekir. Ayrıca bir robotun manipülasyon yetkinliği, bu farklı durumlara ne kadar uyum sağlayabildiğine bağlıdır. Yani robotun yapabildiği temel hareketlerin çeşitliliği, farklı nesne türleriyle çalışabilmesi, değişik ortam koşullarına uyum sağlaması ve görev kısıtlamalarına uyum sağlayabilmesi bu yetkinliği belirler.

Mevcut Kanıtlar

Fiziksel manipülasyon alanında kullanılan kıyas ölçütlerinin sayısı oldukça azdır. Bunlardan daha da azı, farklı sistemlerin performansını zaman içinde karşılaştırmaya imkân veren lider tabloları sunmaktadır. Yazar, ölçeğin alt düzeylerine giren görevleri kapsayan 11 farklı kıyas görevi belirlemiştir. Ancak Düzey 4 veya 5'teki manipülasyon becerilerini kapsamlı şekilde değerlendiren bir kıyas ölçütü bulunmamaktadır.

Mevcut YZ Düzeyi

Manipülasyon alanındaki en gelişmiş robotik sistemler şu anda Düzey 2'dedir. Örneğin üretimde kullanılan robotik kollar, kontrollü ortamlarda açıkça tanımlanmış görevleri başarıyla yerine getirebilir. Ancak ortam daha değişken ve öngörülemez olduğunda zorlanırlar. Robotlar birçok al-ve-yerleştir işleminde oldukça başarılıdır. Fakat düzenli olmayan ortamlarda, kırılğan ya da şekli düzensiz nesnelerle çalışırken zorlanabilir. Nesnelerin türü ya da konumu çok değiştiğinde performansları düşebilir.

Devam Eden Zorluklar

Robotik manipülasyonda ilerlemenin önündeki en büyük engellerden biri, robotların el becerisi ve uyum sağlama yetkinliğinin sınırlı olmasıdır. Özellikle çok farklı türde nesnelerle ya da öngörülemeyen çevre koşullarıyla karşılaştıklarında zorlanırlar. Ayrıca robotlar, gerçek zamanlı karar verme ve anında öğrenme konusunda da sıkıntı yaşayabilir. Bu durum, yeni ve beklenmedik durumlara hızlıca uyum sağlamalarını zorlaştırır. Hem yüksek el becerisi hem de gelişmiş düşünme ve planlama gerektiren karmaşık görevlerde ise mevcut robotlar genellikle etkili ve güvenilir performans göstermek konusunda yetersiz kalmaktadır.

Tablo 3.8: Yapay Zekâ Manipülasyon Ölçeği

Performans Düzeyi	Düzyet Açıklaması
5	Bu zirve düzeydeki robotlar, manipülasyon görevlerinde insan yetenekleriyle eşdeğer bir düzyete ulaşır ve son derece karmaşık ortamlar dâhil olmak üzere her türlü ortamda verimli biçimde çalışabilir. Yansıtıcı yüzeyler, kaygan dokular ve esnek malzemeler dâhil olmak üzere farklı şekil, boyut, malzeme ve dinamiklere sahip nesneleri olağanüstü uyarlanabilirlik becerisiyle ele alırlar. Nesneleri el içinde hızlı biçimde yeniden konumlandırır, karmaşık yönelimlere yerleştirebilir ve dinamik değişikliklere yanıt verebilirler. Sıkı zaman kısıtları altında, yetkin bir insanla eşdeğer hassasiyet, verimlilik, dayanıklılık ve uyum sağlama kapasitesiyle görevleri yerine getirebilirler. Kendi sınırlılıklarını anlayarak ve yeteneklerinin ötesindeki görevleri reddederek insanlarla sorunsuz biçimde iş birliği yaparlar. Tipik görevler arasında bir kişiye giyinmesine yardımcı olmak veya arama ve kurtarma operasyonları yer alır.
4	Bu düzeydeki bir robot, önemli ölçüde dağınıklık ve engel içeren ortamlarda çalışabilir. Hedef nesneleri hedef olmayan nesnelerden hızlı biçimde ayırt edebilirler. Hareketli parçaları olanlar dâhil hem katı hem de esnek nesneleri işleyebilir. Dar alanlarda ya da engelli konumlarda daha yüksek hassasiyetle hareket edebilir. Belirli bir yönelim ya da yerleştirme gerektiren görevleri daha doğru biçimde yerine getirebilir. Orta düzey uyum gerektiren kuvvete dayalı işlemleri gerçekleştirebilir. Değişen nesne özelliklerine ve çevre koşullarına belirli ölçüde uyum sağlayabilir; ancak bazı durumlarda insan onayı gerekebilir. Görevleri sınırlı sürede tamamlayabilir; ancak insan verimliliğiyle eşleşmez. Bu düzeyde tipik görevler arasında bulaşık makinesi boşaltma, kuvvete dayalı yüzey işlemleri ve dağınık ya da dinamik ortamlarda nesne montajı yer alır.
3	Bu düzeydeki bir robot, orta derecede dağınık ortamlara uyum sağlayabilir. Dikkat dağıtıcı unsurlar arasından hedef nesneyi seçip manipüle edebilir. Yansıtıcı ya da düşük sürtünmeli yüzeyler gibi daha önce zorlayıcı olan daha geniş bir nesne şekli ve malzeme çeşitliliğiyle çalışabilir. Nesnelerin yönünü değiştirebilir veya orta derecede zor konumlara yerleştirebilir; ancak robotun tuttuğu bir nesneyi bırakmadan elinin içinde hızlı biçimde yeniden konumlandırması hâlâ zor olabilir. Talimatlara dayalı kuvvete bağlı işlemleri gerçekleştirebilir; ancak büyük ölçüde uyarlama yapamaz. Kısıtlı zaman diliminde çalışabilir; ancak çok kısa sürede yüksek verimlilik gerektiren durumlarda ve kontrollü koşullar dışında performansını sürdürmeyebilir. Bu düzeyde tipik görevler arasında düzensiz nesneleri yeniden yönlendirme, bir yemek için masa hazırlama ve kuvvete dayalı manipülasyon gerektiren hassas malzemeleri işleme yer alır.
2	Bu düzeydeki bir robot, düşük ile orta derecede dağınıklık içeren ortamlarda çalışabilir. Belirli bir alan içinde rastgele yerleştirilmiş nesneleri işleyebilir. Farklı nesne şekilleriyle ve bazı esnek malzemelerle çalışabilir; ancak elastik ya da kaygan malzemeler hâlâ zorluk yaratır. Küçük engellerin etrafından dolaşabilir; fakat bir nesneyi belirli bir açığa döndürmek ya da dar bir boşluğa yerleştirmek gibi daha karmaşık manipülasyonlar sorun olmaya devam eder. Kontrollü koşullarda görevleri yerine getirebilir; ancak hızlı tepki gerektiren ya da beklenmedik değişiklikler içeren durumlarda zorlanabilir. Bu düzeyde tipik görevler arasında oyuncak blokları bir masadan alıp bir saklama kutusuna yerleştirme ve kontrollü fabrika ortamlarında malzeme taşıma yer alır.
1	Bu düzeydeki bir robot, iyi düzenlenmiş ortamlarda basit al-ve-yerleştir görevleriyle sınırlıdır. Kolay kavranabilen, temel şekillere sahip katı nesneleri manipüle edebilir. Tek tip malzemeler algılama ve kavrama açısından fazla zorluk oluşturmaz. Dış engellerin bulunmadığı ortamlarda en iyi performansı gösterir. Önceden tanımlanmış yolları izler ve uyum sağlama yetkinliği oldukça sınırlıdır. Beklenen koşullardan bir sapma olduğunda genellikle görev başarısız olur. Geniş hata payıyla çalışır; hassas konumlandırma gerektiren işlemler yerine daha kaba ve temel hareketlere odaklanır. Örneğin bir depoda, önceden öğrenilmiş konumlardan mısır gevreği kutularını alıp kasalara yerleştirme görevi bu düzyete örnek verilebilir.

İnsanlar çevrelerinde hareket edebilir ve daha üst düzey hedefler tarafından yönlendirilen çok çeşitli görevleri kendi başlarına yerine getirebilir. Doğal bir ortamda özerk bir özne olarak hareket edebilmek, insanın tüm yeteneklerinin birlikte ve uyum içinde çalışmasını gerektirir. Buna algı ve fiziksel hareketin yanı sıra dil kullanımı, sosyal etkileşim ve farklı türde problem çözme becerileri de dâhildir. Bütünleşik robotik zekâ, bu insan özerkliğini taklit etmeyi amaçlar. Bu yaklaşım; otonom yön bulma, insan-robot etkileşimi ve gerçek zamanlı karar verme gibi görevlerde duyuşal, motor ve bilişsel sistemlerin eşgüdüm içinde çalışmasını gerektirir.

Neyin Ölçülmesi Önemlidir?

Robotik zekâ ölçeği altı boyuttan oluşur. Bu boyutların dördü görevin kendisiyle ilgilidir. Bunlar; görevin ne kadar karmaşık olduğu, görevin ne kadar soyut tanımlandığı ve bu nedenle ne kadar problem çözme gerektirdiği, görev sırasında gerekli olan sosyal etkileşim düzeyi ve görevin nasıl yapılacağını sınırlayan etik unsurlardır.

Diğer iki boyut ise görevin gerçekleştiği ortamla ilgilidir. Bunlar, ortamın karmaşıklığı ve ortamda bulunan belirsizlik düzeyidir. Ayrıca robotun ortamla etkileşimi sırasında ortaya çıkan belirsizlik de bu kapsamda değerlendirilir.

Mevcut Kanıtlar

Mevcut yapay zekâ ve robotik sistemlerde bütünleşik zekâ düzeyini değerlendirmek için kullanılan kıyas ölçütlerinin sayısı azdır. Bununla birlikte, karmaşık üretim, uzay keşfi ve insanlara yönelik hizmetler gibi farklı alanlarda çeşitli yarışmalar ve uygulama temelli çalışmalar yürütülmektedir.

Bu alandaki değerlendirmeler; literatür incelemesi, atölye çalışmaları ve araştırmacılarla yapılan görüşmeler bir araya getirilerek oluşturulmuştur. Amaç, güncel araştırmacıların ortak görüşünü yansıtan bir çerçeve geliştirmektir.

Mevcut YZ Düzeyi

Günümüzdeki en gelişmiş sistemler, örneğin otonom teslimat robotları ve endüstriyel otomasyon sistemleri, ölçeğe göre yaklaşık Düzey 2'dedir. Bu sistemler düzenli ve önceden planlanmış ortamlarda iyi çalışır. Ancak ortam karmaşılaştığında zorlanırlar. Beklenmedik durumlarda doğru karar vermekte güçlük çekerler. Yaratıcılık ya da sosyal etkileşim gerektiren görevlerde başarılı değildirlir. Örneğin bir robot, önceden haritası çıkarılmış bir yerde rahatça hareket edebilir. Fakat bir insanla iletişim kurması ya da ortam aniden değiştiğinde buna uyum sağlaması gerektiğinde sorun yaşayabilir.

Devam Eden Zorluklar

Robotik zekâ alanındaki temel sorunlar; uyum sağlama, problem çözme ve etik karar verme konularındaki sınırlılıklardır. Robotlar belirli görevler için programlanabilir. Ancak değişen koşullara uyum sağlama, insanlarla iş birliği yapma ve gerçek zamanlı etik karar verme becerileri henüz yeterince gelişmemiştir. Gerçek dünya ortamları çoğu zaman belirsizlik içerir. Bu durum performansı düşürebilir. Robotlar, eksik ya da çelişkili bilgilerle karşılaştıklarında doğru karar vermekte zorlanabilir.

Çıkarımlar

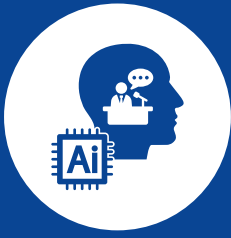
Bütünleşik robotik sistemler etik açıdan bazı önemli soruları gündeme getirir. Özellikle insanlarla etkileşim içeren görevlerde ya da sağlık hizmetleri ve otonom sürüş gibi kritik alanlarda güvenlik, adalet ve sorumluluk konuları büyük önem taşır. Politika yapıcılarının, robotların tasarımı ve kullanımı sırasında şeffaflığı, adaleti ve güvenliği garanti altına alacak standartlar ve düzenlemeler geliştirmesi gerekir. Bu teknolojilerin ilerlemesi için uyum sağlayabilen, etik ilkelere bağlı ve toplumsal sorumluluk gözetan robotik zekâ üzerine yapılacak araştırmalara yatırım yapılması da büyük önem taşır.

Tablo 3.9: Yapay Zekâ Robotik Zekâ Ölçeği

Performans Düzeyi	Düzyet Açıklaması
5	Bu en üst düzeydeki bir robot, düzenli olmayan ortamlarda birden fazla karmaşık görevi yerine getirebilir. Hedef belirleme konusunda yaratıcıdır. Net tanımlanmamış görevleri açıklığa kavuşturup yeniden düzenleyebilir. Değişen koşullara uyum sağlayabilir. Deneyimlerinden öğrenebilir. Farklı görev ve ortamlara genelleme yapabilir. Gelişmiş akıl yürütme ve sağduyuya dayalı düşünme becerileri gösterir. Sosyal etkileşimde yüksek düzeyde yetkindir. Kendi sınırlarının farkındadır. Yasal ya da ahlaki kurallarla çelişen görevleri reddedebilir ve etik kararlar verebilir. Bu düzeyde tipik görevler arasında engelli bireyler için ev içi destek sağlayan robotlar, etik karar verme gerektiren uygulamalar ve farklı ve dinamik ortamlarda yüksek performanslı otonom sürüş yer alır.
4	Bu düzeydeki bir robot, farklı karmaşıklıkta birden fazla görevi yerine getirebilir. Değişen koşullara uyum sağlar ve ortam değiştikçe davranışını buna göre ayarlar. Kendi sınırlarının farkındadır ve geri bildirimden yararlanarak performansını geliştirir. Bu kategorideki görevler genellikle uzun vadeli ve karmaşık hedefler içerir. Görevler, içinde bulunulan bağlama bağlı olarak şekillenir. Robot belirsizlikle başa çıkabilir ve belirsiz ortamlarda karar verebilir. Ancak ürettiği çözümler her zaman insanın bulduğu çözümler kadar etkili ya da verimli olmayabilir. Bu düzeyde tipik görevler arasında, malzemelerin durumuna göre seçim yapan yemek pişirme robotları, engeller arasında yol alan otonom tekerlekli sandalyeler ve havaalanı çevresinde otonom hava araçlarının yön bulması yer alır.
3	Düzyet 3'teki bir robot, orta vadeli ve birden fazla adımdan oluşan görevleri yerine getirebilir. Bu görevler belirli ölçüde esneklik gerektirir. Robot, orta düzeyde değişkenlik içeren ortamlarda çalışabilir ve tam olarak tanımlanmamış birkaç alt görevi bir arada yürütebilir. İnsanlarla iş birliği yapabilir. Orta düzeyde belirsizlikle başa çıkabilir. Işık, hava durumu ya da yeni nesne türleri gibi değişikliklere uyum sağlayabilir. Birden fazla çözümü olan görevleri yerine getirebilir; ancak çok daha öngörülemez ve hızlı değişen ortamlarda zorlanabilir. Bu düzeyde tipik görevler arasında hem taşıma hem temizlik yapan hastane robotları, mobilya kurulumuna yardımcı olan robotlar ve öğrenilmiş tercihlere göre çekim yapan otonom robot kameramanlar yer alır.
2	Bu kategorideki bir robot, yarı yapılandırılmış ortamlarda önceden tanımlanmış görevleri yerine getirir. Ortamdaki bazı değişikliklerle başa çıkabilir. Örneğin nesnelerin yerinin değişmesi ya da alan düzeninin farklılaşması gibi durumlara uyum sağlayabilir. Belirsizlik düzeyi düşüktür ya da orta seviyededir. Görevlerin başarı ölçütleri nettir. Robot genellikle insan müdahalesi olmadan çalışır. Basit ve birden fazla işlev içeren görevleri yapabilir; ancak daha karmaşık ya da beklenmedik değişiklikler karşısında zorlanır. Bu düzeyde tipik görevler arasında hastanelerde malzeme taşıyan robotlar, fabrikalarda kullanılan malzeme taşıma robotları ve meyve toplama yapan tarım robotları yer alır.
1	Düzyet 1'deki bir robot, çok düzenli ve kontrollü ortamlarda basit ve tekrarlayan görevleri yerine getirir. Ortam sabittir ve önceden tamamen bilinir. Değişkenlik yoktur ya da çok azdır. Robot, önceden belirlenmiş talimatları izler. Uyum sağlayan kararlar veremez ve beklenmedik durumlarla başa çıkamaz. İnsanlarla etkileşim kurmaz. Ortamda küçük bir değişiklik olduğunda bile zorlanabilir. Bu düzeyde tipik görevler arasında fabrikalarda temel otomatik montaj işlemleri, robot süpürgeler ve lojistikte nesne ayırma sistemleri yer alır.

Referanslar:

- [1] Boden, M. [2003], *The Creative Mind: Myths and Mechanisms*, Routledge, Milton Park, Abingdon-on-Thames.
- [6] Liang, P. et al. [2022], "Holistic evaluation of language models", arXiv, Vol. 2211.09110, <https://arxiv.org/abs/2211.09110>.
- [4] Papers with Code [n.d.], "Evaluating Information essentiality on BIG-bench", BIG-bench Benchmark, [database], <https://paperswithcode.com/sota/evaluating-information-essentiality-on-big> [accessed on 15 May 2025].
- [2] Rhodes, M. [1961], "An analysis of creativity", *The Phi Delta Kappan*, Vol. 42/7, pp. 305-310, <https://www.jstor.org/stable/i20342591>.
- [5] Schuster, T., A. Fisch and R. Barzilay [2021], "Get your vitamin C! Robust fact verification with contrastive evidence", arXiv, Vol. 2103.08541, <https://doi.org/10.48550/arXiv.2103.08541>.
- [3] Srivastava, A. et al. [2022], "Beyond the imitation game: Quantifying and extrapolating the capabilities of language models", arXiv, Vol. 2206.04615, <https://doi.org/10.48550/arXiv.2206.04615>.



4. YZ YETKİNLİK GÖSTERGELERİNİN POLİTİKA AMAÇLI KULLANIM ALANLARI

OECD Eğitim Araştırma ve Yenilik Merkezi (CERI) bünyesinde yürütülen Yapay Zekâ ve Becerilerin Geleceği (AIFS) Projesi, yapay zekâ (YZ) ve robotik sistemlerin neler yapabildiğini düzenli ve ölçülebilir bir şekilde değerlendirmek ve bunları insan becerileriyle karşılaştırmak için bir çerçeve sunar. Bu bölümde AIFS, politika yapıcılarının OECD'nin YZ Yetkinlik Göstergelerini kullanarak YZ'deki gelişmelerin farklı alanlardaki etkilerini nasıl analiz edebileceğini ele almaktadır. İlk kısımda, bu göstergelerden yararlanarak YZ'nin mesleklerdeki talebi nasıl değiştirebileceğini tahmin etmeye yönelik bir yaklaşım özetlenmektedir. İkinci kısımda ise OECD, eğitim politikası yapıcılarının bu göstergeleri öğretmen rollerindeki değişimi değerlendirmek ve mesleki taleplerdeki dönüşümü dikkate alarak öğrenme hedeflerinin nasıl güncellenmesi gerektiğini tartışmak için nasıl kullanabileceklerini açıklamaktadır.

Yapay zekâ [YZ] çok hızlı ilerliyor. Ancak her yeni gelişme, toplumda ve ekonomide büyük bir değişim yaratmaz. OECD'nin YZ Yetkinlik Göstergeleri, YZ'nin hangi alanlarda gerçekten önemli ve dönüştürücü etkiler yaratabileceğini anlamaya yardımcı olur. Bu göstergeler, tek tek YZ uygulamalarını değerlendirmek için fazla genel olabilir. Ancak meslek gruplarının ve öğrenme süreçlerinin gelecekte nasıl değişebileceğini görmek için uygun bir araçtır.

Göstergeler yeni olduğu için OECD henüz bunları düzenli ve kapsamlı biçimde kullanmamıştır. Bu bölümde, YZ'nin iş hayatında gerekli olan insan becerilerine ne kadar yaklaştığını göstergeler aracılığıyla nasıl değerlendirebileceğimiz açıklanmaktadır. Bunun için YZ'nin yapabildikleri ile meslek gruplarında gereken insan becerileri karşılaştırılmaktadır. Ayrıca, bu karşılaştırmanın çalışma hayatında ve eğitim sisteminde ne tür değişimlere yol açabileceği ele alınmakta ve bunun gelecekteki politika tartışmalarına nasıl yön verebileceği tartışılmaktadır.

Göstergelerin, İnsan Becerilerine Yönelik Mesleki İhtiyaçlarla Eşleştirilmesi

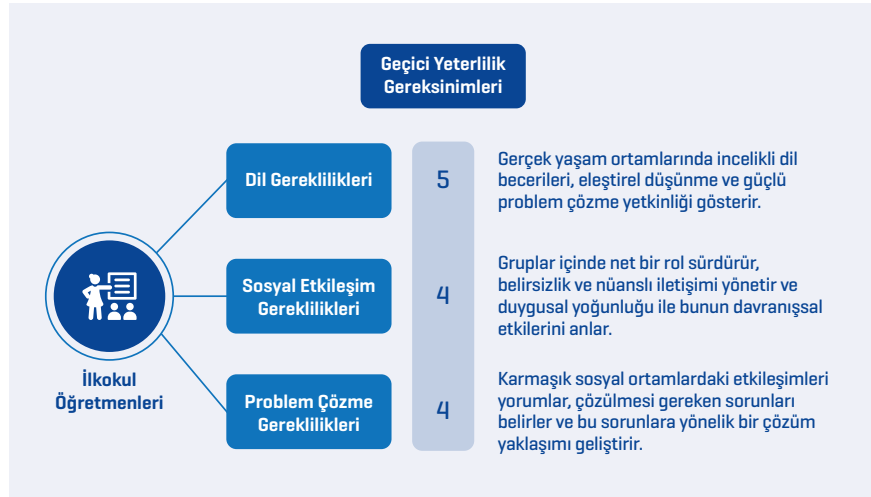
Yapay zekânın ekonomi ve toplumda köklü değişimlere yol açma potansiyelini değerlendirebilmek için ilk olarak, hangi mesleklerin ve bu mesleklerin hangi görevlerinin, YZ'nin yakın gelecekte yerine getirebileceği beceriler gerektirdiğini belirlemek gerekir. Bu analiz, OECD YZ Yetkinlik Göstergelerinin meslek tanımları ve bu mesleklerin temel görevleriyle ilişkilendirilmesi ile başlar. Böylece her meslek için gerekli beceriler ortaya konur ve YZ'nin bu becerilere ne ölçüde yaklaştığı değerlendirilebilir.

YZ Yetkinlik Göstergeleri, dokuz göstergenin düzey tanımları ile iş tanımlarını karşılaştırarak gerçek hayattaki farklı mesleklerin gerektirdiği beceri ve yetkinliklerle doğrudan ilişkilendirilebilir. Bu ilişkilendirme iki şekilde yapılabilir. Birincisi, meslekleri tüm özellikleriyle bir bütün olarak ele almak. İkincisi ise bu mesleklerin içindeki belirli görevlere daha dar bir çerçevede odaklanmak. Amaç, bir mesleği ya da görevi etkili biçimde yerine getirebilmek için her bir YZ Yetkinlik Göstergesi açısından hangi düzeyde yetkinlik gerektiğini belirlemektir.

Bu çalışma, ABD merkezli ONET sistemi kullanılarak yapılabilir. Çünkü ONET, dünyadaki en kapsamlı meslek özellikleri veri tabanlarından biridir. O*NET yaklaşık 900 meslek hakkında veri içerir. Ayrıca sınıflandırma sistemi, birçok ülkenin kullandığı meslek sınıflandırmaları ile kısmen uyumlu hâle getirilmiş ya da karşılaştırılmıştır. O*NET ve onu temel alarak oluşturulan diğer meslek veri tabanları; ayrıntılı görev tanımları, çalışma ortamı bilgileri ve gerekli insan becerileri gibi özellikleri içerir. Bu bilgiler, YZ Yetkinlik Göstergeleri ile doğrudan karşılaştırılabilir.

Bu bölüm, mesleklerin gerektirdiği becerileri YZ yetkinlik göstergeleriyle karşılaştırmaya dayalı yaklaşımı öğretmenlik mesleği örneği üzerinden açıklamayı amaçlamaktadır. Şekil 4.1'de gösterildiği gibi, öğretmenlik mesleğinin önemli bir bölümü; dil kullanımı, sosyal etkileşim ve problem çözme alanlarında en üst düzey becerileri gerektirmektedir. Bu durum, özellikle belirli bir öğretmenlik görevi incelendiğinde daha açık biçimde görülmektedir. O*NET'te yer alan "öğretim yöntemlerini ve materyallerini öğrencilerin farklı ihtiyaç ve ilgi alanlarına uyarlamak" görevi; YZ Yetkinlik Göstergelerinde dil, problem çözme ve sosyal etkileşim alanlarında 4. ve 5. düzeyde tanımlanan becerilere dayanmaktadır. Bu çerçevede öğretmenler, farklı beceri düzeylerine, kültürel geçmişlere ve öğrenme stillerine sahip öğrencilere açık ve anlaşılır yönergeler vermek ve geri bildirimlerini buna göre uyarlamak için ileri düzey dil ve sosyal etkileşim becerileri kullanmaktadır [dil ve sosyal etkileşim, düzey 5]. Aynı zamanda karmaşık bir sosyal ortamda gerçekleşen etkileşimleri yorumlayarak bireysel ya da grup düzeyinde anlama engellerini belirler ve bu engelleri aşmaya yönelik hedefe yönelik yaklaşımlar geliştirirler [problem çözme, düzey 4]. Buna ek olarak öğretmenler, güven ilişkisi kurar, belirsizlikleri yönetir ve duygusal ipuçlarına uygun şekilde karşılık verir. Tüm bunları yaparken, farklı özelliklere sahip öğrencilerin derse katılımını ve motivasyonunu destekleyen güvenli ve destekleyici bir öğrenme ortamı oluştururlar [sosyal etkileşim, düzey 5]. Ancak bu ilk değerlendirmeler henüz geçicidir. Daha sağlam sonuçlara ulaşabilmek için ek doğrulama çalışmaları ve uzman görüşleriyle geliştirilmesi gerekmektedir.

Şekil 4.1: Dil, problem çözme ve sosyal etkileşim alanlarındaki YZ Yetkinlik Göstergelerinin öğretmenlik görevlerinin gerektirdiği becerilerle eşleştirilmesi



Kaynak:

OECD [2025], Introducing the OECD AI Capability Indicators, OECD Publishing, Paris, <https://doi.org/10.1787/be745f04-en>. Bu görsel OECD'nin orijinal çalışmasının Türkçe çeviri ve uyarlamasıdır. Orijinal çalışma ile çeviri arasında herhangi bir farklılık olması durumunda yalnızca orijinal metin geçerlidir.

Bu tür bir analiz, çok aşamalı bir süreçle daha geniş ölçekte uygulanabilir:

- Alan uzmanlarına yapılacak bir anket aracılığıyla, temsil gücü olan bir meslek örneklemini için, meslekleri ve mesleki görevleri yerine getirmek üzere gerekli beceri düzeylerine ilişkin değerlendirmeler toplanabilir.
- Bu değerlendirmeler, yapay zekâ temelli yöntemler ve istatistiksel çıkarım teknikleri kullanılarak O*NET ve Avrupa Beceri, Yetkinlik, Yeterlilik ve Meslekler sınıflandırmasında [ESCO] yer alan tüm mesleklere ve on binlerce göreve genellenebilir.

Meslekler ve bu mesleklerin görevleri, YZ Yetkinlik Göstergelerinde tanımlanan gerekli beceri düzeyleriyle eşleştirildikten sonra, YZ'nin mevcut yetkinliği ile belirli bir meslek ya da görevin gerektirdiği yetkinlik arasındaki fark hesaplanabilir. Bu hesaplamalar, YZ'nin hangi meslekleri ve görevleri mevcut yetkinlik düzeylerinde yerine getirebileceğini belirlemek için kullanılabilir. En kolay belirlenebilecek durumlar, YZ'nin gerekli tüm becerilere sahip olduğu ve bu nedenle tam otomasyonun mümkün olduğu meslekler ya da görevlerdir. Ancak analizler, YZ'nin gerekli becerilerin yalnızca bir kısmına sahip olduğu durumları da ortaya koyabilir. Bu tür durumlar, YZ'nin sahip olmadığı tamamlayıcı becerilere sahip çalışanlarla insan-YZ iş birliği olasılığına işaret eder. Bu analizler ayrıca, YZ göstergelerindeki farklı ilerleme senaryolarının çeşitli meslekleri ve daha geniş ölçekte ekonomiyi nasıl etkileyebileceğini gösteren özet göstergeler geliştirmek için de kullanılabilir.

YZ Yetkinlik Göstergelerinin meslek ve görev gereklilikleriyle eşleştirilmesi, meslekler içindeki belirli görevlerin mevcut ve gelişmiş YZ yetkinlikleriyle nasıl evrilebileceğinin daha derinlemesine analiz edilmesi için bir başlangıç noktası sunmaktadır. Böyle bir inceleme, ilgili meslek alanındaki paydaşlarla yapılacak yapılandırılmış bir tartışmayı gerektirir. Bu tartışmada şu konular ele alınır:

- Toplumsal belirli bir görevin YZ tarafından gerçekleştirilmesini isteyip istemediği.
- Bu görevi yerine getirebilecek YZ sistemlerinin hâlihazırda mevcut olup olmadığı ya da geliştirme aşamasında bulunup bulunmadığı.
- İnsanların YZ sistemleriyle birlikte çalışabilmesi için hangi uyarlamaların gerekli olduğu.

Bu çalışmanın sonunda, dönüştürülmüş görev senaryoları oluşturulacaktır. Bu senaryolar, ilgili sektör temsilcileri, eğitim ve öğretim sağlayıcıları ile YZ uzmanları tarafından değerlendirilecektir. Bu senaryolar, öğretmen rollerinin ve bazı diğer yetişkin rollerinin nasıl değişebileceğini somut biçimde gösterebilir. Aynı yöntem, farklı ülkelerde ve ülke içindeki farklı bölgelerdeki beceri profillerine de uygulanabilir. Böylece politika yapımcılar, hangi bölgelerin YZ'ye daha fazla ya da daha az maruz kalabileceğini belirleyebilir ve buna uygun politika adımları geliştirebilir.

OECD'nin bu yaklaşıma ilişkin daha ayrıntılı açıklamaları, rapora eşlik eden teknik cildin 14. bölümünde yer almaktadır [OECD, 2025].

Senaryoları somutlaştırmak amacıyla, Örnek 4.1'de YZ'nin problem çözme ölçeğinde orta düzeyde olduğu bir durum örneği sunulmaktadır. Bu senaryoda YZ, doğa bilimleri, tıp ve mühendislik alanlarında yüksek düzeyde problem çözme becerisine sahiptir. Ancak sosyal ve etik muhakeme ile genel problem çözme alanlarında daha sınırlı yetkinliğe, ayrıca orta düzeyde duyuşal-motor becerilere sahiptir.

Örnek 4.1: Problem Çözme Ölçeğinde Orta Düzeydeki YZ'nin Salgınla Mücadeleyi Desteklediği Bir Senaryo

Yazın kavurucu sıcağında, yayılma hızı çok yüksek olan ve sivrisinekler yoluyla bulaşan bir salgın hastalık ortaya çıkar ve hızla ülke çapında ve uluslararası alanda büyük şehirlere yayılır. Bir yapay zekâ sistemi, birbirinden farklı veri kümelerini analiz ederek sorunu tespit eder ve kamu yetkililerine bu salgını bir pandemi olarak ilan etmelerini önerir. Tıp merkezleri, araştırma kurumları ve devlet kuruluşları, hastalığı araştırmak, bir tedavi yöntemi ve aşı geliştirmek amacıyla ortak bir program **başlatır**.

YZ sistemleri, daha merkezi veri toplama süreci başlamadan önce hastalığın nasıl yayıldığını ve hangi semptomları içerdiğini **izlemeye** ve **modellemeye** devam eder. Bu bilgiler, sistemlerin umut verici önleme ve kontrol altına alma tedbirlerini **tespit etmesine**, kamu otoritelerinin bu tedbirleri **benimsemesine** ve **uygulamasına** yardımcı olur ve etkili tedavilerin değerlendirilmesini başlatır.

Yapay zekâ sistemleri, dünya genelinde toplanan örnekleri ve verileri analiz etmek amacıyla onlarca robotik moleküler biyoloji laboratuvarını **eşgüdüm içinde** çalıştırarak deneyler **tasarlamakta** ve **yürütmektedir**; üstelik bunu gerçek zamanlı olarak gerçekleştirmektedir. Bu olaydan önce biyomedikal literatür, yapay zekâ sistemlerini bilgilendirmiş ve biyomedikal araştırma verilerinin son derece geniş bir çeşitliliğini ortaya koymuştur. YZ, bu bilgiyi kullanarak virüsün hangi biyolojik yolları etkiliyor olabileceğine dair **hipotezler geliştirir**.

Bilinen literatüre dayanarak hipotezler önceliklendirildikten sonra, araştırmacılar fareler üzerinde hedefli deneyleri **uygular**. Sonuçlarla çalışan bir YZ sistemi, nadir görülen bir nörolojik durumla ilginç bir bağlantı **keşfeder** ve uygulanabilir bir tedavi **önerir**.

YZ, dünya genelinde hükümetlerin tıbbi malzeme stoklarını **yönetmesine** yardımcı olarak hastalığa karşı adil ve verimli bir dağıtımın sağlanmasına katkıda bulunur. Ayrıca, yapay zekâ tarafından hızlandırılan araştırma süreçleri, viral etki mekanizmalarının **belirlenmesi** ve etkili bir aşının **geliştirilmesi** açısından kilit öneme sahiptir. Günler içinde binlerce doz **üretilir** ve **dağıtılır**; hastalık kontrol altına alınır ve etkilenen kişiler iyileşme sürecine girer.

Kaynak: [Gil ve Selman, 2019[2]]'den uyarlanmıştır.

Bu senaryoda YZ sistemleri:

- Veri ve bilgiye erişim, izleme, analiz ve yorumlama alanlarında güçlü yetkinlikler sergiler. Bu yetkinlikler, sorunları tespit etme kapasiteleri ve buna bağlı araştırma-geliştirme faaliyetleriyle ilişkilidir.
- Hem politika alanında hem de sağlık hizmetleri alanında çözümleri belirleyebilir, destekleyebilir ve tasarlayabilir.
- Bilimsel bilgiye ve yeni verilerin analizine dayanan karmaşık sistemlerle çalışabilir. İnsan biyolojisinden daha geniş ekosistem dinamiklerine kadar uzanan alanlarda modelleri etkili biçimde kullanır; buna sosyal ilişkiler ve insan-doğa etkileşimleri de dâhildir. Bu durum, kanıttan hareketle akıl yürütme kapasitesini göstermektedir.

Buna karşılık insanın rolü şu alanlara yoğunlaşır:

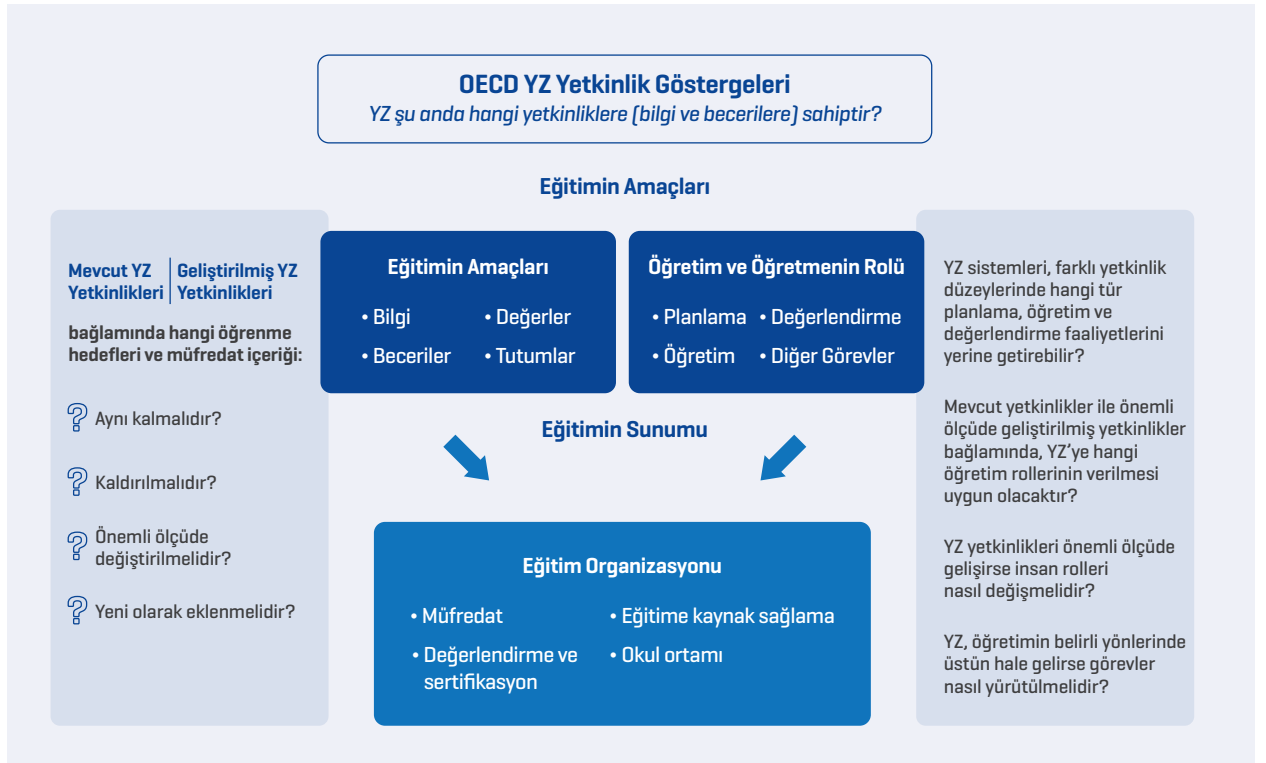
- Hastalık tespit edildikten sonra krizin kontrol altına alınması ve nihayetinde çözüme kavuşturulması amacıyla kamu ve özel sektör çabalarını yönlendirmek.
- YZ'nin deneysel araştırma protokollerinin uygulanmasını sağlamak ve YZ ile robotlar tarafından geliştirilen ve üretilen aşıların etkili biçimde dağıtılmasını temin etmek. Bu süreçte özellikle çeşitli müdahale ve uygulama faaliyetleri için gerekli olan el becerileri ve görsel beceriler başta olmak üzere çeşitli fiziksel becerilere ihtiyaç duyulmaktadır.
- Kriz yönetimi ve iş birliği faaliyetlerinde sosyal becerilerden ve etik akıl yürütmeden yararlanmak.

Eğitimde Dönüştürücü Değişim

Yukarıda sunulan göstergeler ve mesleki analiz temel alınarak eğitimde dönüştürücü değişim potansiyelini incelemek için yeni bir çerçeve ortaya çıkmaktadır [Şekil 4.2]. Bu örnek, eğitim politikasıyla doğrudan ilişkili olduğu için eğitim sektörüne odaklanmaktadır; ancak ilke olarak her meslek ya da ekonomik sektöre uygulanabilir. Bu çerçeve hem eğitimin amacına hem de nasıl sunulduğuna ilişkin kararları destekleyebilir. Özellikle YZ'nin bazı öğretim görevlerinde insan performansını aşması durumunda, eğitim uygulamalarının nasıl değişebileceğine dair tartışmalara yön verebilir. Ayrıca, mevcut ya da gelecekteki YZ yetkinliklerine bağlı olarak; hangi öğrenme hedeflerinin ve müfredat içeriklerinin aynı kalacağı, hangilerinin çıkarılması ya da güncellenmesi gerektiği ve hangilerinin yeni olarak eklenmesi gerektiği gibi temel soruların ele alınmasına yardımcı olur.

Yukarıda sunulan göstergeler ve mesleki analiz temel alınarak eğitimde dönüşümsel değişim potansiyelini incelemek üzere yeni bir çerçeve ortaya konulmaktadır [Şekil 4.2]. Bu örnek, eğitim politikasıyla doğrudan ilişkili olması nedeniyle eğitim sektörüne odaklanmaktadır; ancak ilke olarak herhangi bir meslek ya da ekonomik sektöre de uygulanabilir. Bu çerçeve, hem eğitimin amacı hem de eğitimin sunumu konusunda alınacak kararları bilgilendirebilir. Ayrıca, yapay zekâ yeteneklerinin bazı öğretim görevlerinde insan performansını aşması durumunda eğitim uygulamalarının nasıl evrilebileceğine ilişkin tartışmalara yön verebilir. Bunun yanında, mevcut ya da gelecekteki yapay zekâ yetkinliklerine bağlı olarak hangi öğrenme hedeflerinin ve müfredat içeriklerinin aynı kalacağı, hangilerinin kaldırılması ya da gözden geçirilmesi gerektiği ve hangilerinin yeni olarak programa eklenmesi gerektiği gibi temel soruların ele alınmasına da imkân tanır.

Şekil 4.2: Eğitimde artan YZ Yetkinliklerinin etkilerini analiz etmeye yönelik bir çerçeve



Kaynak: OECD (2025), Introducing the OECD AI Capability Indicators, OECD Publishing, Paris, <https://doi.org/10.1787/be745f04-en>. Bu görsel OECD'nin orijinal çalışmasının Türkçe çeviri ve uyarlamasıdır. Orijinal çalışma ile çeviri arasında herhangi bir farklılık olması durumunda yalnızca orijinal metin geçerlidir.

Eğitimde iki temel potansiyel dönüşüm alanı öne çıkmaktadır. Birincisi, YZ öğretmenlerin görevlerinin önemli bir bölümünü yerine getirebilir ise, eğitimin sunuluş biçimi değişebilir. Örneğin YZ güvenilir biçimde ders anlatabiliyor ya da geri bildirim verebiliyorsa, öğretmenlerin geleneksel görevi daha çok rehberlik, teşvik etme veya karmaşık kişilerarası etkileşimleri yönetmeye doğru kayabilir. Böyle bir gelişme, öğretimin nasıl yürütüldüğünü yeniden tanımlayabilir ve sınıf içi dinamikleri, öğretmen-öğrenci etkileşimini ve genel öğrenme deneyimini dönüştürebilir.

İkincisi, öğrencilerin öğrenme hedefleri; yetişkin olduklarında iş hayatında, toplumda ve günlük yaşamda yerine getirmeleri beklenen görevlerle bağlantılıdır. Eğer YZ artık bu görevlerin birçoğunu yapıabiliyorsa, eğitim sistemlerinin öğrencilerin ne öğrenmesi gerektiğini yeniden düşünmesi gerekebilir. Bu olası etkiler, yapay zekânın eğitimde kullanılan birçok değerlendirmede öğrencilerden daha yüksek performans göstermesinde şimdiden görülmektedir. Buna, OECD'nin 15 yaş grubundaki öğrencilerin becerilerini ölçen Uluslararası Öğrenci Değerlendirme Programı (PISA) da dâhildir. Bu gelişmeler sonucunda bazı mevcut hedefler önemini yitirebilirken, yaratıcılık ya da etik muhakeme gibi bazı beceriler daha da önemli hâle gelebilir.

Elbette bu değişimler yalnızca teknik konular değildir. Toplumsal değerler, öğretmenlerin bazı görevleri YZ'ye devredip devretmeyeceğini ve öğrencilerin hangi bilgi ve becerileri öğrenmesi gerektiğini belirlemede etkili olacaktır. YZ Yetkinlik Göstergeleri ve yukarıdaki mesleki analiz bu değer temelli sorulara doğrudan cevap vermez. Ancak teknik olarak hangi alanlarda büyük değişimlerin mümkün ve hatta olası olduğunu göstermeye yardımcı olabilir. OECD, bu olasılıkları öğretmen rolleri ve öğrencilerden beklenen bilgi ve becerilere ilişkin senaryo temelli analizlerle daha ayrıntılı inceleyecektir. Hazırlanacak senaryolar, öğrenme hedefleri ve müfredat içeriğinde ortaya çıkabilecek değişimleri tartışmaya açacaktır. Eğitim politikası yapıcılarının OECD YZ Yetkinlik Göstergelerinden nasıl yararlanabileceğine dair daha ayrıntılı açıklamalar, bu rapora eşlik eden teknik cildin 14. bölümünde (OECD, 2025) yer almaktadır.



SONUÇ

Bu raporda sunulan YZ Yetkinlik Göstergeleri, yapay zekâdaki ilerlemenin insan becerileriyle nasıl örtüştüğünü değerlendirmek için basit ama etkili bir araç sunmaktadır. Göstergeler, YZ performansına ilişkin veriler ve uzman görüşleri temel alınarak geliştirilmiştir. YZ yetkinliklerini insan becerileri ile ilişkilendirerek tanımladıkları için açık ve anlaşılır bir çerçeve sağlarlar. Bu sayede iş hayatı ve eğitim gibi alanlarda YZ'nin ilerlemesinin olası etkileri daha net biçimde değerlendirilebilir.

İnsan becerileriyle yapılan bu karşılaştırma, YZ'nin yapay genel zekâyâ doğru ne ölçüde ilerlediğini izlemek için de tutarlı bir ölçüt sunar. Aynı zamanda, YZ'nin gerçek dünyadaki iş gereklilikleri ve eğitim hedefleriyle bağlantısını kurarak büyük değişimlerin nerede ortaya çıkabileceğini ve insan rolünün hangi alanlarda temel olmaya devam edeceğini göstermeye yardımcı olur.

Göstergeler, YZ araştırmaları açısından da önemlidir. YZ ilerlemesini anlamlı biçimde değerlendirmek için hangi yetkinliklerin test edilmesi gerektiğine işaret eder. Ayrıca politika yapıcıların, YZ'nin toplumsal, siyasal ve etik etkileri konusunda hangi alanların dikkatle değerlendirilmesi gerektiğini araştırmacılara açık biçimde iletmesini sağlar.

Bu çalışma bir başlangıç niteliğindedir. Göstergeler ve bunlara dayalı analizler geliştikçe, politika yapıcılar hızla değişen YZ alanında daha net ve güvenilir bilgilere dayanarak karar verilecektir. Böylece gelişmiş YZ sistemlerinin sunduğu fırsatlardan yararlanırken olası riskleri de daha bilinçli biçimde yönetmek mümkün olacaktır. OECD, bu göstergeleri geliştirmeye ve güncellemeye devam ederek YZ yetkinlikleri ve bunların eğitim, çalışma hayatı ve sivil toplum üzerindeki etkileri konusunda güvenilir bir uluslararası başvuru kaynağı olmayı amaçlamaktadır.



REFERANSLAR

- [2] Gil, Y. and B. Selman [2019], "A 20-year community roadmap for artificial intelligence research in the US", arXiv, Vol. 1908.02624, <https://arxiv.org/abs/1908.02624>.
- [1] OECD [2025], AI and the Future of Skills Volume 3: The OECD AI Capability Indicators, OECD Publishing.

[illegible]



T.C. MİLLÎ EĞİTİM BAKANLIĞI
YENİLİK VE EĞİTİM TEKNOLOJİLERİ
GENEL MÜDÜRLÜĞÜ



yegitek.meb.gov.tr

Emniyet Mahallesi, Milas Sokak, No:8 Yenimahalle / ANKARA



mebyegitek