

## Bibliyometrik Verilerle Eğitimde Sohbet Robotları

Çağla Ediz<sup>1</sup>

Alpaslan Kibar<sup>2</sup>

### Özet

Son yıllarda, özellikle üretken yapay zeka araçlarına dayalı teknolojilerin ve bu kapsamda hizmet veren büyük dil modellerinin gelişmesi, öğrencilere kariyer planlarını çizmede ve zamanlarını etkin kullanmalarında yol gösterebilecek, eğitim ve mentörlük hizmeti sağlayan sohbet robotlarının geliştirilebilmesini mümkün kılmıştır. Henüz başlangıç aşamasında olan bu çalışmaların, geleceğin eğitim dünyasını yönlendirmesi ve eğitimde yeni bir dönüm noktası olarak kabul edilmesi uzak bir ihtimal değildir. Eğitimde bir çığır açma ihtimali olan eğitim sohbet robotlarının incelenmesi amacıyla bu çalışma, Web of Science'da bu alanda hazırlanan makaleleri bibliyometrik analizlerle değerlendirdi. Bu çalışmada öncelikle eğitim alanında sohbet robotlarına neden ihtiyaç duyulduğu incelendi, sonra ise sırasıyla büyük dil modellerinin gelişimi ve eğitimde sohbet robotlarının kullanımı araştırıldı. Daha sonraki bölümde bibliyometrik analizler için veri indirme ve hazırlama işlemleri anlatıldı. Ardından bibliyometrik analizler ve bu analizlerden elde edilen bulgular paylaşıldı. Çalışılan analizler, yıllar bazında yayın sayıları ve en üretken yayıncıların ülkeleri, en çok yayın yapan dergiler, en çok yayın yapılan wos kategorileri, en üretken üniversiteler ve yazarlar ile en çok atıf alan makalelerdir. Ayrıca eğitimde büyük dil modellerinin işlevlerinin değerlendirilme yöntemleri üzerinde duruldu ve gelecek için önerilerde bulunuldu.

1 Doç. Dr., Sakarya Üniversitesi, cediz@sakarya.edu.tr, <https://orcid.org/0000-0002-0793-3722>

2 Dr. Öğr. Gör., Sakarya Üniversitesi, kibar@sakarya.edu.tr, <https://orcid.org/0000-0002-6898-1736>

## 1. Giriş

Eğitim sohbet robotları üzerine hazırlanmış literatürü inceleyen Wollny vd., bu alandaki gelecek çalışmaların üç açıdan geliştirilmesi gerektiğini belirtir. Bunlar, sohbet robotu değerlendirmelerindeki bilişsel ve duygusal eksikliğin giderilmesine ait çalışmalar, mentörlük için öğrencinin nele ihtiyacı duyduğunun belirleyen çalışmalar ve kişiliğe göre uyarlanabilir adaptasyon yeteneklerini keşfederek buna göre destek veren çalışmalardır (Wollny vd., 2021). Geleneksel eğitim programları, öğrencilerin yapmak istedikleri araştırmalarda veya çalışmak istedikleri alanlarda ihtiyaç duyacakları programları belirleyemez ve bu nedenle öğrenciyi geniş bir müfredata sevk eder. Bu da çoğu zaman öğrencinin zamanını etkin kullanamamasına ve gerçekte ihtiyaç duyduğu gerekliliklere yeterince yönlenebilmesine sebep olur (Berg, 2023). Bu durumu engellemek için öğrencilere atanacak bir mentör, öğrencinin entelektüel yeteneklerini ve eleştirel düşünmesini geliştirir; anlamlı ve önemli görevler verir, mesleğe girişini ve kariyerini kolaylaştırır; tavsiye, teşvik ve geri bildirim sağlar; sahadaki kilit oyuncuları tanıtır; geleceğe yönelik beklentileri artırır (Girves vd., 2005). Genel olarak ikili, akran, kolaylaştırılmış akran, hızlı, işlevsel, grup ve uzaktan olmak üzere yedi mentörlük hizmeti tanımlanmaktadır ve bunlar arasında en yaygın bir danışanı daha kıdemli veya daha deneyimli bir mentörle eşleştiren ikili modeldir (Kashiwagi vd., 2013). Bir mentörün danışmanlık hizmeti verebilmesi ancak kısıtlı sayıda öğrenci için söz konusu olabilmekte ve öğrenciye sağlayabileceği fayda, mentörün kişisel deneyimi, çalışma alanları ve ayırabileceği zaman dilimi gibi kısıtları içermektedir. Bir mentörün sadece ders danışmanlığı hizmeti için bile ders programları, öğrencinin transkripti, öğrencinin ilgi alanları ve öz yeterliliği, intibak durumları, müfredat, yönetmelikler gibi farklı platform ve disiplinler arası verilerden beslenen verilere göre değerlendirmede bulunması gerekir. Çoğu zaman bu çok mümkün olmadığından standart uygulamalarla işlemler yürümektedir. Bunu destekleyici bir şekilde, insan mentörlüğünde yürütülen bir çalışmada danışanlar, mentörlerin zaman eksikliğinin programa zarar verdiğini hissetti, mentörlüğünün külfetli lojistiği, fiziksel mesafeyi, danışanların mentörlük ilişkilerini yüzeysel ve sömürücü olarak algılamasını problemler olarak gördüler ([Benson vd., 2002; Pololi & Knight, 2005; Moss vd., 2008; Alleyne vd., 2009]aktaran Kashiwagi vd., 2013; Wollny vd., 2021). Yaşanan bu tür problemler, eğitim ve danışma işlerinde kullanılacak sohbet robotları tarafından aşılabilir. Bu nedenle geleceğin eğitim sisteminde öğrencilere yol gösterecek ve onlara gerektiğinde doğru eğitim materyallerini sunabilecek sohbet robotları yaygın bir şekilde kullanılacağını tahmin etmekteyiz.

### 1.1. Büyük Dil Modellerinin Gelişimi

İlk sohbet robotu olarak kabul edilen Eliza, girdi mesajından elde ettiği anahtar kelime eşleştirmeleri kullanarak, kural tabanlı bir yaklaşımla, girdi mesajlarına çıktı oluşturmaktaydı (Weizenbaum 1966; Özkol vd., 2019). Eliza, kullanıcı ifadelerini yeniden revize ederek ve kullanıcı girdilerine yeni oluşturduğu sorularla cevap vererek bir terapist gibi davranmaktaydı ve Tureng testini geçen ilk programlardan biri olarak kabul ediliyordu (Sharma vd, 2017). Ardından, farklı model kullanımlarıyla (en çok tercih edilen model, bağlama dayalı olarak kelime olasılık tahminlerinin gerçekleştiği n-gram modelleridir) kelime dizilerinin olasılığını gören ve verilen girdiye dayalı olarak yeni metin üretme yeteneğine sahip olan Dil Modelleri (DM) geliştirildi. Büyük Dil Modellerinin (BDM) atası olarak kabul edilebilecek DM'lerde nadir kelime sorunları, aşırı uyum sorunu ve karmaşık dilsel olguları yakalama problemleriyle karşılaştılar (Chang vd., 2024).

BDM'ler ise olağanüstü öğrenme yeteneklerine sahip gelişmiş dil modelleridir. Ardışık verileri verimli bir şekilde işleme, paralelleştirmeye izin verme ve metindeki uzun menzilli bağımlılıkları yakalama yeteneklerine sahip transformatörler, BDM'lerde temel modül olarak kullanılmaktadır ve transformatörlerin kullanımı NLP alanında bir evrim yaratmıştır (Chang vd., 2024). İnsan tarafından oluşturulan yanıtlarla ince ayar yapılmasına, hatalarda insan geribildiriminden performanslarının geliştirilmesine izin veren öğrenim yetenekleri BDM'lerin bir diğer önemli yönüdür (Chang vd., 2024).

BDM üzerine olan araştırmalar sistematik bir şekilde incelendiğinde, BDM tabanlı otonom ajanın başlıca dört aşamalı harekete sahip oldukları söylenebilir. Bunlar (Wang vd., 2024);

- **Profil aşaması:** Alıcılardan gelen demografik, psikolojik, sosyolojik bilgiler değerlendirilebileceği ve ajan profilinin çıkarılacağı aşamadır. Bu aşamada, manuel veya makine öğrenmesi metotlarından faydalanılarak gerçek dünyadan elde edilen veriler sınıflandırılır. Bazı çalışmalarda sınıflandırma işlemi gerçekleştirmek için BDM'den faydalanılmıştır. Ancak bu çalışmalarda çok sayıda ajan (politikacı/gazeteci/iş insanı; içe dönük/ dışa dönük vb.) tanımlandığında, emek işçiliği çok büyüyebilmektedir.
- **Hafıza aşaması:** Bu aşamada tutulan verilerle geçmiş davranışlarını hatırlar ve gelecek eylemleri için karar alıp hareket ederler ve hareketlerini çıktılara dönüştürürler. Veriler farklı veri tabanlarında,

vektörlerde, doğal dil ifadelerinde veya yapılandırılmış üçlü ifadelerde saklanabilir.

- Planlama Aşaması: Burada sistem bir ağaç yapısı boyunda seçimlerde bulunarak akıl yürütür. Sistemin yaptığı seçimler için geri bildirimler de iyileştirme amaçlı kullanılabilir.
- Eylem aşaması: Bu aşamada ajan görevi gerçekleştirir, diğer ajanlar veya insanlarla bilgi paylaşımında bulunabilir ve öz değerlendirme yapabilir.

## 1.2. Eğitimde Sohbet Robotları

Eğitimde kullanılan sohbet robotları işlevselliğine göre başlıca hizmet desteği ve eğitim desteği olarak ikiye ayrılmaktadır (Wollny vd., 2021). Hizmet desteğinde, özellikle kayıt dönemleri yaşanan aşırı yüklemeyi azaltmak, kampüs ile ilgili bilgi vermek, ders yardımı ve kütüphane hizmetlerine yardımcı olmak için geliştirilen robotlar sayılabilir ([Brustenga vd., 2018; Mckie, & Narayan, 2019; Griol vd., 2017]; akt. Wollny v.d., 2021; Perez vd., 2020 ). Eğitim desteğinde ise muhasebe, ingilizce, bilgisayar bilimleri, matematik gibi farklı konuları öğretmek için geliştirilen sohbet robotları mevcuttur (Perez vd., 2020). Örneğin daha genel konular için eğitilmiş olan ChatGPT, öğrencilere anında açıklamalar sunup, kavramları netleştirip, öğrencilere ek kaynaklar önererek, sanal bir danışman gibi davranır (Kong vd., 2023). Ancak ChatGPT gibi sohbet robotları, Wikipedia, bloglar gibi eğitildikleri materyallere bağlı olarak hareket edeceklerinden, yanlış yönlendirme yapmış olabilirler ve daha bilimsel kaynaklarla doğruluklarının kontrol edilmesine ihtiyaç duyarlar (Berg, 2023). Bu nedenle etkileyici bir şekilde dil üretme yeteneklerine sahip olsalar da, spesifik alanlar için eğitilmediklerinden bu alanlarda güvenilir ve uygulanabilir değildir (Parisi vd., 2019; Özmen & Schwarz, 2024). Öğrencilere eğitim desteği veren bir başka çalışma Neumann vd.'e aittir. Bu çalışmada Alman üniversitesinde hizmet öncesi öğretmen eğitimi müfredatında yer alan eğitim bilimi dersi için, eğitim ve bilgisayar bilimcilerinden oluşan disiplinler arası bir ekip iki sohbet robotu geliştirildi. Bunlar, mentörlük sağlama amacıyla geliştirilmiş sesli yapay zeka sohbet robotlarıydı. Geliştiriciler bu robotları Feedbot ve LitBot olarak adlandırdılar. Feedbot literatür hakkında yazma görevleri sunmakta ve aynı zamanda öğrencilerin ortaya çıkan metinleri hakkında otomatik geri bildirim sağlamaktaydı. Litbot ise öğrencilerin okuma sürecini bireysel olarak desteklemekteydi. Yazarlar hazırlanan sohbet robotlarını kullanıcı anketleriyle ve açık uçlu sorularla değerlendirerek, gelecekteki eğitimle ilgili sohbet robotlarının geliştirilmesinde faydalanılacak bir yol haritası oluşturdu (Neumann vd., 2021).

## 2. Bibliyometrik Analizler

### 2.1. Veri İndirme ve Hazırlama:

Çalışmada, eğitim ve büyük dil modeli ile ilgili kayıtlara ulaşabilmek için, disiplinler arası bir şekilde alanında kaliteli çalışmalarını indeksleyen Web of Science (wos) veri tabanı kullanılmıştır. Wos'dan başlık veya yazar anahtar kelimelerinde, hem büyük dil modeli, sohbet robotu veya üretken zeka, hem de okul, üniversite veya eğitim kelimelerinin yer aldığı çalışmalar 1 Kasım 2024 tarihi itibarıyla filtrelendi. Filtreleme sorgusu şu şekildedir:

(TI=(large language model or GenAi or “Generative Artificial Intelligence” or chatbot) and TI=(school or education or university)) or (AK=(large language model or GenAi or Generative Artificial Intelligence or chatbot) and AK=(school or education or university))

İndirme işleminden sonra 1008 çalışmaya ulaşıldı. Tekrar eden kayıtlar çıkarıldığında kayıt sayısı 1004'e düşürüldü. Sadece bilimsel dergilerde yayınlanmış makaleler bırakılarak 708 kayıt üzerinden istatistiksel frekans sayma yöntemleri kullanılarak bibliyometrik analizler hazırlandı.

### 2.2. Yıllar Bazında Yayın Sayıları ve En Üretken Yayıncıların Ülkeleri

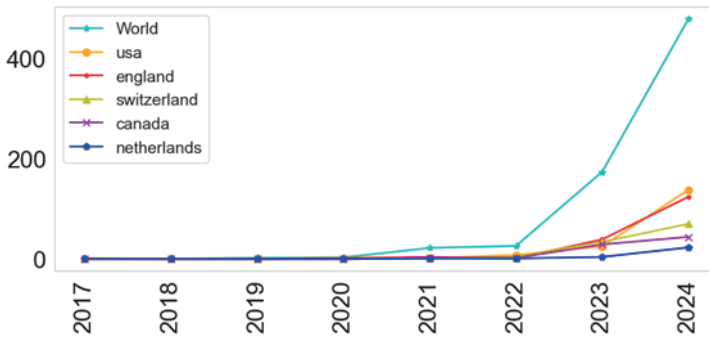
Sohbet robotu kavramı nispeten yeni bir kavram olduğu için eğitim alanında bu konu üzerine hazırlanan çalışmaların da çok büyük oranda son birkaç yılda yapıldığını söyleyebiliriz. Wos'ta bu kapsamda hazırlanan ilk çalışma Computers in Human Behavior dergisinde, 2017 yılında yayınlanmış olan “Stimulating and sustaining interest in a language course: An experimental comparison of Chatbot and Human task partners” adlı çalışmadır. Bu çalışmada, öğrenciler üç hafta aralıklarla on iki haftalık deneysel bir çalışma yürüttüler. Bu süre zarfında, öğrencilerin görevlerinde bir sohbet robotundan ve bir insandan destek almaları sağlandı. Yazarlar, yapısal eşitlik modeli kullanarak, öğrencilerin sohbet robotuyla ilgili görevlere olan ilgilerinde bir düşüş yaşanırken, insandan destek aldıkları görevlere olan ilgilerinde bir düşüş olmadığı sonucunu elde ettiler (Fryer vd., 2017). 2017 yılındaki bu çalışmanın ardından 2018 yılında bu konu kapsamında bir çalışma yapılmadığı görülmektedir. 2019 yılına ait, International Journal Of Emerging Technologies In Learning'de “Sequence to Sequence Model Performance for Education Chatbot”( Eğitim Sohbet Robotu için Sıradan Sıraya Model Performansı) (Palasundram vd., 2019) ve Procesamiento Del Lenguaje Natural'da “Atención automatizada a estudiantes en el proceso de matriculación en la Universidad de Alicante” (Alicante Üniversitesi'ndeki

lisans sürecinde öğrenciler için otomatikleştirilmiş dikkat) (Moreno vd., 2019) adlı sadece iki çalışma bulunmaktadır. 2021 yılı itibarıyla yayınlanan çalışmaların çoğaldığını ve 2024 yılının bu konuda yayınlanan makaleler açısından en verimli yıl olduğu söylenebilir (Figür 1).

Yayıncıların ülkeleri bazında yayın sayılarına (ys) bakıldığında, en çok yayın yapan ülkelerin başında Amerika ve ardından sırasıyla İngiltere ve İsviçre gelmektedir (Tablo 1). Yayın sayılarının yüzdeleri (ys %) değerlendirildiğinde; makalelerin üçte birine yakın bir kısmının Amerika'da, dörtte birine yakın bir kısmının İngiltere'de yayınlandığı görülmektedir.

*Tablo1. Yayınlarda en üretken ülkeler*

sıra	ys	ülke	ys %
1	226	Amerika	31,92
2	173	İngiltere	24,44
3	112	İsviçre	15,82
4	78	Kanada	11,02
5	30	Hollanda	4,24



*Figür 1. Senelik yayın sayıları ve en çok yayın basan ülkeler*

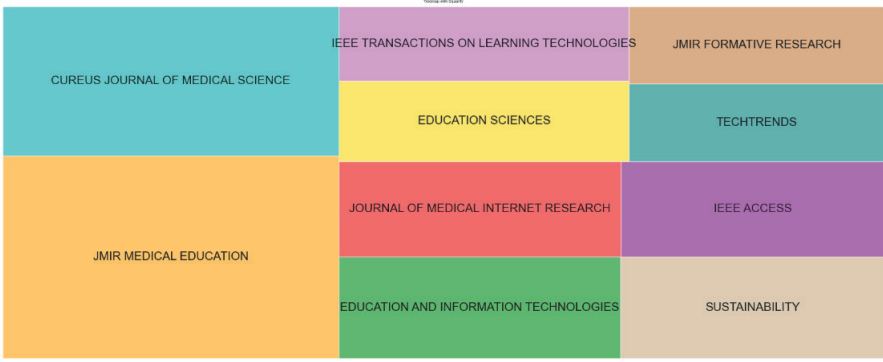
### 2.3. En Çok Yayın Yapan Dergiler

Wos'da en çok yayın yapan dergilere bakıldığında, eğitim ve sohbet robotu çalışmalarında en çok yayın yapan ilk beş dergiden üçünün sağlık alanında olduğu görülmektedir (Tablo 2). Bu konuda en çok yayın yapan dergi, 38 yayın sayısıyla (ys) JMIR Medical Education dergisidir. Bu dergiyi sırasıyla 28 yayınlı Cureus Journal Of Medical Science ve 16 yayınlı Education And Information Technologies dergileri takip etmektedir. Wos'da yayınlanan dergilere ait toplam atıf sayılarına bakıldığında, bu beş dergi içinden

en yüksek atıf sayısına 673 atıfla ikinci sırada yer alan Cureus Journal Of Medical Science sahiptir.

*Tablo 2. En çok yayın yapan dergiler*

Sıra	dergi	ys	as	ys(%)	as/ys
1	JMIR Medical Education	38	535	5,37	14,1
2	Cureus Journal Of Medical Science	28	673	3,95	24,0
3	Education And Information Technologies	16	252	2,26	15,8
4	Sustainability	15	334	2,12	22,3
5	Journal Of Medical Internet Research	15	107	2,12	7,1



*Figür 2. En üretken on dergiye ait ağaç haritalama grafiği*

#### 2.4. En Çok Yayın Yapılan Wos Kategorileri

Eğitim ve sohbet robotu üzerine çalışmalar, en çok eğitim ve eğitim araştırmaları kategorisinde yapılmıştır. Bu kategorideki yayınlar, 213 adet olup, yayınların %30,08'ini oluşturmaktadır. Yayın başına en yüksek atıf da, yaklaşık 13 adet olup, yine bu kategoriye aittir. Yayınlanan makalelerin ikinci ve üçüncü olarak en yüksek atandığı kategoriler “Eğitim, Bilimsel Disiplinler” ve “Sağlık Bilimleri ve Hizmetleri” kategorileridir.

Tablo 3. Yayınların en çok atandığı WoS Kategorileri

sıra	WoS kategorileri	ys	ys (%)	as	as/ys
1	Education & Educational Research	213	30,08	2890	13,6
2	Education, Scientific Disciplines	91	12,85	1196	13,1
3	Health Care Sciences & Services	55	7,77	419	7,6
4	Computer Science, Information Systems	42	5,93	85	2,0
5	Computer Science, Interdisciplinary Applications	40	5,65	264	6,6

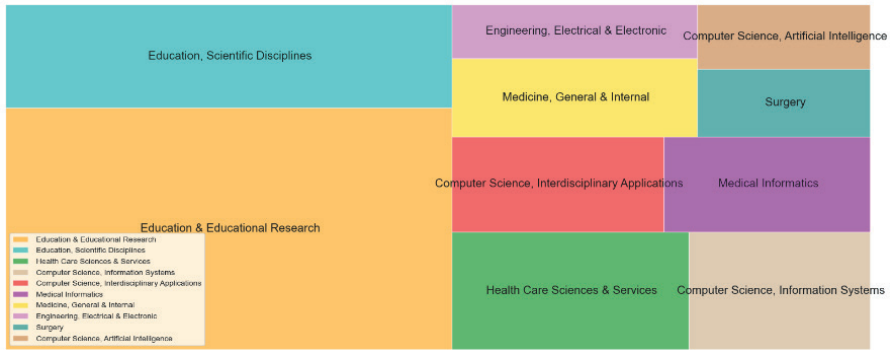


Figure 3. En çok yayın yapılan ilk on WoS kategorisi ağaç haritalama grafiği

## 2.5. En Üretken Üniversiteler

Eğitim ve sohbet robotlarının birlikte ele alındığı makalelerde en üretken üniversiteler arasında birinciliği, 12 adet yayımla iki üniversite paylaşmaktadır. Bu üniversiteler, Amerika'dan University of Hong Kong ve İngiltere'den University of London'dır. Atıf sayılarına bakıldığında University of Hong Kong'un toplam atıf sayısında ve aynı zamanda senelik ortalama atıf sayısında diğerlerinden daha öndedir. Üretkenlikte ilk beşte yer alan üniversitelerin üçü Amerika'da, biri İngiltere'de ve bir diğerinin de Meksika'da olduğu görülmektedir.



**Tablo 4. En Üretken Üniversitelere ait yayın ve atıf sayıları**

sıra	üniversite	ys	ys %	as	ys/as
1	University of Hong Kong	12	1,69	311	25,92
2	University of London	12	1,69	94	7,83
3	University of California System	12	1,69	96	8,00
4	Harvard University	11	1,55	92	8,36
5	Tecnologico de Monterrey	9	1,27	136	1.27

## 2.6. En Üretken Yazarlar

En üretken iki yazar, yayınlarını birlikte yapan ve aynı üniversitede çalışan Punya Mishra ve Danah Henriksen'dir. Bu yazarlar 6 yayın ile eğitim ve sohbet robotları ile ilgili Wos'ta en yüksek sayıda makaleye sahip yazarlardır. Üçüncü sırada ise 5 adet yayın sayısı ile Gwo-Jen Hwang adlı yazar yer almaktadır.

**Tablo 5. En üretken yazarlara ait ülke bilgileri, yayın ve atıf sayıları**

sıra	yazar	ys	ys (%)	as	as/ys	Üniversite
1	Mishra, P	6	0,85	17	2,83	Arizona State University
2	Henriksen, D	6	0,85	17	2,83	Arizona State University
3	Hwang, GJ	5	0,71	246	49,2	National Taiwan University of Science and Technology

## 2.7. En Çok Atıf Alan Makaleler:

Wos'da en çok atıf alan makale, Artificial Hallucinations in ChatGPT: Implications in Scientific Writing (ChatGPT'de Yapay Halüsinasyonlar: Bilimsel Yazımdaki Etkileri) adlı çalışmadır. Bu makalede ChatGPT'yi deneyen Hussam Alkaissi ve Samy McFarlane, sağlık ile ilgili, referans vererek yazılmış iki farklı akademik metin hazırlamak istediler. İlk çalışmada hazırlanan metinlerin içeriğinin çoğu doğru olmasına rağmen, içlerinde yanlış bir bilginin de olduğunu gördüler. Ayrıca, ChatGPT tarafından sunulan kaynakların tümünde yer alan başlıklar aslında mevcut değildi ve verilen PubMed Kimlikleri de başka kaynaklara aitti. İkinci çalışmada da benzer sonuçlar aldılar. Diğer yandan, ChatGPT'nin araştırmacılar tarafından hazırlanan kısa notları tutarlı bir metin oluşturacak şekilde bir araya getirmede çok yetenekli olduğunu gördüler (Alkaissi ve McFarlane, 2023).

Wos'da en çok atıf alan ikinci makale ise *What if the devil is my guardian angel: ChatGPT as a case study of using chatbots in education* (Ya şeytan benim koruyucu meleğimse: ChatGPT, eğitimde chatbot kullanımına ilişkin bir vaka çalışması)'dir. Bu çalışmada, eğitimde chatbot kullanımına ilişkin endişeleri incelemek amacıyla, tweetlerden sosyal ağ analizi ve mülakatlardan içerik analizi yapılmıştır. Ayrıca üç eğitimci bir hafta boyunca ChatGPT kullanarak, kullanıcı deneyimlerini aktarmışlardır. Sonuçlar, ChatGPT'nin eğitimde güçlü bir araç olduğunu ancak dikkatli kullanılmasının gerektiğini göstermiştir. İlaven, chatbot'ların güvenli bir şekilde benimsenmesi için daha fazla yönergeye ihtiyaç duyulduğu ve araştırmacıların bu alanda çalışmasını gerektirecek yeni araştırma konuları bulunduğu vurgulanmıştır (Tlili vd., 2023). Üçüncü en yüksek atıf alan *Examining Science Education in ChatGPT: An Exploratory Study of Generative Artificial Intelligence* adlı makale, ChatGPT'nin, bilim üniteleri, ölçütler ve sınavlar tasarlayan eğitimciler için oldukça yararlı bir araç olabileceğinden bahseder. Ancak eğitimciler, yapay zeka tarafından oluşturulan her kaynağı eleştirel bir bakış açısıyla değerlendirmeli ve bu materyalleri kendi öğretim bağlamlarına uygun şekilde uyarlamalıdır (Cooper, 2023).

*Tablo 6. En çok atıf alan beş makale*

sıra	başlık	ref	as
1	Artificial Hallucinations in ChatGPT: Implications in Scientific Writing	Alkaiissi ve McFarlane, 2023	484
2	What if the devil is my guardian angel: ChatGPT as a case study of using chatbots in education	Tlili vd., 2023	371
3	Examining Science Education in ChatGPT: An Exploratory Study of Generative Artificial Intelligence	Cooper, 2023	248
4	Generative AI and the future of education: Ragnarok or reformation? A paradoxical perspective from management educators	Lim vd., 2023	248
5	Stimulating and sustaining interest in a language course: An experimental comparison of Chatbot and Human task partners	Fryer vd., 2017	174

### 3. Eğitimde Büyük Dil Modellerin İşlevlerinin Değerlendirilme Yöntemleri

BDM'de istemlerin (girdiler, prompts) yanlış sonuçlar üretmesi, halüsinasyonlar olarak adlandırılmaktadır. Başlıca iki çeşit halüsinasyon vardır. Birincisi, çıktıların içeriğinin gerçek dünya verilerine bakıldığında doğru olmamasından kaynaklanmaktadır. Diğer çeşit halüsinasyonlar ise

çıktıların kullanıcı tarafından talep edilenle (istemler) tutarlı olmamasından kaynaklanır.

Eğitim ortamlarında kullanılan chatbotların değerlendirmelerinin nasıl yapıldığını inceleyen bir literatür çalışmasında; teknoloji kabul modeli, öğrenme başarısı, motivasyon artışı, fayda etkisi ve psikolojik faktör ölçümlerinin kullanıldığı görüldü (Hobert, 2019). Bu ölçümler için başlıca kullanılan araçlar şunlardır (Hobert, 2019):

- Saha veya laboratuvar çalışmasından önce veya sonra yapılan, genellikle öğrenme başarısını ölçmeye yönelik nicel anketler,
- Kullanım sonrasında gerçekleştirilen ve daha çok kullanıcının teknoloji kabulünü veya kullanıcı deneyimini ölçmeye yönelik nicel anketler,
- Sohbet robotlarının öğrenme başarısını ve farklı etkilerini değerlendirmek için katılımcılarla yapılan nitel görüşmeler,
- Sohbet robotlarıyla öğrenciler arasındaki etkileşimi değerlendiren sohbet transkriptleri ve teknik log verileri (bu veriler, genellikle söylem analizleri gibi metotlarla değerlendirilmektedir),
- İnsan - insan iletişimiyle, insan – sohbet robotu iletişiminin karşılaştırılması ([Winkler ve Soellner, 2018 ] akt. Wollny vd., 2021).

Literatür incelemelerinden halüsinasyonların; hatalı veri kaynağı kullanımı, kullanılan veri büyüklüğünün yeterli olmaması, model mimarisindeki kusurlar, eğitim amaç içeriklerindeki hatalar veya eksikler, parametrelerde doğru seçimlerin yapılamaması, hata geri bildirimlerinde yetersiz bağlamlar, dikkatsizlikler, kod çözümünde örnekleme hataları, softmax darboğazı gibi problemlerden kaynaklandığı görülmüştür (Huang vd., 2023).

#### 4. Gelecek İçin Öneriler

İnsanların sağlayabileceği hizmetlere olan ihtiyaçların doğru belirlenmesi ve istihdam edilebilirliğin artışının sağlanması; öğrenme, öğretme ve değerlendirme kalitelerinin artırılması ile ilişkilidir. Bu karmaşık ve iddialı hedefi gerçekleştirebilmek yine iddialı ve karmaşık yaklaşımları içerir (Knight, & Yorke, 2004). Öğrencilerin gelecekteki istihdamlarında başarılı olma olasılıklarını; becerileri, kişisel özellikleri, anlayışları, ilgileri, aldıkları eğitim, becerileri kullanabildikleri bir yerde istihdam edilmiş olmaları gibi birçok faktör belirlemektedir. Bu faktörlerin gözetilerek öğrencilere gereken eğitimin sağlanması iyi bir geleceğin inşası için önemlidir. Ancak, ne yazık ki mevcut düzende varolan eğitim alanında kaynakların eşitsiz dağıtımı, istikrarsız eğitim kalitesi, yetersiz içerik, güncel olmayan öğretim yöntemleri,

aşırı öğrenci iş yükü ve hatalı değerlendirme sistemleri gibi birçok sorun bulunmaktadır. Bu problemler, eğitimin adaletini ve kalitesini ciddi şekilde sınırlayarak ilerleme ve gelişimini engellemektedir (Yu, 2023). Türkiye’de eğitim sorunlarını araştıran bir çalışmada da, öğretmenler öğrencilerde vizyon eksikliğinin olduğundan, geleceğe dair planlarının olmayışından ve müfredat programının ağır olmasından yakındılar (Şener, 2018). Öğrencilere gereken becerileri kazandırma, onlara kendi kişisel özellikleri doğrultusunda vizyon sağlama ve kariyerlerini inşa ederken gerekli kariyer danışmanlığı hizmeti verebilme adına bu alanda ulusal politikaların geliştirilmesine acil bir şekilde ihtiyaç duyulmaktadır. Bu kapsamda kullanılabilir üreten yapay zeka araçları, öğrencilerin bilgiyi sentezlemelerine, fikir üretmelerine, araştırmalarındaki hızı, kaliteyi ve verimliliği ciddi bir şekilde arttırmaya yardımcı olacaktır (Berg, 2023). Yükseköğretim kurumları, öğrencileri Üretken Yapay Zeka (GenAI) teknolojilerinin yaygın olduğu bir geleceğe daha iyi hazırlamak için eğitimde söz sahibi kurum ve kuruluşlar politikalarını, müfredatlarını ve öğretim yaklaşımlarını gözden geçirmeli ve bu konuda yapabileceklerini düşünmelidir (Chan & Hu, 2023).

## 5. Sonuç

Danışmanlar ve eğitimciler, öğrencilerin beklentilerine, kişisel becerilerine ve özelliklerine uygun eğitimlere ulaşmalarını sağlayabilir ve onların geleceği için kariyer planları oluşturabilir. Ancak, bir danışman ya da eğitimcinin zamanı, yoğun bir şekilde çalışsa bile sınırlıdır. Bu da danışmanın mentörlük yapabileceği öğrenci sayısını kısıtlar. Her an ulaşılabilir olması mümkün olmadığı gibi, çok sayıda farklı konuda uzmanlık seviyesinde bilgiye sahip olması ve anında yanıt verebilmesi de zordur. Eğitimde fırsat eşitliği açısından bakıldığında her öğrencinin aldığı danışmanlık kalitesinin eşit olup olmadığı da başka bir araştırma konusu olabilir. Bu sorunların çözümü için doğru bir şekilde tasarlanmış ve eğitilmiş eğitim sohbet robotlarına ihtiyaç duyulmaktadır. Eğitimde sohbet robotlarının geliştirilmesi ve kullanımı oldukça yeni bir alan olmasına rağmen, son iki yılda bu konuya yönelik çalışmalar beklenildiği üzere artış göstermiştir. Bu alandaki WOS yayınlarının büyük ölçüde eğitim kategorilerinde ve sağlıkla ilgili dergilerde yayımlandığı görülmektedir.

Mevcut araştırmalar, eğitim sohbet robotlarının verdiği yanıtların zaman zaman halüsinasyonlar içerebildiğini ve bu nedenle bu çıktılara eleştirel bir bakış açısıyla yaklaşılması gerektiğini vurgulamaktadır. Ancak, bu problemlerin aşıldığı ve sohbet robotlarının yaygın bir şekilde kullanıldığı yakın gelecekte, bu teknolojiler eğitimin kalitesini önemli ölçüde artıracaktır. Geleceğin dünyasında güçlü bir konum elde edebilmek için bu alandaki çalışmalara ivedilikle önem verilmesi gerekmektedir.

## Kaynaklar

- Alkaissi, H., & McFarlane, S. I. (2023). Artificial hallucinations in ChatGPT: implications in scientific writing. *Cureus*, 15(2).
- Alleyne, S. D., Horner, M. S., Walter, G., Fleisher, S. H., Arzubi, E., & Martin, A. (2009). Mentors' perspectives on group mentorship: a descriptive study of two programs in child and adolescent psychiatry. *Academic Psychiatry*, 33, 377-382.
- Berg, C. (2023). The case for generative AI in scholarly practice. *Available at SSRN 4407587*.
- Benson, C. A., Morahan, P. S., Sachdeva, A. K., & Richman, R. C. (2002). Effective faculty preceptoring and mentoring during reorganization of an academic medical center. *Medical teacher*, 24(5), 550-557.
- Brustenga, G. G., Alpiste, M. F., & Castells, N. M. (2018). Briefing paper: Chatbots in education. *Open University of Catalonia(UOC)*, 1(1), 1-33.
- Chan, C. K. Y., & Hu, W. (2023). Students' voices on generative AI: Perceptions, benefits, and challenges in higher education. *International Journal of Educational Technology in Higher Education*, 20(1), 43.
- Chang, Y., Wang, X., Wang, J., Wu, Y., Yang, L., Zhu, K., ... & Xie, X. (2024). A survey on evaluation of large language models. *ACM Transactions on Intelligent Systems and Technology*, 15(3), 1-45.
- Cooper, G. (2023). Examining science education in ChatGPT: An exploratory study of generative artificial intelligence. *Journal of Science Education and Technology*, 32(3), 444-452.
- Fryer, L. K., Ainley, M., Thompson, A., Gibson, A., & Sherlock, Z. (2017). Stimulating and sustaining interest in a language course: An experimental comparison of Chatbot and Human task partners. *Computers in human behavior*, 75, 461-468.
- Girves, J. E., Zepeda, Y., & Gwathmey, J. K. (2005). Mentoring in a post-affirmative action world. *Journal of Social Issues*, 61(3), 449-479.
- Griol, D., Molina, J. M., & Callejas, Z. (2017). Incorporating android conversational agents in m-learning apps. *Expert systems*, 34(4), e12156.
- Hobert, S. (2019). How are you, chatbot? evaluating chatbots in educational settings—results of a literature review.
- Huang, L., Yu, W., Ma, W., Zhong, W., Feng, Z., Wang, H., ... & Liu, T. (2023). A survey on hallucination in large language models: Principles, taxonomy, challenges, and open questions. *arXiv preprint arXiv:2311.05232*.
- Kashiwagi, D. T., Varkey, P., & Cook, D. A. (2013). Mentoring programs for physicians in academic medicine: a systematic review. *Academic Medicine*, 88(7), 1029-1037.

- Knight, P., & Yorke, M. (2004). *Learning, curriculum and employability in higher education*. London: RoutledgeFalmer.
- Mckie, I. A. S., & Narayan, B. (2019). Enhancing the academic library experience with chatbots: An exploration of research and implications for practice. *Journal of the Australian Library and Information Association*, 68(3), 268-277.
- Moreno, I., Gutiérrez, Y., & Montoyo, A. (2019). Atención automatizada a estudiantes en el proceso de matriculación en la Universidad de Alicante.
- Moss, J., Teshima, J., & Leszcz, M. (2008). Peer group mentoring of junior faculty. *Academic Psychiatry*, 32, 230-235.
- Neumann, A. T., Arndt, T., Köbis, L., Meissner, R., Martin, A., de Lange, P., ... & Wollersheim, H. W. (2021). Chatbots as a tool to scale mentoring processes: Individually supporting self-study in higher education. *Frontiers in artificial intelligence*, 4, 668220.
- Ozmen, B. B., & Schwarz, G. S. (2024). Future of artificial intelligence in plastic surgery: Toward the development of specialty-specific large language models. *Journal of Plastic, Reconstructive & Aesthetic Surgery*, 93, 70-71.
- Özkol, İ., Doğan, K., & Köseali, G. (2019). EBYS Uygulamalarında Yapay Zekâ Destekli Chatbot (Sohbet Robotu) Kullanımı. *Yalçınkaya B. (Editör), Ünal MA (Editör), Yılmaz B. (Editör), Özdemirci F. (Editör) Bilgi Yönetimi ve Bilgi Güvenliği*, 229-250.
- Palasundram, K., Sharef, N. M., Nasharuddin, N., Kasmiran, K., & Azman, A. (2019). Sequence to sequence model performance for education chatbot. *International journal of emerging Technologies in Learning (iJET)*, 14(24), 56-68.
- Parisi, G. I., Kemker, R., Part, J. L., Kanan, C., & Wermter, S. (2019). Continual lifelong learning with neural networks: A review. *Neural networks*, 113, 54-71.
- Pérez, J. Q., Daradoumis, T., & Puig, J. M. M. (2020). Rediscovering the use of chatbots in education: A systematic literature review. *Computer Applications in Engineering Education*, 28(6), 1549-1565.
- Pololi, L., & Knight, S. (2005). Mentoring faculty in academic medicine: a new paradigm?. *Journal of general internal medicine*, 20, 866-870.
- Sharma, V., Goyal, M., & Malik, D. (2017). An intelligent behaviour shown by chatbot system. *International Journal of New Technology and Research*, 3(4), 263312.
- Şener, G. (2018). Türkiye’de yaşanan eğitim sorunlarına güncel bir bakış. *Milli Eğitim Dergisi*, 47(218), 187-200.
- Tlili, A., Shehata, B., Adarkwah, M. A., Bozkurt, A., Hickey, D. T., Huang, R., & Agyemang, B. (2023). What if the devil is my guardian angel:

- ChatGPT as a case study of using chatbots in education. *Smart learning environments*, 10(1), 15.
- Wang, L., Ma, C., Feng, X., Zhang, Z., Yang, H., Zhang, J., ... & Wen, J. (2024). A survey on large language model based autonomous agents. *Frontiers of Computer Science*, 18(6), 186345.
- Weizenbaum, J. (1966). ELIZA—a computer program for the study of natural language communication between man and machine. *Communications of the ACM*, 9(1), 36-45.
- Winkler, R., and Soellner, M. (2018). Unleashing the Potential of Chatbots in Education: A State-Of-The-Art Analysis. in Academy of Management Annual Meeting Proceedings 2018 2018 (1),
- Wollny, S., Schneider, J., Di Mitri, D., Weidlich, J., Rittberger, M., & Drachler, H. (2021). Are we there yet?-a systematic literature review on chatbots in education. *Frontiers in artificial intelligence*, 4, 654924.
- Yu, H. (2023). Reflection on whether Chat GPT should be banned by academia from the perspective of education and teaching. *Frontiers in Psychology*, 14, 1181712.

