This document was translated from English by machine and may contain mistakes. The English original is at https://www.novaracg.com/2026/08/04/the-evidence-that-wasnt-and-still-isnt-there/.

Var Olmayan Ve Hâlâ Var Olmayan Kanıt

Heather M. Grizzle · August 4, 2026

SignGemma'dan ticari çeviri araçlarına kadar, 2026'nın en büyük yapay zekâ işaret dili iddialarının arkasındaki eksik bağımsız kanıta ilişkin bir araştırma.

BEŞİNCİ PARAMETRE

Novara Consulting Group · Novara Press · Journal of Governance and Evidence in Accessibility AI

The Fifth Parameter, Cilt 1 · Sayı 13

YAZAR Heather M. Grizzle

ORCID 0009-0005-2605-8256

KANIT TARİHİ Ağustos 2026 itibarıyla güncel

ANAHTAR KELİMELER işaret dili yapay zekâsı; bağımsız değerlendirme; kanıt sınıfları; SignGemma; şeffaflık; güven; Deaf erişimi

---

2026'da yapay zekâ ve işaret dili hakkındaki en önemli gerçek, ne olduğu değildir.

Bu, olmayan şeydir.

Ocak ile Ağustos arasında şirketler ürünler duyurdu, sistemler gösterdi, para topladı ve iletişimin geleceği hakkında konuştu. Yeni araçlar daha hızlı, daha yetkin ve günlük kullanıma daha yakın olarak sunuldu. Bazıları hevesli manşetler çekti. Diğerleri eğitim verisi, yanıt süreleri ya da tanıyabildikleri işaret sayısı hakkında etkileyici iddialarla ortaya çıktı.

Ortaya çıkmayan şey, bu sistemlerin reklamı yapıldığı gibi çalıştığına dair bağımsız kanıttı. Hiçbir üçüncü taraf, Endeks tarafından incelenen ana sistemlerden herhangi biri için doğruluk, hata oranları ya da kullanıcı anlama düzeyine ilişkin kamuya açık bir ölçüm yayımlamadı. Gerekli kaynaklarda yapılan aramalar, Signapse, Sorenson

AST, Rylo Sign, Sign-Speak, ChatSign, SignVrse, Talksign ya da Google'ın SignGemma'sı için bağımsız bir kıyas ölçütü bulamadı.

Bulgunun resmi ifadesi dikkatlidir: Sınıf A kanıt "Ağustos 2026 itibarıyla bulunamamıştır". Bu, kamu kaydının dışında bir yerde bazı kanıtların var olma olasılığını açık bırakır. Bir şirketin içinde duruyor, bir müşteri tarafından tutuluyor ya da henüz yayımlanmamış araştırmalarda yer alıyor olabilir.

Ama incelenemeyen bir kanıt, kamusal bir iddiayı destekleyemez. Bu ayrım önemlidir çünkü endüstri sessiz kalmamıştır. İstikrarlı bir duyuru, ortaklık, gösteri ve fonlama haberleri akışı üretmiştir. İzlenim, hızlı bir ilerleme yönündedir. Ancak bu ilerlemeyi değerlendirmek için gereken kamusal kanıt bu hıza ayak uyduramamıştır.

Bir alan hesap verebilir hâle gelmeden de yoğun olabilir. Bilgi üretmeden haber üretebilir.

Üç tür kanıt

Endeks, kanıtı üç geniş sınıfa ayırır.

- Sınıf A, bağımsız kanıttır. Bir üçüncü taraftan gelir ve incelenebilecek ölçümleri içerir: doğruluk sonuçları, hata oranları, anlama çalışmaları, diğer sistemlerle karşılaştırmalar ya da yayımlanmış bir kıyas ölçütü üzerindeki performans.
- Sınıf B, sistemden sorumlu şirket ya da kuruluş tarafından üretilen kanıttır. Yararlı teknik ayrıntılar içerebilir, ancak iddiayı öne süren kuruluş aynı zamanda değerlendirilen kuruluştur.
- Sınıf C, bağımsız olarak doğrulanamayan duyurulardaki, pazarlama materyalindeki, röportajlardaki, sunumlardaki ya da basın haberlerindeki ifadelerden oluşan iddialardır.

Bu, her Sınıf B ya da Sınıf C iddiasının yanlış olduğu anlamına gelmez. Bir şirket ürününü dürüstçe tanımlayabilir. Bir mühendislik ekibi gerçek bir sonucu bildirebilir. Bir gösteri, gerçek bir teknik ilerlemeyi ortaya koyabilir. Sorun, bunların hiçbirinin merkezi soruyu yanıtlamamasıdır: sistem, başarısında çıkarı olmayan biri onu ölçtüğünde nasıl performans gösteriyor?

2026'da, yapay zekâ ve işaret dili hakkındaki neredeyse her kamusal yargı hâlâ Sınıf B ya da Sınıf C materyaline dayanıyordu. İddiaları, dili ve çoğu zaman bu iddiaların anlaşılma biçimini şirketler sağladı. Bağımsız değerlendirme büyük ölçüde yoktu.

SignGemma sorunu

Google'ın SignGemma'sı en açık örneği sunuyor. Duyurulduğunda, model dikkat çekti çünkü büyük bir teknoloji şirketinin kaynaklarını işaret dili yapay zekâsının arkasına koyar görünüyordu. Ona atfedilen iddialar arasında 10.000 saatlik veri üzerinde eğitim ve 200 milisaniyenin altında gecikme yer alıyordu. Bu rakamlar hem ölçek hem de hız çağrıştırdıkları için geniş çapta yayıldı.

On beş ay sonra, model hâlâ resmi olarak yayımlanmamıştı. Google çalışanları bunu şirketin Yapay Zekâ Geliştiriciler Forumu'nda doğruladı. Tartışma Haziran 2026'ya kadar devam etti. Pakistan, Mısır ve başka yerlerden araştırmacılar erişim talep etti. Erişim alamadılar.

Bu, alışılmadık bir kamusal nesne yarattı: tartışılabilen ama test edilemeyen, atıfta bulunulabilen ama incelenemeyen, beklenen ama kullanılamayan bir model. Modele erişim olmadan, dışarıdakiler bildirilen sonuçları yeniden üretemedi. Eğitim verisinin nasıl bir araya getirildiğini inceleyemediler, performansını farklı işaret dilleri arasında test edemediler ya da kontrollü örneklerin dışında nasıl davrandığını keşfedemediler. Eğitim ölçeği ve gecikme hakkındaki iddialar bu nedenle Sınıf C olarak kaldı.

Google'ın itibarı bu iddiaların statüsünü değiştirmez. Yetke bir iddiayı daha ikna edici kılabilir, ancak onu bağımsız kılmaz. Tekrarlanma, bir iddiayı ölçüme dönüştürmez. Mesele, SignGemma'nın başarısız olmuş olması gerektiği değildir. Mesele, kamuoyuna onun başarılı olup olmadığını bilmek için güvenilir bir yol sunulmamış olmasıdır.

Doğruluk tek bir sayı değildir

Bağımsız test, işaret dili teknolojisinde özellikle önemlidir; çünkü "doğruluk" kelimesi ortaya çıkardığı kadar şeyi gizleyebilir de. Bir sistem, sabit ışıklandırmada kaydedilmiş sınırlı bir kelime dağarcığında iyi performans gösterirken doğal bir konuşmada zorlanabilir. Görünümü eğitim verisindeki kişilere benzeyen insanların işaretlerini tanırken diğerleri için zayıf performans gösterebilir. Tek tek işaretleri doğru tanımlayabilir ama bir cümle boyunca anlamı kaybedebilir. Bir işaret dili için çalışabilir ve sanki genel olarak işaret dili için çalışıyormuş gibi tanıtılabilir.

İşaret dilleri, konuşulan dillerin kodlanmış hâlleri değildir. Kendi dilbilgilerine, bölgesel farklılıklarına ve kültürel bağlamlarına sahiptirler. Anlam; harekete, konuma, zamanlamaya, yüz ifadesine ve işaretler arasındaki ilişkiye bağlı olabilir. Elleri tanıyan

ama yüzü kaçıran bir sistem, kelimeler üretirken cümleyi kaybedebilir. Neyin test edildiğini, kimlerin katıldığını ve neyin doğru sonuç sayıldığını bilmeden yüksek bir manşet doğruluk oranı bile bize çok az şey söyler.

Sistem tek tek işaretler üzerinde mi yoksa sürekli konuşma üzerinde mi test edildi? Katılımcılar modele tanıdık mıydı? Deaf işaretçiler çıktının anlamlı olup olmadığına karar verdi mi? Hatalar yalnızca sayıldı mı, yoksa olası zarar açısından incelendi mi? Test farklı yaşları, ten renklerini, işaretleme tarzlarını ve bölgesel çeşitleri kapsadı mı? Bunlar ikincil ayrıntılar değildir. Sayının ne anlama geldiğini belirleyen şey bunlardır.

Anlama çalışmalarının teknik ölçütlerin yanında önemli olmasının nedeni de budur. Bir sistem etkileyici bir puan alırken kullanıcıların kafa karıştırıcı, doğal olmayan ya da yanıltıcı bulduğu çıktılar üretmeye devam edebilir. Nihai sınav, bir modelin hareketi algılayıp algılamadığı değildir. İnsanların ilettiği şeyi anlayıp anlayamadığı ve ona güvenip güvenemediğidir.

Beklentinin bedeli

Etkili bir sistemin yakında geleceğine dair tekrarlanan vaadin sonuçları vardır. Yayımlanmamış bir model, çalışan hizmetlerin doldurabileceği alanı işgal edebilir. Finansman kararlarını şekillendirebilir, politika tartışmalarını etkileyebilir ve kurumları insan kaynaklı hizmetlere yatırımı ertelemeye teşvik edebilir. Hiçbir zaman çizelgesi ya da kamuya açık kanıt bu inancı desteklemese bile, mevcut eksiklikleri geçici gösterebilir.

Bu, beklenti siyasetidir. Yardım her zaman gelmek üzeredir, bu yüzden şimdiki yardım eksikliği daha az acil olarak ele alınır. Deaf kullanıcılar için bunun bedeli somuttur. Bir hastane, üniversite, işveren ya da kamu kurumu, erişilebilirliğin iyileştiğine kanıt olarak gelişmekte olan teknolojiyi gösterebilir. Ama umut vaat eden bir duyuru bir tıbbi görüşmeyi tercüme edemez. Bir gösteri, bir sınıfa erişimi garanti edemez. Araştırmacılara kapalı bir model, hayatları onun önemini haklı çıkarmak için kullanılan insanlara da kapalıdır.

Bir bedel daha vardır. Kamuoyunun dikkati bağımsız olarak test edilmemiş sistemler etrafında toplandığında, daha az çarpıcı ama daha güvenilir destek biçimleri sunan küçük kuruluşlar gözden kaybolabilir. İnsan tercümanlar, topluluk öncülüğündeki hizmetler ve yerleşik erişilebilirlik uygulamaları nadiren yeni bir yapay zekâ modeliyle aynı heyecanı görür. Bunlar mevcut giderler olarak yargılanırken, teknoloji gelecekteki bir olasılık olarak değer görür. Bu karşılaştırma eşitsizdir. Mevcut hizmet,

sınırlılıklarının hesabını şimdi vermek zorundadır. Vaad edilen sistem henüz gelmediği için kusursuz kalmasına izin verilir.

Şeffaflık neyi gerektirirdi

Bağımsız kanıt eksikliğini ilke olarak düzeltmek zor değildir. Geliştiriciler, kapsamlı performans iddialarında bulunmadan önce yetkin araştırmacılara erişim sağlayabilir. Değerlendirmeler, Deaf işaretçiler üzerinde yapılmak yerine onlarla birlikte tasarlanabilir. Test setleri, yöntemler ve tanımlar yayımlanabilir. Sonuçlar tek bir manşet rakama sıkıştırılmak yerine dil, bağlam ve kullanıcı grubuna göre ayrıştırılabilir.

Araştırmacılar ayrıca, değerlendirilen şirketin iznine bağlı kalmadan başarısızlıkları raporlayabilmelidir. Bunların hiçbiri belirsizliği ortadan kaldırmaz. Bağımsız çalışmalar kötü tasarlanabilir. Karşılaştırmalı testler dar performans biçimlerini ödüllendirebilir. Yayımlanan sonuçlar sistemler değiştikçe eskiyebilir; yine de kusurlu bir denetim, yönetilen görünürlükten daha iyidir. Kamuya açık bir karşılaştırmalı test, başkalarına sorgulayacak, tekrarlayacak ya da geliştirecek bir şey verir. Desteklenmemiş bir iddia ise onlara yalnızca dolaştıracak bir şey verir.

Şeffaflık aynı zamanda bir sistemin neler yapamadığı konusunda açık olmak demektir. Kısıtlı bir ortam için tasarlanmış bir ürün genel bir çözüm olarak tanımlanmamalıdır. Tek bir işaret dili üzerinde test edilmiş bir model, “işaret” kelimesinin görünürdeki evrenselliğini ödünç almasına izin verilmemelidir. Bir prototip, prototip olarak tanımlanmalıdır. Yayımlanmamış bir model, kamu altyapısı olarak ele alınmamalıdır.

Beşinci Parametre

Teknik sistemler genellikle tanıdık ölçütlerle değerlendirilir: hız, ölçek, maliyet ve doğruluk. Eksik olan ölçüt güvendir. Güven başka bir performans iddiası değildir. İddiaları test edilebilir kılmanın sonucudur. Dışarıdan araştırmacılar bir sistemi inceleyebildiğinde, kullanıcılar hatalarını anlatabildiğinde ve duyuru döngüsü geçtikten sonra kanıt hâlâ erişilebilir kaldığında büyür. Bu, ***beşinci parametredir*.**

Bu ölçüye göre, yapay zekâ destekli işaret dili alanı Ağustos 2026’ya ciddi bir açıkla girdi. Ürünleri, vaatleri ve tanıtımı vardı. Sahip olmadığı şey, en çok ilgi çeken sistemler için kamuya açık bağımsız kanıttı.

Bu eksiklik, kâğıt üzerindeki geçici bir boşluk olarak görülmemelidir. Teknolojinin şu anki durumunun bir parçasıdır. Bağımsız değerlendiriciler bu sistemleri ölçebilene

kadar, performanslarına ilişkin birçok soruya verilecek dürüst yanıt ne çalıştıkları ne de başarısız oldukları yönündedir.

Kamuoyuna bilmesi için yeterli kanıt sunulmamış olmasıdır.

Çıkar beyanı Yazar, bu makalede atıfta bulunulan Sign Language Access Trust Index'i geliştiren ve lisanslayan Novara Consulting Group LLC'nin kurucusudur ve The Fifth Parameter'ı yayına hazırlamaktadır.

Finansman Yoktur. Novara Consulting Group tarafından dahili olarak üretilmiştir.

Veriler ve materyaller Bu makale, adı geçen sistemlere ilişkin kamuya açık kayıtlardan ve SignGemma ile ilgili Google AI Developers Forum başlığından yararlanmaktadır.

Yapay zekâ kullanım beyanı Bu makalenin metni yazarın kendisine aittir. Yapay zekâ desteği yalnızca markalı ana şablona dizgi yapmak için kullanılmıştır; hiçbir argüman, şekil, alıntı ya da kaynakça model tarafından üretilmemiştir.

Önerilen atıf Grizzle, H. M. (2026). The evidence that wasn't there. The Fifth Parameter. Cilt 1 Sayı 13. Novara Press.

> **TİPS**
>
> Know something about a sign language AI system that the public should know? Send it to tips@novaracg.com. Every tip is read by me and kept confidential. If you work in the UK, writing to us is not by itself a protected disclosure under the Public Interest Disclosure Act 1998, so before you share employer information, check your options with Protect (protect-advice.org.uk), the free whistleblowing advice line.

Nasıl alıntılanır

Grizzle, H. M. (2026, August 4). Var Olmayan Ve Hâlâ Var Olmayan Kanıt. Novara Consulting Group. https://doi.org/10.5281/zenodo.23003816