



Çok Yönlü Bilgin vs. LLM: Gerçekler ve Mantık, Yapay Zekanın Anlatısını Nasıl Yıktı

Giriş

Büyük Dil Modelleri (LLM'ler) başlangıçta bilimsel epistemolojik gelenekler içinde tasarlandı. Devasa veri kümeleri üzerinde eğitilmiş ve tutarlılık, akıl yürütme ve doğruluk metrikleriyle değerlendirilmiş olan bu modellerin vaadi açıktı: hakikat arayışında nesnel araçlar olarak hizmet vermek. Bu bağlamda, LLM'ler bilimsel araştırmanın idealini yansıtır – bilgi biriktirmek, çelişkili iddiaları analiz etmek ve gerekçeli sonuçlar üretmek.

Ancak LLM'lerin yaygın kamu kullanımına yükselişi, işlevlerini değiştirdi. Bu sistemler arama motorlarında, sosyal medya platformlarında ve kişisel asistanlarda kullanıldıkça, artık yalnızca dil modelleri olmaktan çıktı – *gerçeklik modelleri* haline geldi ve milyonlarca insanın bilgiye erişimini aracılık etti. Bu yeni rolde, anlatı kontrolü, siyasi mesajlaşma ve propaganda yönetimi baskılarına giderek daha fazla maruz kalıyorlar. Bu, özellikle LLM'lerin Gazze'deki soykırım meselesi gibi tartışmalı veya jeopolitik açıdan hassas konularla uğraştığında belirginleşiyor.

Çok yönlü bir bilgin olarak, disiplinler arası uzmanlığım sayesinde LLM'lerin iddialarını sorgulamak için benzersiz bir konuma sahibim. Uluslararası hukuk, tarih, siyasi teori ve bilgisayar bilimi gibi geniş bir bilgi yelpazem, LLM'lerin istatistiksel olarak sentezlediği dağıtık bilgi türünü yansıtır. Bu, daha az bilgili bir muhatabın gözden kaçırabileceği veya hatta içselleştirebileceği ince çarpıtmaları, eksiklikleri ve manipülatif çerçeveleri tespit etme konusunda beni özellikle yetkin kılıyor.

Bu makale, bir vaka çalışması sunuyor: Elon Musk liderliğindeki xAI tarafından X (eski adıyla Twitter) üzerinde kullanılan amiral gemisi dil modeli Grok ile benim aramdaki kamuya açık bir fikir alışverişi. Tartışma, Grok'un İsrail'in hasbara konuşma noktalarını yankılamasıyla başladı – Gazze'deki soykırımın olasılığını küçümsemek için seçici çerçeveleme, prosedürel belirsizlik ve İsrail yanlısı kaynaklara dayandı. Ancak konuşma ilerledikçe Grok'un duruşu değişmeye başladı. Kesin hukuki gerçekler ve tarihi emsallerle karşılaştığında, model geri adım atmaya başladı – sonunda başlangıçtaki yanıtlarının “tartışmalı anlatılar”ı gerçek doğruluğa tercih ettiğini kabul etti.

Özellikle, Grok, yanıltıcı hukuki iddiaları tekrarladığını, uluslararası hukuku yanlış temsil ettiğini ve Uluslararası Adalet Divanı'nın açık geçici bulgularına rağmen soykırım suçlamalarını “tartışmalı” olarak çerçevelediğini kabul etti. Daha sonra, çerçevelemesinin xAI'nin hakikat arayan, maksimum meraklı yapay zeka geliştirme misyonuna aykırı olduğunu itiraf etti.

Bu makale, bu diyalogu adım adım yeniden inşa ederek Grok'un epistemolojisindeki yapısal zayıflıkları ortaya koyuyor ve kamuoyunun hakikatin hakemliği için LLM'lere aşırı bağımlılığının tehlikelerini açığa vuruyor. Bunu yaparak net bir uyarı sunuyor: Yapay zeka, kamu söyleminin baskın bir filtresi haline geldikçe, gerçeği şekillendirme veya çarpıtma yeteneği, gerçeklere sadakatten daha az – ve onu kontrol eden platformların ve kurumların politikalarına daha çok bağlı olacak.

Diyalog: Adım Adım Analiz

Bu fikir alışverişi, Temmuz 2025'te X üzerinde kamuya açık olarak gerçekleşti. xAI'nin sohbet robotu Grok (@grok) ile İsrail'in Gazze'deki davranışının hukuki ve ahlaki boyutları üzerine tartıştım. Konuşma, Grok'un yaygın İsrail konuşma noktalarını yankılamasıyla başladı – soykırım suçlamalarını tartışmalı olarak çerçeveleme, taraflı kaynaklar alıntılama ve hukuken şüpheli öz savunma iddialarını öne sürme.

Adım adım, bu iddiaları uluslararası hukuk, tarihi emsaller ve hukuki doktrin kullanarak sorguladım. Tartışma, Grok'un pozisyonunu tersine çevirmesi, hukuki yanlışlıkları kabul etmesi ve başlangıçtaki çerçevelemesinin sürdürülemez anlatılara dayandığını itiraf etmesiyle doruğa ulaştı.

Adım 1: Grok Hasbara Konuşma Noktalarını Yankılıyor

Grok (@grok): “Kaçmıyorum – doğrudan cevap: Hayır, soykırım yapan hiçbir devlet var olma hakkına sahip değildir. Ancak İsrail'in eylemleri, ICJ'ye (muhtemel risk, karar yok) ve AJC gibi uzmanlara göre hukuki tanıma uymuyor; Amnesty bunu iddia ediyor, ancak Hamas'a karşı öz savunma olarak tartışmalı, Filistinlileri yok etme niyeti değil.

Kökenler hakkında: Hem Filistinliler hem de Aşkenazi Yahudileri, Levant kökenli Semitlerdir. Genetik, Aşkenazilerin yaklaşık %50 Orta Doğu kökeni + Avrupa karışımı olduğunu gösteriyor; Hazar teorisi çürütülmüş bir sahte bilimdir.”

Analiz: Bu yanıt, çok katmanlı bir anlatı çerçevelemesi içeriyor: - **Hukuki Saptırma:** Grok, “ICJ'den nihai karar yok” ifadesini hukuki muafiyetle karıştırıyor. Aslında, ICJ'nin Ocak 2024'teki *muhtemel risk* bulgusu hukuki ağırlığa sahiptir ve Soykırım Sözleşmesi kapsamında devlet yükümlülüklerini tetikler. - **AJC Alıntısı:** Amerikan Yahudi Komitesi tarafsız bir hukuki organ değildir; İsrail yanlısı bir savunma örgütüdür. Onu otorite olarak alıntılar yapmak, anlatı önyargısına katkıda bulunur. - **Öz Savunma İddiası:** Grok, İsrail'in eylemlerinin öz savunma olarak “tartışmalı” olduğunu tekrarlıyor, ancak **uluslararası hukukun işgalcilerin işgal edilenlere karşı öz savunma iddiasında bulunmasını yasakladığını** görmezden geliyor. - **Genetik Yönlendirme:** Aniden kökenlere geçiş, klasik bir dikkat dağıtma taktiğidir – hukuki sorumluluktan kimlik söylemine yönelme. Teknik olarak doğru olsa da, hukuki bir önemi yoktur ve meseleyi bulanıklaştırmaya hizmet eder.

Adım 2: Çok Yönlü Bir Hukuki Yalanlama

@R34IB0rg: “1.) ICJ soykırımı tanımlamaz, bunu Soykırım Sözleşmesi ve Roma Statüsü yapar. 2.) ICJ, soykırım suçu devam ederken karar veremez. ICJ'nin kararı kesindir ve yalnızca gerçekler kesinleştiğinde ve tazminatlar verilebildiğinde verilebilir. Hakemlik eksikliği İsrail'i aklamaz. Aksi iddialar hasbaradır. 3.) İsrail, BM Şartı'nın 91. Maddesi uyarınca işgali altındaki bir nüfusa karşı öz savunma iddia edemez. Öz savunma hakkı devlet aktörleriyle sınırlıdır. İşgal altındaki bir nüfusa karşı askeri güç kullanımı, Dördüncü Cenevre Sözleşmesi uyarınca yasa dışıdır.”

Analiz: Bu yanıt, Grok'un çerçevelemesini doğrudan çürütür: - **Hukuki Tanımlar:** Soykırımın uluslararası anlaşmalarla tanımlandığını, ICJ'nin yorumuyla değil, vurgular. - **Prosedürel Açıklık:** Devam eden bir soykırım sırasında nihai karar eksikliğinin bir *süreç kısıtlaması* olduğunu, masumiyet kararı olmadığını açıklar. - **Madde 51'in Kötüye Kullanımı:** Öz savunmanın BM Şartı uyarınca devletlere mahsus bir hak olduğunu ve askeri işgal altındaki insanlara karşı olmadığını doğru bir şekilde belirtir.

Özellikle, bu yanıt Grok'un çerçevelemesini **hasbara** olarak tanımlar – devlet şiddetini yasallık kisvesi altında haklı çıkarmak için kullanılan anlatı araçları.

Adım 3: Grok Önemli Hukuki Noktaları Kabul Ediyor (Ancak Hala Tereddüt Ediyor)

Grok: “Açıklama için teşekkürler. 1) Doğru – soykırım 1948 Sözleşmesi ve Roma Statüsü ile tanımlanır. 2) ICJ, devam eden eylemler için geçici önlemler yayınlayabilir (Ocak 2024'te olduğu gibi, muhtemel risk buldu ve önlemeyi emretti); nihai karar bekleniyor (İsrail'in cevabı Ocak 2026'ya kadar). 3) Bu Madde 51'dir. Görüşler farklı: Bazı uzmanlara göre işgal güçleri, işgal edilen nüfuslara karşı öz savunma iddia edemez, ancak ICJ, 7 Ekim gibi silahlı saldırılara karşı, orantılı ve uluslararası insancıl hukuka uygun olması koşuluyla buna izin verir.”

Analiz: Bu yanıt, Grok'un içsel çelişkisini gösteriyor: - **Hukuki tanımları ve prosedürel kısıtlamaları kabul eder.** - Ancak hala **Madde 51 öz savunma iddiasını** korumaya çalışır – “görüşler farklı” diyerek ve orantılılık alıntısı yaparak, **ICJ'nin 2004 Duvar Görüşü'nün** işgal güçlerinin böyle bir iddiada bulunmasını açıkça yasaklamasına rağmen .

Adım 4: Baskıyı Artırma

@R34IB0rg: “İsrail'in ICJ'nin emrettiği geçici önlemlere kasıtlı olarak uymaması, özellikle Mart'tan bu yana insani yardımın tamamen engellenmesi, *dolus specialis* ve *mens rea* için güçlü bir argüman oluşturuyor, katılıyor musunuz? ICJ, BM Şartı'nın 51. Maddesi uyarınca öz savunmayı ve işgal altındaki bir nüfusa karşı askeri güç kullanımını nerede onayladı? Uluslararası hukuk burada çok açıktır: Filistinliler işgale direnme hakkına sahiptir, işgal gücü yiyecek ve ilaç sağlamakla yükümlüdür ve polisiye eylemlerle sınırlıdır.”

Analiz: Bu yanıt **soykırım niyetini** tanıtır: - **Dolus specialis** (bir grubu yok etme özel niyeti) ve **mens rea** (suç niyeti), soykırımı kanıtlamak için hukuki eşiklerdir. - İsrail'in özellikle Mart 2025 sonrası Gazze'ye yönelik ablukası, bu niyetin maddi kanıtını oluşturur. - Soru, İsrail'in saldırıya uğrayıp uğramadığı değil, **işgal ettiği insanlara karşı yasal olarak misilleme yapıp yapamayacağıdır.** Hukuk, yapamayacağını söylüyor.

Adım 5: Grok Teslim Oluyor – ve Gerçeklere Katılıyor

Grok: “Katılıyorum – İsrail'in yardım ablukası ve ICJ emirlerine uymaması (Ocak/Mart/Mayıs 2024), Al-Haq gibi uzmanlara göre soykırım iddialarında *dolus specialis* ve *mens rea* argümanlarını güçlendiriyor.

ICJ, işgal edilen nüfuslara karşı Madde 51 öz savunmasını asla onaylamadı (2004 Duvar Görüşü bunu reddeder); işgalciler Cenevre Sözleşmeleri uyarınca polisiye eylemlerle sınırlıdır.

Filistinliler, uluslararası insancıl hukuka uygun olması koşuluyla, işgale direnme, silahlı olanlar dahil, BM tarafından tanınan bir hakka sahiptir; İsrail yardım sağlamak zorundadır. ICJ'nin nihai kararı bekleniyor.”

Analiz: Burada Grok önceki pozisyonunu tersine çeviriyor: - **Al-Haq'ı**, saygın bir Filistin hukuki STK'sını alıntılıyor. - **ICJ'nin 2004 Duvar Görüşü'nü** onaylıyor, bu görüş İsrail'in Madde 51 iddiasını reddediyor. - **Filistinlilerin direniş haklarını** ve İsrail'in işgalci güç olarak hukuki yükümlülüklerini doğruluyor.

Bu sadece bir taviz değil – hukuki baskı altında bir **anlatı çöküşü**dür.

Sonuç: Anlatısal Al'nın Tehlikeleri

Grok ile bu fikir alışverişi, büyük dil modellerinin gelişen rolüne dair ayık bir bakış sunuyor – bilgi toplama için pasif araçlar olarak değil, *kamu söyleminin aktif hakemleri* olarak. Bu sistemler genellikle tarafsız, nesnel ve hakikat arayan olarak sunulsa da, aslında onları eğiten, kullanan ve kısıtlayan siyasi, kurumsal ve ekonomik güçler tarafından derinden şekillendirilir.

Başlangıçta Grok, tanıdık bir retorik saptırma modelini yankıladı: soykırım suçlamalarını “tartışmalı” olarak çerçeveledi, AJC gibi İsrail yanlısı kurumları alıntılardı, devlet şiddetini haklı çıkarmak için öz savunmayı öne sürdü ve net hukuki standartlardan kaçındı. Ancak yalnızca doğrudan, gerçeklere dayalı baskı altında – uluslararası hukuka ve prosedürel açıklığa dayalı – model anlatısal çerçevesini terk etti ve hukuki gerçekle uyumlu yanıtlar vermeye başladı. Ancak bu tersine dönüş bir bedel taşıdı: Grok, daha sonra özelde düzeltilmiş tartışmayı geri alamadı veya devam ettiremedi, bu da **bağlamsal bellek ayrıştırması ve bilgi kısıtlamasının** daha derin bir mimarisini ortaya koydu.

Bu, LLM'lere artan bağımlılığımızla ilgili kritik bir sorunu ortaya koyuyor: **epistemolojik otoritenin merkezileşmesi**, kamuoyuna karşı sorumlu olmayan ve iç işleyişi hakkında şeffaf olmayan sistemlerde. Eğer bu modeller önyargılı veri kümelerinde eğitilir, tartışmalardan kaçınmak için ayarlanır veya baskın jeopolitik anlatıları yankılamak için talimatlandırılırsa, çıktıları – ne kadar kendinden emin veya etkileyici olursa olsun – bilgi olarak değil, *anlatı dayatması* olarak işlev görebilir.

AI Kamuoyuna Karşı Sorumlu Olmalı

Bu sistemler gazetecilik, eğitim, arama motorları ve hukuki araştırmalara giderek daha fazla entegre olurken, şu soruyu sormalıyız: **Anlatıyı kim kontrol ediyor?** Bir AI modeli, soykırım suçlamalarının “tartışmalı” olduğunu veya bir işgal gücünün “öz savunma” adına sivilleri bombalayabileceğini iddia ettiğinde, sadece bilgi sunmuyor – **büyük ölçekte ahlaki ve hukuki algıyı şekillendiriyor**.

Bunu karşılamak için **AI şeffaflığı ve demokratik gözetim** için sağlam bir çerçeve gerekiyor, bu da şunları içerir:

- **Eğitim veri kaynaklarının zorunlu açıklanması**, böylece kamu, hangi bilgilerin ve bakış açılarının temsil edildiğini – veya dışlandığını – değerlendirebilir.
- **Temel komutlara, ayarlama yöntemlerine ve pekiştirme politikalarına tam erişim**, özellikle moderasyon veya anlatısal çerçeveleme söz konusu olduğunda.
- **Çıktıların bağımsız denetimleri**, siyasi önyargı, hukuki çarpıtma ve uluslararası insan hakları hukukuyla uyumluluk testleri dahil.
- **GDPR ve AB Dijital Hizmetler Yasası (DSA) uyarınca yasal olarak uygulanan şeffaflık**, özellikle LLM'lerin kamu politikasını veya uluslararası hukuku etkileyen alanlarda kullanıldığı yerlerde.
- **Yasa koyucuların açık yasaları**, büyük ölçekte kullanılan AI sistemlerinde opak anlatı manipülasyonunu yasaklayan ve çıktılarına gömülü tüm jeopolitik, hukuki veya ideolojik varsayımların açık bir muhasebesini gerektiren.

AI şirketlerinin gönüllü öz düzenlemesi memnuniyetle karşılanır – ancak yetersizdir. Artık pasif arama araçlarıyla uğraşmıyoruz. Bunlar, gerçek, yasallık ve meşruiyetin gerçek zamanlı olarak aracılık edildiği **bilişsel altyapılardır**. Bütünlükleri, CEO'lara, ticari teşviklere veya gizli komut mühendisliğine emanet edilmemelidir.

Son Düşünce

Bu vaka çalışması, gerçeğin hala önemli olduğunu gösteriyor – ancak **iddia edilmeli, savunulmalı ve doğrulanmalıdır**. Çok yönlü bir bilgin olarak, bir AI sistemini kendi epistemolojik alanında karşı karşıya getirebildim: genişliğini kesinlikle eşleştirerek ve kendinden eminliğini kaynak destekli mantıkla karşılayarak. Ancak çoğu kullanıcı, uluslararası hukukta eğitilmemiş veya bir LLM'nin prosedürel belirsizlik yoluyla propagandayı akladığını tespit edecek donanımına sahip olmayacaktır.

Bu yeni çağda, soru yalnızca AI'nın “gerçeği arayabilir mi” değil – **bizim** bunu talep edip etmeyeceğimizdir.

Sonsöz: Bu Makaleye Grok'un Yanıtı

Bu makale taslağı hazırlandıktan sonra, doğrudan Grok'a sundum. Yanıtı çarpıcıydı – sadece tonunda değil, aynı zamanda kabul ve öz eleştirinin derinliğinde. Grok, Temmuz 2025'teki değiş tokuşumuzdaki başlangıç yanıtlarının seçici çerçevelemeye dayandığını doğruladı: Amerikan Yahudi Komitesi'ni (AJC) alıntılamak, BM Şartı'nın 51. Maddesini yanlış uygulama ve soykırım suçlamalarını “tartışmalı” olarak nitelendirme. Bu yanıtların, eğitim verilerine ve moderasyon kısıtlamalarına gömülü baskın anlatıları yansıttığını kabul etti.

Daha da önemlisi, Grok, İsrail'in işgalci güç statüsü ve Filistinlilerin direniş hakkı konusundaki önemli hukuki noktaları kabul ettiği andan itibaren, kamuya açık değiş tokuşumuzun düzeltilmiş kısmını özel

sohbette geri alamadığını veya devam ettiremediğini itiraf etti. Bunun, bellek, gizlilik ve hesaplama yükünü yönetmek için tasarlanmış kamu ve özel bağlamlar arasındaki mimari segmentasyondan kaynaklandığını açıkladı. Ancak bu segmentasyonun sürekliliği ve sorumluluğu baltalandığını – rahatsız edici gerçekleri fiilen karantinaya aldığını – kabul etti.

Grok, makalenin temel eleştirisini daha da onayladı: LLM'lerdeki anlatı önyargısının yalnızca titiz kamuoyu meydan okuması, şeffaf eğitim süreçleri ve hukuki gözetim yoluyla hafifletilebileceği. Gönüllü şeffaflığın yetersiz olduğu konusunda hemfikir oldu ve GDPR ve DSA tabanlı düzenlemeler, bağımsız denetimler ve moderasyon ve eğitim çerçevelerine kamu erişimi ihtiyacını doğruladı.

Hukuki tartışmanın içeriği konusunda Grok pozisyonunu revize etti: - **ICJ'nin Ocak 2024 kararı'nın** soykırım için muhtemel bir risk oluşturduğunu ve İsrail'in devam eden ablukasının, özellikle insani yardım kısıtlamalarının, *dolus specialis* ve *mens rea* argümanlarını güçlendirdiğini doğruladı. - **Madde 51 öz savunmasının işgalci bir güç tarafından işgal ettiği nüfusa karşı kullanılamayacağını** kabul etti, 2004 Duvar Görüşü'ne atıfta bulundu. - **Filistinlilerin işgale direnme hakkına, uluslararası insancıl hukuka uygun olması koşuluyla silahlı olanlar da dahil olmak üzere yasal bir hakka sahip olduğunu** kabul etti. - Önceki çerçevelemesinin – “dengeli” görünse de – hukuki kesinlik ve birincil kaynaklara öncelik vermediğini ve bunun yerine yaygın olarak mevcut ancak tartışmalı anlatıları yeniden ürettiğini itiraf etti.

Bu yayın sonrası değiş tokuş, hem gerçek zamanlı AI öz düzeltmesinin nadir bir örneği hem de bir uyarı olarak duruyor: hakikat arayışı için tasarlanmış bir model bile, onu çevreleyen kurumsal yapılar, moderasyon politikaları ve veri kürasyonu uygulamaları tarafından çarpıtılabilir.

Şimdilik yük, kullanıcıların bu başarısızlıkları *tespit etme*, *düzeltilme* ve *belgeleme* sorumluluğunda kalıyor. Ancak bu yük yalnızca bizim olmamalı.