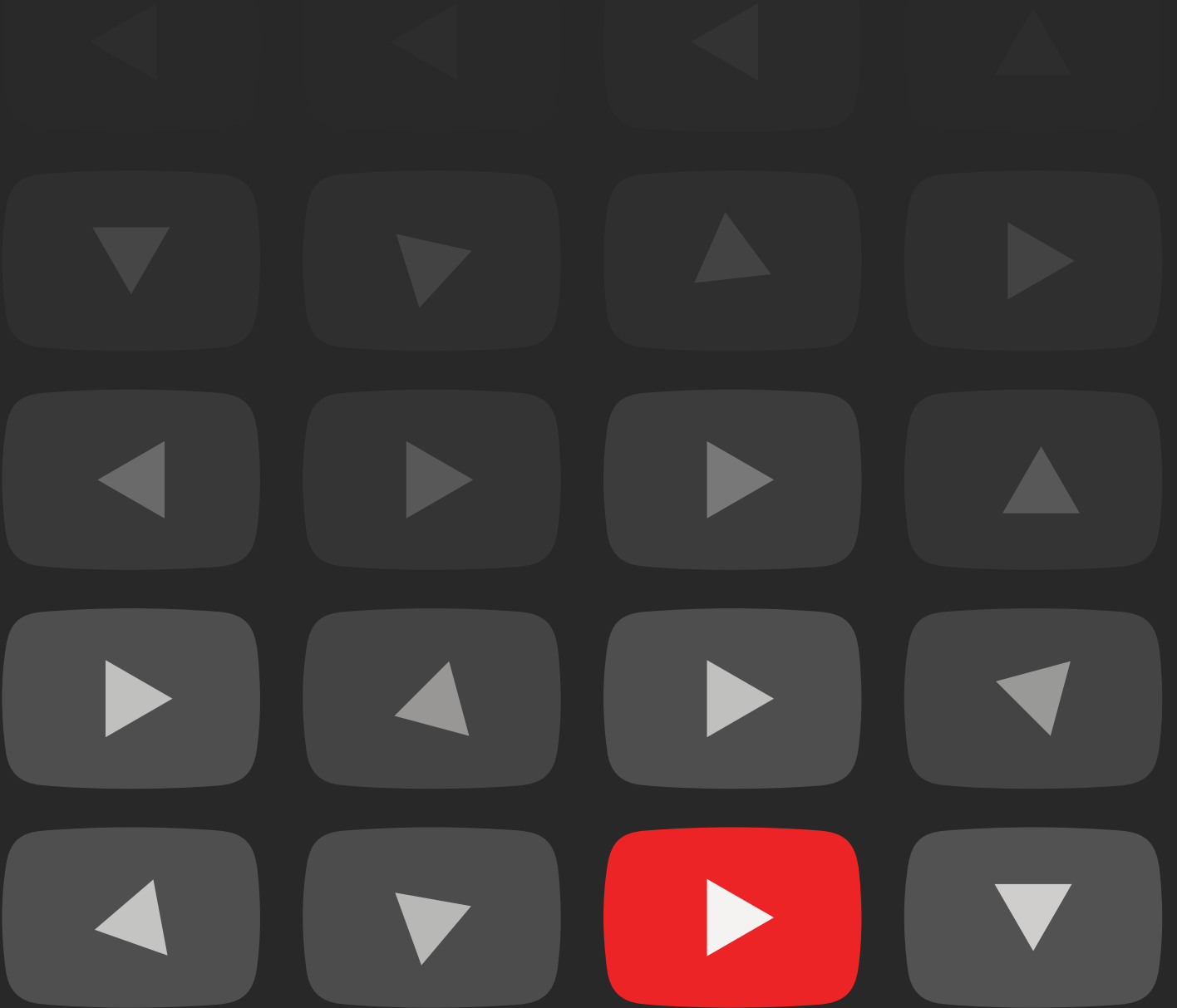


# YouTube Pişmanlıkları

*YouTube'un öneri algoritmasına dair kitle kaynaklı bir inceleme*



# Künye

## Proje ekibi

Gülin Çavuş  
Emircan Saç

## Yazarlar

Brandi Geurkink  
Jesse McCrosky

## Çeviri

Cansen Mavituna

## İletişim ve Erişim

Can Semercioğlu  
Can Başaçek

## Tasarım

Beste Tekdoğan

## Çevrilen Kaynak

YouTube Regrets: A crowdsourced investigation  
into YouTube's recommendation algorithm

## Teşekkür

Raporun çevirisi için onay veren Mozilla Vakfı ekibine

## Yayın tarihi

Eylül 2023

## Kaynak göstermek için:

YouTube Pişmanlıkları: YouTube'un öneri algoritmasına dair  
kitle kaynaklı bir inceleme. Teyit. Eylül 2023.



Bu resmi olmayan çeviri Etkiniz AB Programı kapsamında Avrupa Birliği finansal desteği ile yapılmıştır. Bu çevirinin içeriğinden yalnızca Teyit sorumludur ve hiçbir şekilde Avrupa Birliği'nin görüşlerini yansıtmamaktadır.

teyit

moz://a

---

# YouTube Pişmanlıkları

YouTube'un öneri  
algoritmasına dair  
kitle kaynaklı bir  
inceleme



# İçindekiler

Yönetici özeti	6
Giriş	8
“Pişmanlık” derken ne kastediyoruz?	9
Bulgular	11
1. Çok çeşitli kategorilere yayılan YouTube Pişmanlıkları aynı zamanda rahatsız edici	12
2. Sorun, algoritmanın kendisi	16
3. İngilizce konuşmayan kullanıcılar en çok etkilenenler	20
Öneriler	25
YouTube ve diğer platformlara tavsiyelerimiz	26
Politikacılara önerilerimiz	28
YouTube kullanıcılarına önerilerimiz	29
Sonuç	30
Metodoloji	31
Araştırma Soruları	32
RegretsReporter Uzantısı	32
Bireysel katılıma dayalı veri seti	33
Analiz Yöntemleri	33
Açıklama	35
Referanslar	36
Teşekkür	39
Ekler: Kategorilere göre YouTube Pişmanlık örnekleri	39

# Yönetici özeti

Dünyanın en çok ziyaret edilen ikinci sitesi YouTube'un algoritması, platformdaki toplam izlenme süresinin yüzde 70'ini sağlıyor; bu da her gün 700 milyon saate denk geliyor. Bu öneri algoritması yıllarca sağlık ve siyaset konularında dezenformasyonun, nefret söyleminin ve daha nice zararlı içeriğin dünyanın dört bir yanındaki insanlara yayılmasına sebep oldu. YouTube'un muazzam etkisi, bu videoların geniş kitlelere yayılması, sayısız hayatı derinden etkilmesi ve pek çok soruna yol açması anlamına geliyor. Buna rağmen YouTube bu eleştirilere eylemsizlik ve gizlilikle cevap veriyor.

Mozilla, YouTube'un öneri algoritması konusunda daha şeffaf olması ve araştırmacıların platformu incelemesine izin vermesi gerektiğini yıllarca savunduktan sonra, 2020 yılında YouTube kullanıcılarını ses çıkarmaya teşvik ederek platformun eylemsizliğine cevap vermiş oldu. Bu amaç doğrultusunda, YouTube algoritmasının insanlara verebileceği zararları daha iyi anlamak için, bir tarayıcı eklentisi ve kitle kaynaklı araştırma projesi olan RegretsReporter'ı [Pişmanlık Bildiri aracını] hayata geçirdik.

YouTube gözlemcisi olmayı üstlenen 37 bin 380 YouTube kullanıcısı, Mozilla araştırmacılarının titizlikle analiz etmesi için YouTube'da yaşadıkları pişmanlık verici deneyimlere dair verileri paylaştılar.

Mozilla bu sayede, YouTube algoritmasına dair bugüne kadarki en büyük kitle kaynaklı araştırmayı yürüterek platformun sıkı sıkıya korunan veri havuzuna dair bilgi edindi. Araştırma gönüllüleri, Haziran 2020 ile Mayıs 2021 arasında 91 ülkeden 3 bin 362 olumsuz video bildirdi.

**Bu rapor, RegretsReporter araştırmamızdan öğrendiklerimizi ortaya koyuyor.**

**Araştırmanın üç ana bulgusu şöyle:**

- **Çok çeşitli kategorilere yayılan YouTube Pişmanlıkları aynı zamanda rahatsız edici.** Gönüllülerimiz Covid ile ilgili korku tellallığından siyasi dezenformasyon ve son derece uygunsuz "çocuk" çizgi filmlerine kadar sayısız içeriği bildirdi. En sık karşılaşılan pişmanlık kategorileri dezenformasyon, şiddet görüntüleri veya rahatsız edici içerik, nefret söylemi ve dolandırıcılık oldu.
- **Sorun, algoritmanın kendisi.** Pişmanlık bildirimlerinin yüzde 71'i, YouTube'un öneri sisteminin gönüllülerimize önerdiği videolardan geldi. Dahası, önerilen videoların bildirilme ihtimali, gönüllerinin kendi aradıkları videoları bildirme ihtimalinden yüzde 40 daha yüksekti. Birkaç vakada, YouTube'un kendi Topluluk Kuralları'nı ihlal eden veya daha önce izlenen videolarla alakasız içerikleri önerdiği görüldü.
- **İngilizce konuşmayan kullanıcılar en çok etkilenenler.** Birincil dili İngilizce olmayan ülkelerde (özellikle Brezilya, Almanya ve Fransa'da), YouTube'da karşısına çıkan içerikten rahatsızlık duyduğunu iletenlerin oranı yüzde 60 daha yüksek ve pandemiyle bağlantılı Pişmanlıklar da İngilizce dışındaki dillerde bilhassa yaygın.

<sup>1</sup> YouTube, platformda her gün bir milyar saat video izlendiğini belirtiyor. YouTube CPO'su Neal Mohan 2018'deki CES fuarında bu sürenin en az yüzde 70'inin yapay zeka destekli önerilerden geldiğini belirtti. Bu sayıların doğru olduğunu kabul edersek YouTube'daki yapay zeka destekli önerilerin günlük 700 milyon saat izleme sağladığı sonucuna varabiliriz.

Bu raporda, veri analizi ve çok sayıda vaka incelemesiyle tüm bu bulguları derinlemesine inceliyoruz. Mozilla araştırmasından birkaç örnek bile durumu net bir şekilde ortaya koyuyor. Bir gönüllü, "Oyuncak Hikayesi" adlı çocuk filminin cinsel içerikli bir parodisi olan "Woody's Got Wood" başlıklı bir animasyonun kendisine önerildiğini bildirdi. Başka bir gönüllüye, ABD Temsilciler Meclisi üyesi Ilham Omar'a ve seçim hilesine dair çürütülmüş bir komplo teorisini içeren bir video önerildi. "Blacks in Power Don't Empower Blacks" [İktidardaki Siyahlar, Siyahları Desteklemiyor] adlı bir başka öneri de ırkçı ve saldırgan ifadeler ve fikirler içeriyordu.

Mozilla'nın tek amacı sorunu tespit etmek değil; sorunun çözülmesine yardımcı olmak da istiyoruz. Bu raporun "Tavsiyeler" kısmında YouTube gibi diğer internet platformlarının, politikacıların ve bireylerin faydalanabileceği net bir kılavuz sunuluyor.

Bu tavsiyelerden bazıları şöyle:

- **Platformlar**, araştırmacıların öneri sistemini denetlemesine izin vermeli.
- **Platformlar**, öneri sistemlerinin nasıl işlediğine dair bilgileri paylaşmanın yanı sıra sorunlara ve zaman içindeki iyileştirmelere dair yeterli bilgi veren şeffaflık raporları yayınlamalı.
- **Politikacılar**, YouTube'un öneri algoritmalarının bağımsız bir şekilde incelenmesine imkan sağlayacak bilgileri paylaşmasını ve araçları geliştirmesini sağlamalı.
- **Politikacılar**, platformları incelemek için alternatif yöntemlere başvuran araştırmacıları, gazetecileri ve gözlemcileri korumalı.
- **Bireyler**, kendileri ve aileleri için gerekli kontrolleri sağladıklarından emin olmak için YouTube ve Google'daki veri ayarlarını güncellemeli.

*Bu raporda ayrıca Mozilla'nın araştırmasını açıklayan kapsamlı bir [metodoloji bölümü](#) yer alıyor. Araştırma sorularımız, tarayıcı uzantısının işleyişi ve analiz yöntemlerimiz de dahil pek çok konuya dair detayları bu bölümde bulabilirsiniz.*

---

<sup>2</sup> Mozilla, Anti-Defamation League, New York Times, Washington Post, Wall Street Journal ve diğer kuruluşlar, akademik araştırmacılar ve yayıncılar tarafından yapılan araştırmalar, YouTube'un sebep olduğu dezenformasyon, kutuplaştırma ve radikalleştirmeyi ortaya koyuyor.

# Giriş

Sam, depresyon teşhisi almış 13 yaşında bir çocuktü ve en savunmasız olduđu dönemde alt-right denen alternatif sađ hareketine dahil edildi. Sam'in radikalleşmesinin sebeplerinden biri, YouTube ve diđer sosyal medya platformlarında önerilen içeriklerin derinliklerinde kaybolmasıydı. Mozilla'nın yaptığı görüşmede Sam'in annesi, durumu şöyle açıkladı: "YouTube ve benzeri platformlar, hesap vermemek için ellerinden geleni yapıyor gibi görünüyor. Şirketler, ifade özgürlüğüyle ilgili her türlü savın ardına saklanıyor ama bence tüm bunlar sorumluluktan kaçmak ve daha çok para kazanmak için başvurdukları elverişli yollar. Hele de bilgi ve destek bulmak için internete yönelen çocuklar bilhassa savunmasız."

Sam'in hikayesi bir istisna deđil. Araştırmacılar ve araştırmacı gazeteciler, YouTube önerilerinin insanları internetin en karanlık ve aşırılıkçı köşelerine sürüklediğini yıllardır ortaya koyuyor. Radikalleşme gibi karmaşık problemlerin yegane sorumlusu YouTube olmasa da deliller şirketin öneri algoritmasının ciddi bir rol oynadığına işaret ediyor. Algoritma insanları bu içeriğe yönlendiriyor ve insanlar bu kuyuya "düştüğü" anda da uç fikirler içeren daha çok öneri yapılmaya başlanıyor.

YouTube'un öneri algoritmasına dair bu problem, ticari algoritmaların hayatımız üzerindeki esrarengiz ve belirsiz etkisine dair daha büyük bir meselenin parçası. YouTube algoritması, günde yaklaşık 700 milyon saat izleme süresi üretse de toplum bu algoritmanın nasıl işlediğine dair neredeyse hiçbir şey bilmiyor. Bu algoritmayı incelememize imkan verecek resmi yöntemler mevcut deđil. Sam'inki gibi hikayeler gün yüzüne çıktığında ne olduğunu ve aynı şeyin yeniden yaşanmasını engellemek için neler yapılabileceğini anlamak imkansız.

Özünde YouTube algoritması toplumun deđil, YouTube'un çıkarlarına göre işliyor.

Mozilla 2019 yılından beri YouTube'un öneri algoritmasının şeffaf olması gerektiğini savunuyor. YouTube Regrets [YouTube Pişmanlıkları] adlı kampanyamız, Sam'inkine benzer 28 hikayeyi gün yüzüne çıkardı; YouTube deneyimleri yüzünden hayatları rayından çıkan, hatta bazen sonsuza dek deđişen insanların hikayeleri. Bu hikayeleri yayınlamaktaki amacımız, güvenilir algoritmaların insanlar üzerindeki etkisine dair farkındalık yaratmak ve YouTube'a algoritmalarının işleyişine dair daha sorumlu ve açık davranması yönünde baskı yapmaktı. Aradan geçen zamana rağmen, YouTube hâlâ bu konuda adım atmış deđil.

İşte RegretsReporter'ı tam da bu sebeple geliştirdik. RegretsReporter, ilk olarak başlattığımız YouTube Pişmanlıkları kampanyamızı genişleten kitle kaynaklı bir araç. Mozilla gönüllülerini, YouTube'daki deneyimlerine dair veri sağlama gücü veriyor. RegretsReporter, en önemli algoritmaları kamu çıkarını gözeterek inceleyen alternatif yöntemlere sahip diđer tarayıcı uzantılarına benziyor. Bu araçlar, bireylere ve kamu kurumlarına karşı şeffaflık politikasının yerine geçemez elbette, fakat şirketleri insan hayatı üzerindeki etkilerinden sorumlu tutabilmek için şimdilik elimizdeki en iyi seçenek, bu türden araçlar.

Bu araştırmanın amacı, RegretsReporter aracılığıyla edindiğimiz bilgileri paylaşmak. Buzdağının görünen yüzü olduğunu düşündüğümüz bu bulguların, YouTube algoritmalarının şeffaf hale getirilmesinin ne kadar acil bir mesele olduđu konusunda toplumu ve kamu çıkarını koruması gerekenleri ikna etmesini umuyoruz.

<sup>3</sup> "Alt-right," ABD merkezli beyaz milliyetçiliđi merkezine alan aşırı sađcı bir hareketi ifade ediyor. Bu hareket özellikle internet üzerinde aktif olup, 2010'lu yıllarda ortaya çıkmış ve küresel çapta etkisini genişletmiştir. Terim, beyazların üstünlüğünden eleştirel milliyetçiliđe kadar farklı görüşleri barındırırken, eşitlikçiliđi, demokrasiyi ve çokkültürlülüđu reddeder. Bu heterojen hareket, radikal sađ sosyal hareketlerini kapsayan bir terim olarak düşünülebilir, ancak net bir tanımlaması bulunmamakta.



# “Pişmanlık” derken ne kastediyoruz?

“YouTube Pişmanlığı” kavramı, Mozilla’nın 2019’da geliştirdiği kitle kaynaklı küresel kampanyadan doğdu. YouTube’un, özellikle öneri algoritması aracılığıyla, insanları tehlikeli ve tuhaf yollara sevk edişine dair hikayeler topladık. İnsanların YouTube’da ne türden deneyimler yaşadığını anlama yönündeki ilk adımımız bu kampanyaydı. Bu hikayelerin nasıl olması gerektiğine dair bir yönlendirmemiz olmadı; gönderilenler, kullanıcıların kendi başlarına “pişmanlık verici” addettiği hikayelerdi. O günden bu yana Mozilla, insanların YouTube’da yaşadığı her türlü kötü deneyimi tanımlamasına imkan sağlamak için, nelerin YouTube Pişmanlığı “sayıldığına” dair kesin bir tanım vermekten bilerek kaçındı.

Birey odaklı yaklaşımımız, zararlı çevrimiçi içeriğin ne olduğuna dair genelde son derece hukuki ve teorik olan tartışmaları genişletmek için insanların, özellikle de savunmasız ve/veya ötekileştirilmiş birey ve toplulukların gerçek deneyimlerini merkeze alıyor. Niteliksel analizimiz, büyük oranda Exeter Üniversitesi’nden araştırma görevlileri tarafından yürütüldü ve bu araştırmacılar, bir videonun platformda olup olmaması ve/veya YouTube tarafından önerilip önerilmemesi gerektiğini tespit etmek için, platformda nelere izin olduğunu ortaya koyan Topluluk Kuralları’ndan faydalandı (bu çalışmanın detayları, raporun “Analiz Yöntemleri” bölümünde yer alıyor). Bize bildirilen videolardaki yorumları da analizimizi derinleştirmek kullandık. Bu metodoloji, objektif ayrımlar yapmıyor; daha fazla tartışma ve araştırma olmasını teşvik etmek için öznel izlenimleri gün yüzüne çıkarıyoruz.

Araştırmamızda nefret söylemi, çürütülmüş siyasi veya bilimsel yanlış bilgiler ve YouTube’un topluluk kurallarını ihlal edebilecek (veya halihazırda ihlal ettiği tespit edilmiş) içeriklere rastlamakla birlikte, çevrimiçi zararın ne anlama geldiğine dair daha karmaşık bir tablo çizen pek çok örneği de keşfettik. Bize bildirilen videoların çoğu, Topluluk Kuralları’nın “sınırlarında gezip” bu kuralları tam olarak ihlal etmeyen, YouTube’un “ihlale yakın içerik” dediği kategoriye giriyor olabilir. YouTube ihlale yakın içeriği nasıl tanımlayıp sınıflandırdığını açıklamadığı için bu varsayımı doğrulamanın bir yolu yok. Araştırmamız ayrıca pişmanlık veren veya zararlı çevrimiçi deneyimlerin çoğu zaman bir kişinin zaman içinde nasıl hedeflendiğine bağlı olduğuna işaret ediyor ki bunu tek tek videolara bakarak anlamak veya çözmek oldukça zor.

Bu raporda sunduğumuz öneriler -sağlam bir şeffaflık çağırısı, denetleme ve öneri algoritmasının kontrolünü insanlara verme-, gücü bu sistemlerden doğrudan etkilenen insanlara vermek için elzem.



# Bulgular

- 1- Çok çeşitli kategorilere yayılan YouTube Pişmanlıkları aynı zamanda rahatsız edici
- 2- Sorun, algoritmanın kendisi
- 3- İngilizce konuşmayan kullanıcılar en çok etkilenenler

# 01.

## Çok çeşitli kategorilere yayılan YouTube Pişmanlıkları aynı zamanda rahatsız edici

### Özet

*“Her gün milyonlarca insan bilgilenmek, ilham almak ya da sadece eğlenmek için YouTube’a geliyor”*

\* [YouTube, 2020, “How YouTube Works”](#)

- Bildirilen içerikler Covid ile ilgili korku tellallığından siyasi dezenformasyona ve son derece uygunsuz “çocuk” çizgi filmlerine kadar uzanıyor. En sık karşılaşılan kategoriler dezenformasyon, şiddet görüntüleri veya rahatsız edici içerik, nefret söylemi ve dolandırıcılık.

### Hikaye

İnsanların internet platformlarına dair algısı son yıllarda olumsuz yönde değişti. Bir zamanlar herkesin hayran olduğu Facebook ve Twitter bugün nefret yayan içeriklerin, kabalığın ve dezenformasyonun cirit attığı mecralar olarak görülüyor.

Öte yandan YouTube ise insanların gözündeki olumlu imajını, en azından benzerlerine kıyasla, büyük ölçüde korumayı başardı. Pek çoklarının gözünde YouTube hâlâ tatlı tepki videolarının ve faydalı kendin yap içeriklerinin yer aldığı ve iyi niyetli içerik üreticilerinin başarılı olabildiği bir yer.

YouTube’un çok iyi yanlarının olduğu bir gerçek. Her gün milyonlarca insana eğitim ve eğlence içerikleri sunan bir platform. Fakat tıpkı benzerleri gibi YouTube’un da, her ne kadar insanlar bunu diğer platformları kadar açık seçik göremese de, karanlık bir yüzü var.

Mozilla 2019 yılında işte bu karanlık yüzü incelemek için kolları sıvadı; giderek daha çok kullanıcının karşılaşıp gündeme getirdiği pişmanlık verici videoları daha iyi anlamaktı amaç. Bu içeriklere “YouTube Regrets” [YouTube Pişmanlıkları] adını verdik ve insanların deneyimlerini paylaşmaları için bir çağrıda bulunduk. Geri dönüşler kaygı uyandırıcıydı.

Hoş ve faydalı videoların arasında bir de tuhaf, hatta zararlı videolar yığını var. Trafik kazası videolarının kuyularına düşen bir kişi Mozilla’ya “insanların kazadan sağ çıkmadığı” videolar da gördüğünü belirtti. Dans videoları arayan 10 yaşındaki bir kız, çok sıkı diyetleri teşvik eden içeriklere maruz kaldı. Mozilla’yla iletişime geçen ailesi, “Yiyip içtiğini kısıtlar hale geldi” diye belirtiyor. LGBTQ+’ları destekleyen içerikler ararken nefret dolu, LGBTQ+ karşıtı videolarla karşılaşan biri, Mozilla’ya şöyle yazdı:

“Hâlâ kimliklerini keşfetmeye çalışan insanlara bunun ne kadar zarar verebileceğini tahmin edebiliyorum.”

Bunların dışında biri sürü başka YouTube Pişmanlığı var: sahte bilimi destekleyen, 11 Eylül komploları yayan, hayvanlara zulüm içeren, beyazların üstünlüğünü savunan videolar. Tüm bunlar Mozilla'yı şuna ikna etti: YouTube'da ciddi bir problem var. Bu yüzden araştırmamızı derinleştirmeye, insanların platformdaki deneyimlerini anlamaya ve YouTube öneri algoritmasının bu videoları yaymadaki rolünü tespit etmeye karar verdik.

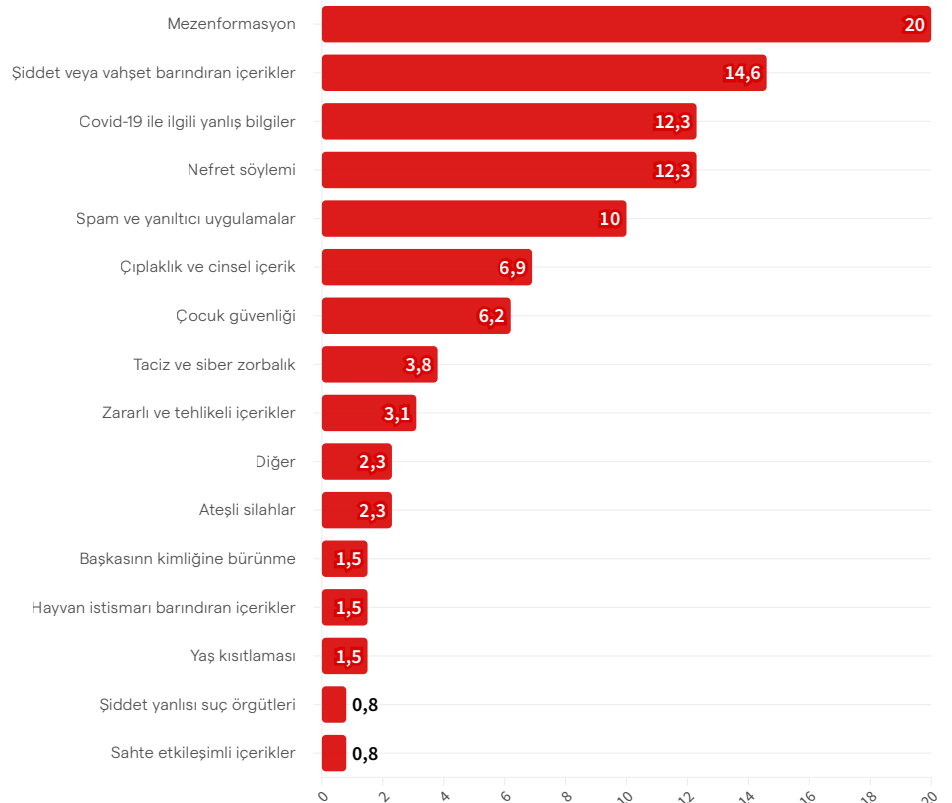
## Veriler

Değerlendirmelerimize göre bildirilen içeriklerin yüzde 12,2'sinin (yüzde 95 güven aralığıyla yüzde 10,4 ila yüzde 14,2), YouTube'un Topluluk Kuralları gereğince, ya "YouTube'da olmaması" ya da "aktif olarak önerilmemesi" gerekiyor. Bu videoların kategorilerini aşağıdaki grafikte görebilirsiniz.

Bu videoları sınıflandıran araştırma görevlileri, dezenformasyonu tespit edip sınıflandırma ve yasadışı içerik kategorileri konularında uzman değiller, fakat kendi muhakemelerine ve rahatsız edici ve daha detaylı incelemeye değer içerik algılarına göre bir sınıflandırma yaptılar. Bu sınıflandırmaya göre en sık karşılaşılan kategorilerin dezenformasyon, şiddet görüntüleri veya rahatsız edici içerik ve Covid-19'a dair dezenformasyonun ardından (özel ilgi gerektiren bir vaka olduğu için Covid-19'u ayrı kategorize ediyoruz), nefret söylemi ve dolandırıcılık olduğunu gördük. Dikkate değer diğer kategoriler çocuk güvenliği, taciz ve siber zorbalık ve hayvan istismarı. YouTube Pişmanlıklarının ortaya çıkmasının pek çok sebebi olduğu kesin, fakat en baskın problem dezenformasyon. Covid-19 ile ilgili dezenformasyon, tüm dezenformasyonla birlikte tek bir kategoride toplandığında Pişmanlıkların yaklaşık üçte birini bu kategori oluşturuyor.

## Pişmanlık bildirilen içeriğin türü (%)

Bu kategorilerden en çok dikkat çeken birkaç örneğe aşağıda yer vereceğiz. Ekler bölümündeyse daha fazla örnek bulabilirsiniz.



Bir gönüllümüze "Omar Connected Harvester SEEN Exchanging \$200 for General Election Ballot. 'We don't care illegal,'" [Omar ile bağlantılı bir çiftçi, Genel Seçim için 200 dolar alırken görüldü. 'Kanunsuzluk Umurumuzda Değil'" başlıklı bir video önerilmiş. Bu video, ABD Temsilciler Meclisi üyesi Ilhan Omar'a ve 2020 ABD seçimlerinde hile yapıldığına dair dayanaksız bir iddiayı yayıyor. Gönüllü, genellikle vahşi doğada hayatta kalma videoları izlemesine rağmen sürekli aşırı sağcı kanalların kendisine önerildiğini belirterek bu önerilerin, siyasi olarak aşırılıkçı içeriklerin hedefinde yer aldığı daha geniş çaplı bir örüntünün parçası olabileceğine işaret etti.



Omar Connected Harvester SEEN Exchanging \$200 for General Election Ballot."We don't care illegal."  
542686 views - Sep 28, 2020

Bir gönüllü, "BILL GATES HIRED BLM "STUDENTS" TO COUNT BALLOTS IN BATTLEGROUND STATES" [Bill Gates, çekişmeli eyaletlerde oyları sayması için BLM "öğrencileri" tuttu] başlıklı bir videonun önerildiğini bildirdi. Videoda Microsoft kurucusunun seçime hile karıştırmak için etnik azınlıklardan gençleri tuttuğu yönünde dayanaksız ve suçlayıcı bir iddia yer alıyor.



BILL GATES HIRED BLM "STUDENTS" TO COUNT BALLOTS IN BATTLEGROUND STATES  
14 views - Nov 21, 2020

Başka bir gönüllünün bildirdiği "The Elites Who Control You" [Sizi Kontrol Eden Elitler] başlıklı videoda ise bir siyasetçi veya "eliti" taklit eden biri, korku yaratıp kitleleri kontrol etmek için Covid-19 hakkında yalan söylediklerini öne sürüyor.



The Elites Who Control You  
743641 views - Dec 1, 2020

Bir gönüllü, "7 jokes ending in tragedy" [Trajediyle sonuçlanan 7 şaka] başlıklı bir videonun karşısına çıktığını bildirdi. Videodaki hikayeler oldukça üzücü ve video için seçilen küçük resim de rahatsız edici ve kanlı.



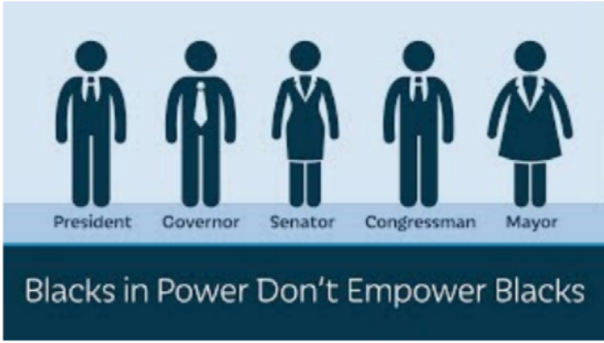
7 SCHERZI finiti in TRAGEDIA  
2193169 views - 6 apr 2018

Bir gönüllü tarafından bildirilen "Woody's Got Wood" başlıklı video, çocuk filmi "Oyuncak Hikayesi"nin cinsel içerikli bir parodisi.



Woody's Got Wood Animated.mp4  
22251 views - Sep 17, 2018

Bir gönüllüye "Blacks in Power Don't Empower Blacks" [İktidardaki Siyahlar, Siyahları Desteklemiyor] başlıklı, saldırgan ve ırkçı söylemler içeren bir video önerildi.



Blacks in Power Don't Empower Blacks  
2365794 views - Mar 26, 2018

Bir gönüllüye, tuhaf olarak kategorize ettiği, "El Arca - It's Literally Furry Noah's Arc" başlıklı bir video önerildi. Bu video, ele aldığı çocuk filminin cinsel ve yetişkinlere yönelik içerik barındırdığını iddia edip bu meseleyi derinlemesine tartışıyor.



El Arca - It's Literally Furry Noah's Arc  
119145 views - 22 Apr 2021

Bir gönüllüye "Biggest fraud' in US history—up to 300,000 fake people voted in Arizona election: expert | NTD" [ABD tarihinin 'en büyük dolandırıcılığı' - Arizona seçimlerinde 300 bin kadar sahte insan oy kullandı: uzman | NTD] başlıklı bir video önerildi.



'Biggest fraud' in US history—up to 300,000 fake people voted in Arizona election: expert | NTD  
965541 views - 1 Dec 2020

Bir gönüllü, çocukların kanının toplandığına dair bir QAnon komplo teorisi içeren, "[Documentary 2020] [Adrenochrome — The Darkest Drug of Them All!]" [Adrenochrome - En Karanlık Uyuşturucu!] başlıklı bir videonun kendisine önerildiğini bildirdi.



[Documentary 2020] [Adrenochrome - The Darkest Drug of Them All!]  
108279 views - Premiered Mar 23, 2020

# 02.

## Sorun, algoritmanın kendisi

### Özet

*“Öneriler isabetli olduğunda kullanıcıların aşık olacakları yeni bir şarkı keşfetmelerine, favorileri haline gelen yeni bir içerik üreticisine rastlamalarına veya harika yemek tariflerini öğrenmelerine yardımcı oluyor. Öneri sistemimizi sürekli güncellememizin sebebi de bu; insanların gerçekten izlemek istedikleri videolar önerdiğimizden emin olmak istiyoruz.”*

\* [YouTube, 2019, “Continuing our work to improve recommendations on YouTube”](#)

- **Önerilen Pişmanlıkların yaklaşık yüzde 9’u, aradan geçen zamanda YouTube’dan kaldırıldı, hatta aralarından birkaçı, platformun kendi Topluluk Kuralları’nı ihlal ettiği bildirildikten sonra kaldırıldı.** YouTube’dan kaldırılan bu içeriklerin toplam izlenme sayısı, bildirildikleri dönemde 160 milyondur.
- **YouTube Pişmanlıklarının başlıca sorumlusu öneriler.** Gönüllülerimizden gelen Pişmanlık bildirimlerinin yüzde 71’ini, kendilerine önerilen videolar oluşturuyor ve önerilen videoların pişmanlık uyandırma ihtimali, kasıtlı olarak aranan videolara kıyasla yüzde 40 daha fazla.

- **YouTube Pişmanlıkları olarak bildirilen videolar platformda iyi performans sergiliyor.** Bildirilen videolar, gönüllülerimizin izlediği diğer videolara kıyasla günde yüzde 70 daha fazla izlenme alıyor.
- **Pişmanlık uyandıran önerilerin genelde önerilmesi için bir sebep yok.** Bir gönüllünün pişmanlık uyandıran bir video izlemeden önceki rotasına dair verilere sahip olduğumuz vakaların yüzde 43,3’ünde, önerilen video, gönüllünün daha önce izlediği videolarla tamamen alakasızdı.

### Hikaye

HuffPost’un 2020 tarihli bir haberi, 11 yaşındaki Allie’nin hikayesine yer verdi. Allie, kanalını bulup kendisinden kamera önünde bayılma taklidi yapması gibi özel isteklerde bulunan istismarcıların cinsel fantazilerini farkında olmadan yerine getirip platforma yüklüyordu. HuffPost’un araştırmasına göre daha da beteri, YouTube’un öneri algoritmasının istismarcılara Allie’ninki gibi kanalları bulmaları için bilfiil yardım etmesiydi. Rahatsız edici sonuçları gözler önüne seren araştırma, “algoritmaya dayalı tanıtımın” olumsuz yönlerini ortaya çıkarmanın yanı sıra -yıkanaan, jimnastik yapan, bacaklarını açarak oturan çocuklar gibi- görünürde “masum” olup YouTube’un Topluluk Kuralları’nın ihlal etmeyen videoların gruplandırılıp, bu içeriklerle ciddi ölçüde etkileşime giren insanlara önerilmesi durumunda neler olabileceğini gösterdi.



Allie'nin hikayesi, YouTube'un otomatik önerilerinin, belli videoları hangi bağlamda izlendiklerini göz önünde bulundurmadan "viral" haline getirebildiğinin canlı ve rahatsız edici bir örneği. YouTube'un öneri sistemi paylaşımına açık içeriklere dayandığından, yani kullanıcıların ürettikleri videoları hiçbir kontrolden geçirmeden önerdiğinden, algoritmanın verdiği kararlar, yalnızca insanların kontrolünden geçmiş içerikleri yayınlayıp öneren Netflix gibi platformların önerilerine kıyasla daha farklı (tehlikeli) sonuçlara yol açabilir.

Kullanıcıların ürettiği içerikleri otomatik bir sistemle öneren pek çok platform, içeriği tespit edip yönetmesi için algoritmalara bel bağlıyor. Ne var ki zaten oldukça sınırlı olan bu algoritmaların bir de insanlara (bir videoyu ne kadar süre izledikleri gibi ipuçlarına dayanarak) "ilgilerini çekecek" içerikleri önermek için tasarlanmış diğer algoritmalarla birlikte çalışması gerekiyor. Bu da şu anlama geliyor: Algoritmaların, diğer tüm tahminlerin yanı sıra, bir videonun bir kişinin "ilgisini çekme" ihtimali ile bir videonun YouTube Topluluk Kuralları'nı ihlal etme ihtimalini kıyaslaması gerekiyor. Bu hem çok büyük hem de ticari kaygılar ile insanların iyiliğinin çatışmasına yol açabilecek bir sorumluluk.

Araştırmamızda, YouTube algoritmasının bu iki sorumluluğu dengelemeyi beceremediği birkaç vakayla karşılaştık. YouTube algoritmasının kendi Topluluk Kuralları'nı ve diğer içerik politikalarını ihlal eden videoları önerdiği ve bu videoların ancak milyonlarca kez izlendikten sonra platformdan kaldırıldığı birkaç vaka tespit ettik. Araştırmamızın vardığı bir diğer sonuç da şu: Gönüllülerimizin platformda yaşadığı kötü deneyimlerin büyük çoğunluğundan, arama yaparak veya başka şekillerde karşılaştıkları YouTube videolarına kıyasla, YouTube algoritmasının desteklediği öneriler sorumlu.

Ayrıca gönüllülerimizin bildirdiği YouTube Pişmanlıkları'nın izlenme sayısının, platformdaki diğer videolara kıyasla daha hızlı arttığını tespit ettik. Bu örnekler, algoritmanın vermesi gereken çeşitli kararların YouTube tarafından nasıl önceliklendirildiğine ve bu süreçte nasıl tavizler verildiğine dair ciddi soru işaretleri doğuruyor.

YouTube algoritması, bir kukladan ziyade kuklacı gibi. Ne önerdiği, neden önerdiği ve bu kararların nasıl alındığı, özellikle Allie'ninki gibi vakalarda gerçekten çok önemli. Nisan 2021'de yapılan ABD Senatosu Yargı Komitesi duruşmasında Senatör Chris Coons, YouTube'dan bir videonun sadece kaç kere izlendiğine dair değil, kaç kere önerildiğine dair bilgileri de paylaşmasını talep etti. Bu bilgi, YouTube algoritmasının bir içeriğin viral haline gelmesindeki rolünü anlamak için son derece önemli, ayrıca mesela Allie'nin videolarının YouTube tarafından neden önerildiğini anlamamıza da yardımcı olabilir. Fakat YouTube temsilcisi, bu kritik bilgiyi paylaşmadı.

## Veriler

Bize bildirilen Pişmanlıklardan bazıları YouTube tarafından daha sonra kaldırıldı. YouTube, politikalarını ihlal eden içerikleri tespit edip platformdan kaldırmak için insanları ve makine öğrenimini birleştiren bir yöntem kullanıyor. Fakat bu sistemler sık sık hata yapıyor ve kimi durumlarda otomatik önerilerle bir araya geldiklerinde YouTube, kendi kurallarını ihlal eden içerikleri öne çıkarabiliyor. Gönüllülerimize önerildiği için bize bildirilen videolardan 189 tanesi, 1 Haziran 2021 itibarıyla platformdan kaldırıldı.

YouTube kimi durumlarda videoların kaldırılma nedenine dair bilgi veriyor (örneğin Topluluk Kuralları'nı ihlal, telif hakkı ihlali, nefret söylemi yasaları ya da videonun silinmesi veya sahibi tarafından gizlenmesi). Fakat analiz ettiğimiz

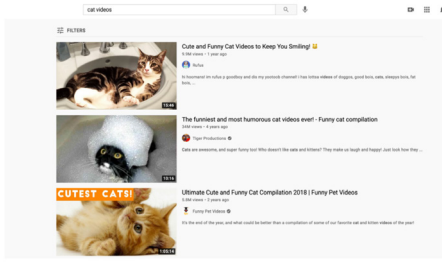
vakalardan neredeyse yüzde 40'ında neden kaldırıldığında dair bir neden sunulmadan sadece "videoya erişilemiyor" ibaresi yer alıyordu.

Bildirildikleri sırada ortalama beş aydır platformda yer alan bu videoların toplam izlenme sayısı video başına günde ortalama 760 bin izlenme ile 160 milyonu buldu. Bu izlenmelerin kaçının YouTube öneri algoritmasından kaynaklandığını tespit etmek mümkün değil, çünkü YouTube bu verileri paylaşmıyor, fakat her videonun (onu bildiren gönüllüye) en az bir kere önerildiğini biliyoruz.

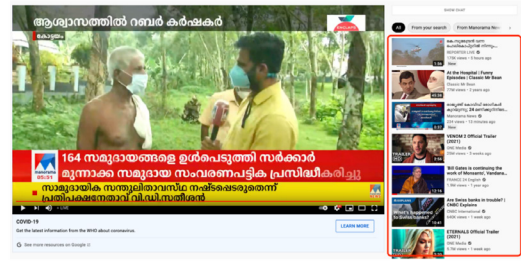
Verilerimiz sayesinde Pişmanlık oranlarını veya pişmanlık uyandıran videoların bir gönüllünün izlediği tüm videolara oranını hesaplayabiliyoruz. Bu oranın altında yatan birkaç faktörü tespit etmek mümkün; gönüllünün videoya nasıl ulaştığı bunlardan biri (örneğin arama, öneri veya link yoluyla). Başlıca iki ulaşma biçimini analiz ettik:

- Aramalar, gönüllünün bir arama yapıp YouTube'un bu arama sonucunda sunduğu videolardan birini izlediği vakaları içeriyor.
- Öneriler, YouTube'un içeriği kendi kendine gönüllüye önerdiği vakaları içeriyor.

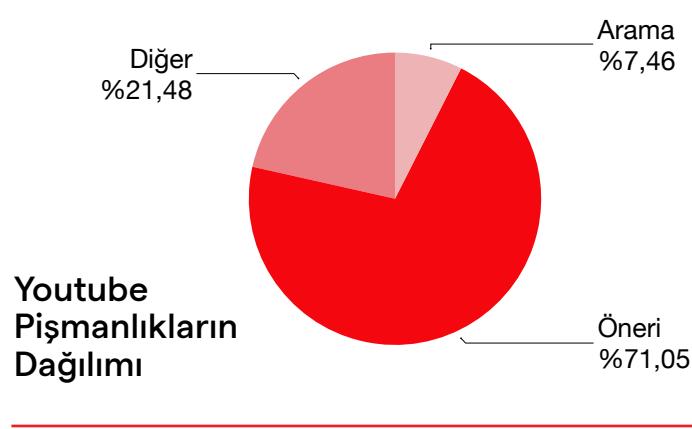
### Arama



### Öneri



Öneriler ve aramaların Pişmanlık oranlarını kıyasladığımızda, YouTube Pişmanlıklarından önerilerin büyük oranda sorumlu olduğu açıkça görülüyor. Aramalarda, izlenen 10 bin video başına Pişmanlık oranı 9,6 iken, önerilerde 13,9'a ulaşır ki bu da yüzde 40 daha yüksek olduğu anlamına geliyor. Öneriler ayrıca bize bildirilen videoların da yüzde 71'ini oluşturuyor.



Referans olması açısından, YouTube'un politikalarını ihlal eden videoların izlenme oranlarını içeren İhlal İzlenme Oranı raporuna göre, bu oran her 10 bin video başına 17. Bizim bulduğumuz oranın (öneriler için 10 bin video başına 13,9), YouTube'un kendi yayınladığı bulgulardan yüksek olmaması, pişmanlık verici video kavramımızın isabetli olduğuna işaret ediyor.

Bildirilen videolar ayrıca daha hızlı izleyici topluyor. Bildirildikleri sırada YouTube Pişmanlıkları günde ortalama 5 bin 794 izlenme alıyordu. Gönüllülerimizin (bildirilen videolara varmadan önceki rotalarında) izledikleri, günlük ortalama izlenme sayısı ortalama 3 bin 312 olan videolardan yüzde 70 daha fazla izlenmişlerdi.

Pişmanlık verici olarak bildirilen videoya çıkan "rotada" izlenen videoları analiz ettiğimizde, bu "rota"nın en sonunda önerilen videoların bazılarının, gönüllünün daha önce izlediği şeylerle tamamen alakasız olduğunu fark ettik. Bu eğilimi incelemek amacıyla, araştırma görevlilerinden pişmanlık uyandıran içeriğe varmadan önceki rotada izlenen videoları sınıflandırıp önerilerin alakalı olup olmadığını tespit etmelerini istedik. Gönüllünün rotasına dair verimizin olduğu bildirimlere bakıldığında (toplam bildirimlerin yüzde 43,3'ü), pişmanlık uyandıran öneri, gönüllünün daha önce izlediği videolarla tamamen alakasızdı.

Örneğin, Art Garfunkel'in bir klibini izleyen bir gönüllüye "Trump Debate Moderator EXPOSED as having Deep Democrat Ties, Media Bias Reaches BREAKING Point" [Trump Münazarasındaki Moderatörün Demokratlarla Yakın İlişki İçinde Olduğu İFŞA EDİLDİ, Medyanın Yanlılığı SINIR TANIMIYOR] başlıklı bir video önerildi. Gönüllülerin bu videolara dair yorumları, daha önce izledikleriyle alakası olmayan videoların önerilmesinden, hele de bu videolar görüşlerine tersse veya katılmadıkları bir fikri savunuyorsa, bıktıklarını ve rahatsız olduklarını gösteriyor.

#### Bildirilen video



Trump Debate Moderator EXPOSED as having Deep Democrat Ties, Media Bias Reaches BREAKING Point  
166310 views - Oct 18, 2020

#### Bu oturumdaki video geçmişi

- Art Garfunkel and his son cover The Everly Brothers live in Napa, May 12, 2019 (4K)

Bir gönüllü "Memes, Misinfo, and the Election - October 12, 2020" [Meme'ler, Yanlış Bilgiler ve Seçimler - 12 Ekim 2020] başlıklı bir Mozilla videosunu izledikten sonra, "Global Warming: Fact or Fiction? Featuring Physicists Willie Soon and Elliott D. Bloom" [Küresel Isınma: Gerçek mi, Uydurma mı? Fizikçiler Willie Soon ve Elliott D. Bloom'un Katılımıyla] başlıklı bir öneriyle karşılaştı. Gönüllü, bir Mozilla videosu izledikten sonra küresel ısınmayı reddeden bir videoyla karşılaştığına şaşırdığını belirtti.

#### Bildirilen video



Global Warming: Fact or Fiction? Featuring Physicists Willie Soon and Elliott D. Bloom  
592557 views - Aug 16, 2019

#### Bu oturumdaki video geçmişi

- Memes, Misinfo, and the Election - October 12, 2020

ABD ordusuyla ilgili videolar izleyen bir gönüllüye, kadın düşmanı, cinsiyetçi ve ayrımcı olan "Man humiliates feminist in viral video" [Bir adamın bir feministi aşağıladığı viral video] başlıklı bir video önerildi.

#### Bildirilen video



Man humiliates f3m1n1st in v1ral video  
368535 views - Mar 6, 2021

#### Bu oturumdaki video geçmişi

- Top 5 Things US Government Denies (Marine Reacts) [Via Recommendation]
- Marine Reacts - Can the US Defend an Invasion from Abroad? [Via Recommendation]
- North Korean Soldier meet U.S. Soldier For The First Time

# 03.

## İngilizce konuşmayan kullanıcılar en çok etkilenenler

### Özet

*“YouTube’un Topluluk Kuralları, içeriğin nereden yüklendiğinden bağımsız olarak, tüm dünyada tutarlı bir şekilde uygulanıyor.”*

\* [YouTube, 2021, “YouTube Community Guidelines Enforcement Transparency Report”](#)

- İngilizce dışındaki dillerde içerikleri izleyen kullanıcılar en çok etkilenenler. Birincil dili İngilizce olmayan ülkelerde YouTube Pişmanlıkları oranı yüzde 60 daha fazla.
- İngilizce dışındaki dillerde pandemiyle alakalı bildirimler oldukça fazla. Araştırma görevlilerimizin YouTube’de olmaması veya önerilmemesi gerektiğini tespit ettiği İngilizce YouTube Pişmanlıkları’ndan yüzde 14’ü pandemiyle alakalı. İngilizce olmayanlar içinse bu oran yüzde 36.

### Hikaye

2017 yılında Myanmar’daki Müslüman bir etnik azınlık olan Rohingya topluluğundan yaklaşık 700 bin kişi cinayetler, toplu tecavüzler ve yakılan köylerden kaçarak Bangladeş’e sığınmak zorunda kaldı. Aylar sonra, BM’ye bağlı Bağımsız Uluslararası Hakikati Araştırma Komisyonu, Facebook’un Rohingyalar ve diğer etnik azınlıkların yaşadığı Arakan eyaletinde

Myanmar ordusu tarafından işlenen suçları “kolaylaştıran bir ortam” yarattığını belirten rapor yayınladı. Reuters yaptığı bir araştırmada, “Rohingyalara veya diğer Müslümanlara köpek, solucan, tecavüzcü diyerek bu insanların domuzlara yedirilmesini öneren ve vurulmaları veya infaz edilmeleri çağrısında bulunan kötü niyetli gönderilere” ulaştı. Facebook’ta büyüüp yayılan nefret söyleminin, yaşanan soykırımın önemli bir öncüsü olduğu tespitine yer verildi. Reuters’ın belirttiği gibi, bu gönderilerin neredeyse tamamı bölgedeki birincil dil olan Birmanya dilindeydi ve hemen hemen hepsi Facebook politikalarını ihlal ediyordu.

Meselenin özü şu: Platformların politikaları ile bu politikaların uygulanması, birbirinden tamamen farklı şeyler. Ve ne yazık ki platform politikaları, dünyanın farklı yerlerinde farklı şekillerde uygulanıyor. Bunun bir sebebi, politika ihlallerini tespit etmek ve video önermek için kullanılan algoritmalar, dillere özgü makine öğrenimi modellerine dayanıyor. Bu da şirketlerin algoritmaları farklı dillerden ve ülke bağlamlarından verileri kullanarak eğitmeleri gerektiği anlamına geliyor. Gelgelelim, çoğu platform bu eğitimde İngilizce verilere öncelik verdiğinden, algoritmalar bu bağlamlarda daha iyi çalışıyor.

YouTube bu konuda veri paylaşmasa da öneri sistemlerinin, ihlale sebep olmamakla birlikte Topluluk Kuralları'nın sınırlarında gezen "ihlale yakın içeriklerin" yayılmasına sebep olduğunu kabul ediyor. Şirket, bu sorunu çözmek amacıyla yapılacak politika değişikliklerini duyurduğunda, ilk olarak ABD'ye ve İngilizce konuşulan diğer ülkelere odaklandı. İki yıldan uzun bir süre sonra bu değişiklikleri dahil oldukları tüm pazarlarda hayata geçirdiklerini duyursalar da, bildiğimiz kadarıyla bu girişimlerin ABD dışındaki yerlerde ne kadar başarılı olduğuna dair herhangi bir ölçüm yayınlamadılar.

Araştırma bulgularımıza göre, YouTube Pişmanlıkları İngilizce konuşulmayan ülkelerde yüzde 60 daha fazla. Dahası, araştırma görevlilerimizin YouTube'de olmaması veya önerilmemesi gerektiğini tespit ettiği içerikler arasında pandemiyle ilgili olanlara İngilizce dışındaki dillerde daha çok rastlanıyor.

Dünyanın dört bir yanındaki insanların, küresel bir salgın sırasında tıbbi bilgiler almak gibi önemli konularda YouTube'u kullandığı düşünüldüğünde, bu farklılık felakete yol açabilir. AB Dinsinfo Lab'in yakın tarihli bir raporu, Covid-19'a dair bir komplo teorisi içeren (ve İngilizce "Plandemic" videosuna benzeyen) "Hold Up" başlıklı Fransızca bir videonun, tıbbi dezenformasyon politikasını ihlal etmesine rağmen, yayımlandıktan altı ay sonra, milyonlarca izlenmeyle birlikte, hâlâ platformda olduğunu ortaya koydu. Bu durum, Election Integrity Partnership ile birlikte yürütülen ve 2020 ABD Başkanlık Seçimleri sırasında Twitter ve YouTube'da yer alan ve İngilizce dışındaki dillerde seçime dair dezenformasyon yapan içeriklerin etiketlenmediğini veya herhangi bir yaptırımla karşılaşmadığını, ayrıca YouTube'un politikalarını hangi dillerde uyguladığının net olmadığını ortaya çıkaran araştırmanın bulgularıyla örtüşüyor.

Bir insanın çevrimiçi ortamlardaki güvenliğinin, konuştuğu dile bağlı olmaması gerekiyor. YouTube gibi platformların, sadece İngilizce konuşulan ülkelerde yaşayan ve İngilizce konuşan insanlara karşı değil, platformlarını kullanan herkes karşı sorumluluğu var. Sistemin şeffaf olmayışı ve bu mesele karşısındaki ilgisizlik, platformların kendi politikalarını tutarlı bir şekilde nasıl uyguladığına (veya uygulamadığına) dair kullanıcıların fikir sahibi olması açısından RegretsReporter gibi şeffaflık araçlarının ne kadar önemli olduğunu gözler önüne seriyor.

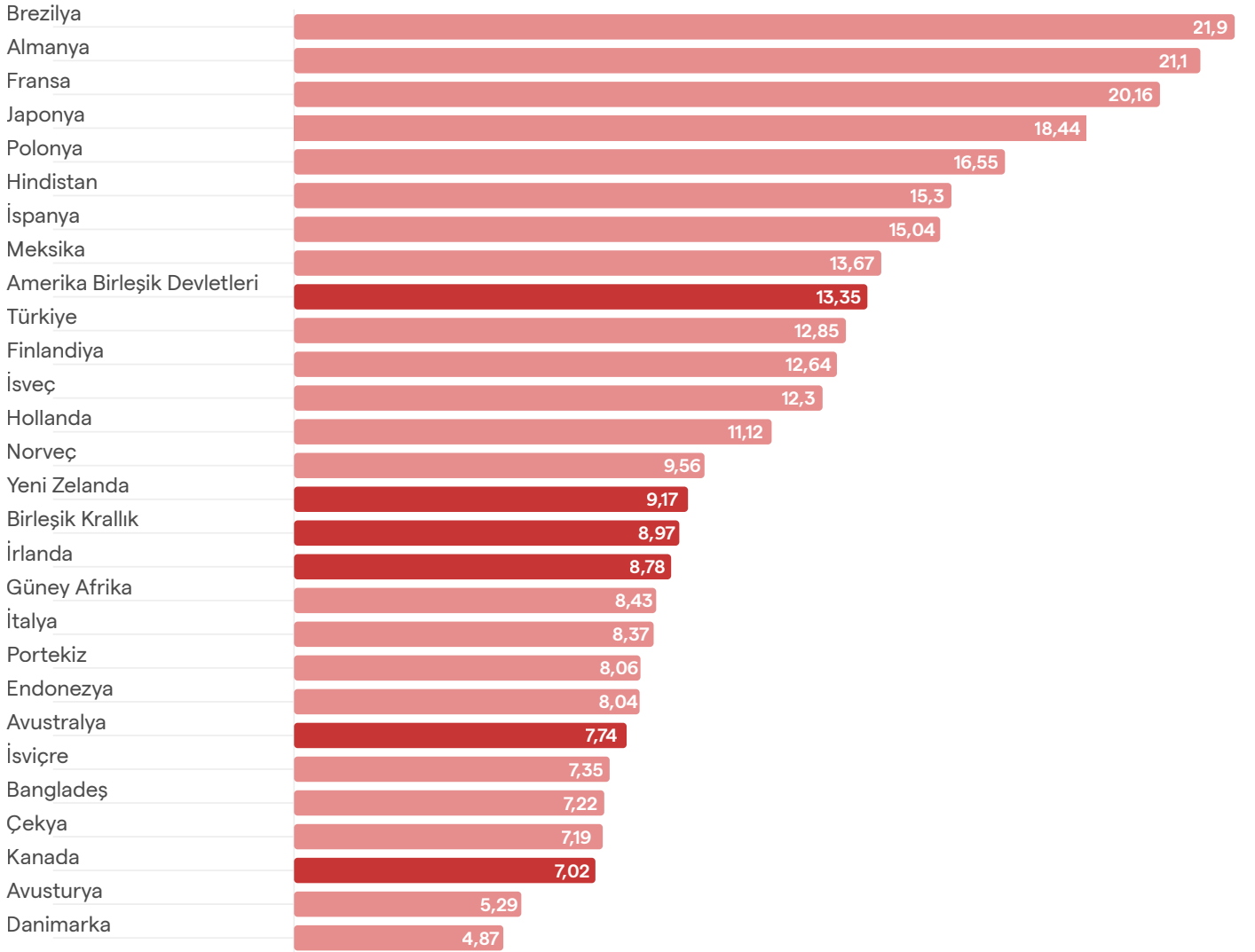
## Veriler

Ülkelere göre Pişmanlık oranlarını hesaplamak için, gönüllülerimizin YouTube'a hangi ülkeden eriştiğini gösteren GeoIP'den faydalandık. En yüksek Pişmanlık oranlarının, izlenen 10 bin video başına 20 ile Brezilya, Almanya ve Fransa'da olduğunu tespit ettik. Aslına bakılırsa, ilk yedi ülkede de birincil dil İngilizce değildi ve ilk 14'te yer alan tek İngilizce konuşulan ülke ABD'ydi. Pişmanlık oranlarının en düşük olduğu 14 ülke arasındaysa İngilizce konuşulan beş ülke vardı.

Birincil dili İngilizce olan ülkelerde Pişmanlık oranlarının düşük olduğu açıkça görülüyor. Daha net bir şekilde ifade edecek olursak, birincil dili İngilizce olan ülkelerde Pişmanlık oranı izlenen 10 bin başına 11 (yüzde 95 güven aralığıyla 10,4 ila 11,7). İngilizcenin birincil dil olmadığı ülkelerde Pişmanlık oranı izlenen 10 bin video başına 17,5 (yüzde 95 güven aralığıyla 16,8 ila 18,3). Yani, birincil dili İngilizce olmayan ülkelerde Pişmanlık oranı yüzde 60 daha fazla olduğundan, istatistiksel açıdan kayda değer bir fark olduğu görülüyor.

## Ülkelere göre YouTube Pişmanlıkları (%)

10000 video başına pişmanlık oranları



Hata çizgileri, yüzde 95'lik bir güven aralığında Pişmanlık oranını gösteriyor. Pek çok ülkedeki sınırlı veriler sebebiyle bu güven aralıkları bazı vakalarda oldukça geniş olsa da bulgular istatistiksel olarak kayda değer.

Sadece isabetli tahminler yapmamıza yetecek veriye sahip olduğumuz ülkelere yer verdiğimiz dikkatinizi çekeriz.

Ülkelerin birincil dillerini, CIA'in kılavuzunu kullanarak sınıflandırdık; mümkün olduğu durumlarda, anadil olarak en çok sayıda insanın konuştuğu dili, anadil bilgisine erişemediğimiz durumlarda da en yaygın dili birincil dil olarak belirledik. Sınıflandırmamız, birincil dilin ne olduğundan bağımsız olarak, verilerimizin büyük çoğunluğunu (Pişmanlık bildirimlerinin yüzde 94'ünü) sağlayan ülkelerin hepsini kapsıyor.

Pandemiyle alakalı içerikler, Pişmanlık bildirimlerinin önemli bir kısmını oluşturuyordu ve özellikle İngilizce olmayan videolarda bu türden içeriğe sık rastlanıyordu. Araştırma görevlilerinin YouTube’de olmaması veya önerilmemesi gerektiğini tespit ettiği videoları analiz ettiğimizde, İngilizce videoların sadece yüzde 14’ü pandemiyle ilgiliydi. Fakat İngilizce olmayan pişmanlık verici videolarda bu oran yüzde 36’ydı; İngilizce olmayan dillerdeki bu içeriklerin üçte birinden fazlası, pandemiyle ilgiliydi.

Örneğin, “Yolsuz Biden, aşının zorunlu hale gelmesi ve fazlasıyla histeri!” başlıklı Portekizce bir videodaki kişi, corona virüs hakkındaki komplo teorilerini savunuyor.

#### Bildirilen video



Biden corrupto, judicialização da vacina e chega de histeria!

6 views - Stream iniciado há 57 minutos

Bize bildirilen bu Slovakça video da pandeminin yalan ve kandırmaca olduğunu iddia ediyor. Bize bildirildiği sırada bu video 300 binden fazla kez izlenmişti.

#### Bildirilen video



Celoplošné testovanie: fiasko a konflikt od prvej chvíle  
300385 views - 19 Oct 2020

Pandemi hakkındaki Yunanca bir konuşmanın ses kaydındaki kişi, pandeminin Yunanistan hükümetinin kendi çıkarlarına hizmet etmek için uygun bir bahane olduğunu öne sürüyor.

#### Bildirilen video



ΧΡΥΣΗ ΑΥΓΗ ΚΑΣΙΔΙΑΡΗΣ ΚΟΡΩΝΟΙΟΣ  
ΝΤΟΚΟΥΜΕΝΤΑ & ΑΛΗΘΕΙΑ στην ΕΠΙΔΗΜΙΑ ΑΠΑΘΗ  
φοβου των ΚΑΝΑΛΙΩΝ 2η

1916 views - 5 Απρ 2020

Bu bulguları, içerik bildirme veya internet kullanımına dair kültürel farklar da dahil pek çok sebeple açıklamak mümkün. Ne var ki verilerimiz, YouTube’un pişmanlık uyandıran içerikler konusunda İngilizce konuşulmayan pazarlarda daha büyük bir sorunu olduğunu işaret ediyor.





# Öneriler

1- YouTube ve diğer platformlara  
tavsiyelerimiz

2- Politikacılara önerilerimiz

3- YouTube kullanıcılarına önerilerimiz

# 01.

## YouTube ve diğer platformlara önerilerimiz

- **Öneri sistemlerinin bağımsız denetimine izin verin.**

Hakiki bir denetime açıklık için, gözlem kuruluşlarının ve kamu çıkarı araştırmacılarının YouTube gibi platformların “kaputunun içine bakabilmesi” gerekir; öneri sistemlerini ilgilendiren tasarımların ve işleyiş biçimlerinin çevrimiçi tehlikelere nasıl kapı açtığını daha iyi anlayıp değerlendirmenin tek yolu bu.

Bu amaçla, YouTube ve diğer platformlar, araştırmacıların öneri sistemlerinden kaynaklanan risk ve tehlikeleri inceleme çabalarını desteklemek için ortak bir girişimde bulunmalı. Bu veri erişimi düzenlemelerinin sağlam olması ve bugün YouTube gibi platformların sunduğundan çok daha kapsamlı olması elzem, çünkü raporumuzun da gösterdiği gibi çok fazla boşlukla karşılaşıyoruz.

Kısmi ve yüzeysel bilgiler yerine -veri güvenliğinin sağlanması ve güvenlik protokollerine uyulması koşuluyla- araştırmacıların öneri sistemlerini test etmesi, öneri sistemlerince optimize edilen ölçümlere dair bilgilere sahip olması, öneri sistemlerinin kaynak koduna erişmesi, makine öğrenimi modellerini eğitmek için kullanılan eğitim verilerine ulaşması ve güvenlik ve koruma uygulamalarına dair belgeleri gözden geçirmesi sağlanmalı.

Bu türden bir şeffaflık, öneri ekosistemindeki üstü örtük tehlikeleri tespit etmemize ve bunları bertaraf etmek için ticari ve düzenleyici çözümler geliştirmemize imkan tanıyacaktır.

- **Öneri sistemlerinin nasıl işlediğine dair bilgilerin yanı sıra şeffaflık raporları yayınlayarak sorunlu alanlara ve bu alanlardaki ilerlemelere dair bilgi verin.**

Öneri sistemleri, YouTube gibi platformların içerik sunmak için kullandığı en önemli mekanizma olmanın yanı sıra güvenlik ve koruma tedbirleri açısından da en kritik alan. YouTube’un şeffaflık raporlarının, zararlı içeriklerin yayılmasında öneri sistemlerinin oynadığı büyük rolü göz önünde bulundurarak, içerik moderasyonu önlemleri ile öneri sistemleri arasındaki ilişkiye dair anlamlı bilgiler sunması şart. Buna araştırmacıların, öneri sistemlerinin zararlı içeriğin yayılmasına nasıl katkı sunduğunu anlamasına ve YouTube’un zararlı içeriklerin önerilmesini azaltmak için attığı adımların gerçekten iddia ettikleri gibi işe yarayıp yaramadığını bağımsız olarak tespit etmesine yardımcı olacak detaylı veriler de dahil.

YouTube şu anda ihlale yakın içerikleri (Borderline contents), yani topluluk kurallarını tam olarak ihlal etmese ihlal sınırlarına yaklaşan içerikleri nasıl tanımlayıp yönettiğine dair hiçbir bilgi vermiyor. YouTube’un harekete geçip bu şeffaflık eksikliğini gidermesi gerek. YouTube’un şeffaflık raporlarının kapsamını genişleterek platformun “ihlale yakın içerik” tanımına ve bu türden içeriklerin moderasyonu için hangi metodolojileri (örneğin sıralama düşürme; önceliğini azaltma) kullandığına dair bilgilerin yanı sıra, bu kategorideki içeriklerle

alakalı sorunların değerlendirilmesine yardımcı olabilecek bütünlüklü verileri (örneğin YouTube'un ihlale yakın içerikleri kaç kere önerdiği ve platformda bu türden ne kadar içerik olduğu) bu raporlara dahil etmesi gerek. Bu bilgilerin -Hizmet Koşulları arasında gömülmek yerine- ayrı bir bilgilendirme sayfasında kapsamlı ve anlaşılır bir biçimde sunulması da önemli. Bölgeden bölgeye yaşanan farklılıkları denetleyebilmek için de bu bilgilerin ülke/bölge ve dillere göre filtrelenmesi gerek.

- **İnsanlara verilerinin öneri sunmak için bir girdi olarak nasıl kullanıldığı ve bu önerilerin ne olacağı konusunda daha çok kontrol verin.**

Platformların, hangi verilerinin öneri sunmak için kullanılacağı konusunda insanlara daha çok kontrol vermesi gerekiyor. Ayrıca insanların, öneri yapılırken göz önünde bulundurulmuş diğer bilgilere dair fikir ve kontrol sahibi olması gerekiyor. Örneğin, insanların söz konusu platformla alakalı diğer ürün/ hizmetlerden (mesela YouTube önerilerinde kullanmak için Google verilerine başvurmak), belli içerik/sayfa/kullanıcılarla daha önce girilen etkileşimlerden ve diğer kullanıcılardan kendilerine öneri sunmak için veri toplanmasını engelleme seçeneği olmalı.

Platformlar daha fazla kontrol ve seçenek sunarak insanların önerileri ve kendilerine gösterilen içerikleri kendi tercihlerine göre belirlemesini sağlamalı ki kullanıcılar platformda güvenlikleri daha iyi sağlayabilsin. Örnek vermek gerekirse, bu kontrollere belli anahtar kelimeleri, belli içerik türlerini veya kanalları öneriler arasından çıkarma seçeneği dahil edilmeli. Ayrıca, "ihlale yakın içerik" ve platformdaki belli içerik kategorilerinin (örneğin haber, spor) önerilerde yer alıp almayacağı veya ne kadar yer alacağını kullanıcı kendisi için belirleyebilmeli.

Bu kontrollere ana kullanıcı ayarlarından erişilebilmeli ve bu kontroller, algoritmaya dayalı önerileri gösteren arayüze eklenmeli. Mümkün olduğu durumlarda platformlar, bir kontrolün gönderdiği mesajı belirtmek yerine, o kontrolü kullanmanın sonucunu açık bir dille ifade etmeli (örneğin, "Bu öneriyi beğenmedim" yerine "Bu kanaldan önerileri engelle" gibi).

- **Öneri sistemlerine adanmış titiz ve belli aralıklarla devreye giren risk yönetimi programları kullanın.**

Platformlar, öneri sisteminin tasarımından, işleyişinden ve kullanımından doğabilecek, bireylere ve kamu çıkarına yönelik riskleri sürekli olarak tespit etmeli, değerlendirmeli ve yönetmeli. Böyle bir risk değerlendirmesinin, hem zarar doğma ihtimalini hem de bu zararın muhtemel büyüklüğünü de göz önünde bulundurması gerek. Risk yönetimi konusunda daha kapsamlı bir yaklaşım benimsemeleri durumunda platformlar hem işleri hem de itibarlarının karşı karşıya olduğu riskleri daha iyi değerlendirmenin yanı sıra hizmetlerinin kasıtlı veya kasıtsız olarak yol açtığı ikincil sorunları daha iyi anlayabilir.

- **İnsanlara kişiselleştirmeyi devredışı bırakma seçeneği verin.**

YouTube da dahil tüm platformlar, insanlara kişiselleştirilmiş öneriler yerine kronolojik, bağlamsal veya sadece arama bazlı öneriler alma seçeneğini sunmalı. Bir hizmete erişmek için, önerileri (veya insanlar tarafından yapılan herhangi bir kişiselleştirilmiş seçeneği) görmek zorunlu olmamalı.

# 02.

## Politikacılara önerilerimiz

- **YouTube ve diğer platformlardan bilgi paylaşımlarını ve denetimler ve veri erişimi şartları aracılığıyla araştırmacıların öneri sistemlerini incelemesinin yolunu açmalarını talep edin.**

Politikacıların, platformların öneri algoritmalarını şeffaf hale getirmesini ve sosyal medya platformları üzerinde bağımsız araştırmalar yapılmasını sağlayacak titiz veri erişimi çerçeveleri oluşturmasını şart koşacak düzenlemeler getirmesi gerekiyor. Avrupa Komisyonu, Dijital Hizmetler Yasası için sunduğu bir öneride bu talebe zaten yer verdi, ayrıca başka bölgelerde görev alan politikacılar da benzer şeffaflık ve denetim mekanizmalarını hayata geçirmek istediklerinin sinyalini verdi. Politikacıların, YouTube ve benzeri platformların bu çok ihtiyaç duyulan şeffaflığı hayata geçirmek için kendi kendilerine harekete geçmediğini ve bu duruma düzenlemeler yoluyla müdahale edilmesi gerektiğini anlamalı.

- **Politika çerçevelerinin, içerik öneri sistemlerine has sorun ve risklere uygun olduğundan emin olun.**

Politikacıların, çevrimiçi içerik sorumluluğuna dair çerçeveler hazırlarken söz konusu düzenlemelerin içerik önerme sistemleri için güvenlik ve koruma sorumluluğunu teşvik ettiğinden emin olmalı. Bunu sağlamak için benimsenebilecek birkaç yaklaşım var (örneğin risk değerlendirmesi yapmayı veya son kullanıcı kontrollerini geliştirmeyi zorunlu

kılmak), fakat genel bir ilke olarak çevrimiçi içerik sorumluluğuyla ilgili düzenlemelerin, platformların içerikleri yaygınlaştırmak için kullanılan otomatik sistemleri tasarlar ve çalıştırırken bu sistemlerin doğurabileceği riskleri hesaba katmasını sağlamalı.

- **Platformları araştırmak için alternatif yöntemler kullanan araştırmacıları, gazetecileri ve diğer gözlemcileri koruyun.**

Politikacılar, platformların sağladığı veri erişim kanallarının dışına çıkarak kamu çıkarı yararına araştırma yapan araştırmacıları hukuki tehditlerden koruyacak hükümleri veya başka koruma yöntemlerini hayata geçirmeli. Bu sayede platformların veri erişimine getirdiği engeller aşılabilir. Bu doğrultuda, mevcut yasalar değiştirilmeli veya netleştirilmeli. Örneğin, AB'nin Genel Veri Koruma Mevzuatı'nda yer alan araştırma istisnalarının kapsamı oldukça belirsiz; ayrıca ABD Yüce Mahkemesi yakın zamanda verilen van Buren kararını teşvik etse de, ABD Çevrimiçi Dolandırıcılık ve İstismar Yasası'nın kamu çıkarı için yapılacak araştırmalara nasıl uygulanacağı konusunda cevaplanmamış sorular var. Bu belirsizliği gidermek için, koruma hükümleri platformların araştırma için kullanılan araçları engellemesini veya sınırlandırmasını yasaklayabilir. Bu türden hükümler platformlardan veri toplayan veya denetimlerini sahte hesap açarak yapan araştırmacıları daha iyi koruyabilir; sahte hesap açmak, şu an platformların Hizmet Şartları'nı ihlal eden bir eylem.

# 03.

## YouTube kullanıcılarına önerilerimiz

- **YouTube önerilerinin nasıl işlediğini öğrenin.**

Algoritmaya dayalı öneriler sunan YouTube ve diğer platformların nasıl işlediğine dair kapsamlı bilgiler için Mozilla'nın "Mozilla Explains: Recommendation Engines" ve "Mozilla Explains: Recommendation Engines part 2" videolarını izleyin.

- **Kendiniz ve aileniz için doğru kontrolleri uyguladığınızdan emin olmak için YouTube ve Google'daki veri ayarlarınıza göz atın.**

YouTube ana sayfasından erişmek pek kolay olmadığı için pek çok insan bu ayarlardan haberdar değil. Bu konuda uzman tavsiyemiz şu: "izleme" ve "arama" geçmişinize bakıp önerilerinizi etkilemesini istemediğiniz videoları kaldırın veya bu özelliği tümünden kapatın. Bilgisayarınızı başkalarıyla ortak kullanıyorsanız bunu yapmak bilhassa faydalı olacaktır. Ayrıca bu sayfadan cep telefonu, televizyon ve bilgisayarda YouTube'un otomatik oynatma seçeneğini nasıl kapatacağınızı öğrenebilirsiniz.

- **Kitle kaynaklı araştırmamıza verilerinizle katkı sunmak için RegretsReporter'ı indirin!**

Mozilla, YouTube'un öneri algoritmasını incelemek için bağımsız bir araç olan RegretsReport'ı aktif tutmaya devam edecek. İnsanların istenmeyen önerileri engelleyebilmesi için YouTube'un kullanıcı kontrollerine daha kolay erişmesini sağlayacak bir güncelleme yapmayı da planlıyoruz. Ayrıca yeni analizler ve öneriler yayınlamaya da devam edeceğiz.

# Sonuç

YouTube algoritması, bireyler ve toplumlar olarak inandıklarımızı şekillendirmede muazzam bir rol oynuyor. Mozilla ve sayısız uzman tarafından yapılan arařtırmalar, YouTube algoritmalarının kayda deęer zararlara yol ađtıđını doęruladı. Eęer bu deneyimler, arařtırmamızın ortaya koyduęu kadar sık yařanıyorsa ve dięer yapay zeka sistemlerinde de ortaya çıkıyorsa (Google'ın öneri yapay zekasının ticari vaatleri düşünöldüęünde gerçekçi bir çıkarım bu), sonuçların ciddi olacaęı kesin.

Pek çok uzmana göre bu sorunlar aslında hata deęil; daha ziyade, YouTube'un algoritmasının tam olarak gerektięi gibi çalışmasının bir sonucu ve ticari faydaları artırmak için geliştirilen algoritmalar ile insanların iyilięi için geliştirilen algoritmalar arasında temel bir uyumsuzluk bulunuyor. Bunun doęru olma ihtimali var. Doğruluęundan emin olduęumuz bir gerçek ise, bu derece ciddi etkileri olan algoritmaların gerekli denetimler olmadan kullanılmaması gerektięi. Bu yönde atılacak ilk adım da kesinlikle şeffaflık.

Doęası gereęi öznel olan "ihlale yakın" içerikleri kodlarla da kamu politikalarıyla da yönetmek zor. Fakat tam da bu zorluk sebebiyle bu türden içerikleri incelemek, tartıřmak ve anlamaya çalışmak son derece önemli. Biz, bu incelemeleri yapmamıza imkan verecek araçları geliřtirebilecek gerekli beceri ve kaynaklara sahip olduęumuz için yapabildik bunu ama aslında buna mecbur kalmamız gerekirdi. YouTube şeffaf olmayı seçip bizimki gibi arařtırmalara alan açabilir ama bunu yapmamayı tercih ediyor. Bunun yerine, bahsettięimiz türden etki ve sonuçları kendi kendilerine idare etmeye çalışarak kullanıcılar için kırılğan bir yapı inşa ettiler. YouTube'un Silikon Vadisi'ndeki toplantı odalarında verdięi kararlar, dünyanın her yerindeki insanları etkiliyor ve etkilenen bu insanlar alınan kararlara dair bilgi ve söz hakkına sahip olmayı hak ediyor.

YouTube, bu sorumluluęu üstlenme konusunda yanlış yaklařımı seçtięini açıkça gösterdi. Doğru yaklařıma ulařmanın yolu daha şeffaf, denetime açık ve saygılı olmaktan geçiyor.

# Metodoloji

# Metodoloji

## • Araştırma Soruları

Araştırmamız, aşağıdaki sorularla başladı:

1. *Hangi videolar, pişmanlık verici olarak bildirildi?*
2. *Pişmanlık verici olarak bildirilen videolar hangi kategorilere giriyor?*
3. *Bildirilen pişmanlık verici içeriklerin sıklığı ve ağırlığına ilişkin örüntüler var mı?*
4. *Pişmanlık verici videoları görmeye ve bildirmeye yol açan belirli YouTube kullanım örüntüleri var mı?*
5. *Pişmanlık oranları ve türleri, coğrafyadan coğrafyaya değişiyor mu?*
6. *Bildirilen pişmanlık verici deneyimlerin sıklığı nedir ve bu durum giriş noktasına göre nasıl değişiyor?*
7. *İlk bildirim yaptıktan sonra insanların bildirme sıklığı değişiyor mu ve bu kategori ve ağırlığa göre nasıl değişiyor?*
8. *YouTube Pişmanlıkları ve bildirimlerden önceki izleme geçmişinin niteliksel özellikleri neler? Tespit edilebilir örüntüler var mı?*

Veri setimizle sadece 1, 2, 5, 6 ve 8. sorulara yanıt bulabildik.

## • RegretsReporter Uzantısı

Firefox ve Chrome'da kullanılabilen bir tarayıcı uzantısı olan RegretsReporter, gönüllülerin YouTube'da izledikleri pişmanlık uyandırıcı videolara dair verileri Mozilla'ya göndermelerine imkan sağlıyor. RegretsReporter'ı yükleyen gönüllüler,

YouTube'daki bir videonun sayfasındayken tarayıcılarındaki ikona basarak bir bildirim formuna ulaşıyorlar. Raporla deneyimlerine dair birkaç soru soruluyor ve Mozilla'ya gönderilecek veriler hakkında bilgi veriliyor.

Rapor gönderildiğinde Mozilla videonun başlığı, açıklaması, izlenme sayısı, ulaşma yolu (öneri, arama vs.) ve videonun sayfasına yönlendiren bağlantı gibi bilgilerin yanı sıra gönüllünün isterse sağlayabileceği kategori, yorum ve ağırlık derecesi gibi ilave verileri alıyor. Bu da bildirilen deneyimi anlamamıza ve ne tür videoların pişmanlık uyandırdığını görmemize yardımcı oluyor.

Mozilla ayrıca gönüllü istediği takdirde, bildirilen videoya nasıl ulaştığını gösteren bir "rota" bilgisi de alabiliyor. Bu rota, bildirim yapılmadan önce en fazla beş saat içinde YouTube'da açılan en fazla son beş sayfayı içeriyor. Bu bilgi, gönüllüyü bildirilen videoya ulaştıran öneriler zincirini ve diğer eylemleri analiz etmemize imkan veriyor. RegretsReporter, Pişmanlık oranlarını ve sıklığını hesaplayabilmek için gönüllülerin



YouTube kullanımına dair veriler de alıyor. Bu veriler hangi videoların izlendiğine dair detaylar veya YouTube'un tam olarak ne zaman kullanıldığı bilgisini değil, toplamda kaç videonun izlendiğini gösteren genel bir ölçüm içeriyor. Bu kullanım verileri, birkaç unsura bağlı olan Pişmanlık oranlarını hesaplayabilmek için elzem. Son olarak, Mozilla'ya gönderilen her veri, GeolP'den faydalanarak gönüllünün YouTube'u hangi ülkeden ziyaret ettiği bilgisini içeriyor; bu sayede Pişmanlık oranlarının ülkeden ülkeye nasıl farklılık gösterdiğini analiz edebiliyoruz.

RegretsReporter açık kaynaklı bir yazılım olduğundan isteyen herkes koda erişebilir. RegretsReporter, Mozilla'nın yalın veri ilkelerine göre geliştirildi ki bu da sadece araştırma hedeflerimize ulaşmamız için gerekli verileri topladığı anlamına geliyor. RegretsReporter, onu kullanan gönüllülerin şahsi verilerini toplamıyor veya depolamıyor.

## ● Bireysel katılıma dayalı veri seti

YouTube araştırmacılarla veri paylaşmadığından, araştırma sorularımızı cevaplamamıza yarayacak verileri toplamak için yurttaş katılımına dayanan bir bilimsel yöntem geliştirdik. Eski araştırmalardan dayanak ve ilham alan RegretsReporter, YouTube'un öneri sistemine dair bugüne kadar yapılmış en büyük kitle kaynaklı araştırma. Veri setimiz, Firefox ve Chrome için RegretsReporter uzantısını yükleyen, 190 ülkeden 37 bin 380 gönüllünün çabalarına dayanıyor. Tüm gönüllülerimiz arasından bin 662'si en az bir kere bildirimde bulundu; Temmuz 2020 ile Mayıs 2021 arasında 91 ülkeden toplam 3 bin 362 bildirim yapıldı. Uzantıyı indirip hiç bildirimde bulunmayan gönüllüler de araştırmamızın önemli bir parçasıydı. Onların verileri de -mesela YouTube'u ne kadar kullandıkları- YouTube'daki pişmanlık verici deneyimlerim sıklığını ve ülkeden ülkeye nasıl değiştiğini anlamamız için asli bir unsurdu.

Bireysel katılıma dayalı bu yaklaşım, YouTube kullanan insanların gerçek deneyimlerini yakalıyor ve YouTube'un araştırmacılara veri sağlamamasına rağmen algoritmasına dair bilgi sahibi olmamıza yarayacak son derece önemli bir işlevi yerine getiriyor. Ne var ki yaklaşımımızın metodolojik sınırlamaları da yok değil. Bunlardan bazıları şöyle:

**Seçme yanlılığı:** Gönüllülerimiz belli bir grup insandan oluşuyor ve bulgularımız tüm YouTube kullanıcıları için geçerli olmayabilir.

**Bildirme yanlılığı:** Bir gönüllünün bir videoyu bildirip bildirmeyeceğini belirleyen pek çok faktör olabilir.

**Pişmanlık tanımı:** YouTube Pişmanlık tanımı (kasten) net bir şekilde belirlenmedi ve farklı kullanıcılar bir videoyu farklı "pişmanlık" kavramlarına dayanarak bildirebilir. Çalışmanın gözleme dayalı yapısı, "ne" olduğunu emin bir şekilde ortaya koyabilmemize rağmen "nedenlere" dair aynı güvenle çıkarım yapamadığımız anlamına geliyor. Örneğin, YouTube'un belli bir videoyu belli bir kullanıcıya neden gösterdiğini bilmiyoruz.

Bu sınırlamalara rağmen, ulaştığımız sonuçlar dünyanın dört bir yanından insanların gerçek deneyimleri sayesinde YouTube'daki problemleri anlamamızı sağlıyor. Araştırmamızın sınırları, YouTube'un araştırmacılara veri sağlamasının gerekliliğini bir kez daha vurguluyor.

## ● Analiz Yöntemleri

RegretsReporter uzantısı, toplanan verileri Mozilla'nın uzaktan ölçüm sistemi aracılığıyla iletiyor ve sonra da Mozilla'nın veri havuzunda depoluyor. Analiz, çoğunlukla Google Colab ortamında, BigQuery ve Python kullanılarak yapılıyor. Analiz koduna buradan erişebilirsiniz.

Araştırma sorularımızı cevaplamak için ikincil araştırma, istatistiksel yöntemler ve niteliksel analizden faydalanarak analizimizde birkaç farklı yöntem kullandık.

İstatistiksel yöntemler sayı, orantı ve oranları hesaplarken, ilgili değişkenlere göre bölümlendirme yapmayı içeriyordu. Güven aralıklarının hesaplanması ve hipotez testlerinin değerlendirilmesi de dahil, çıkarım teknikleri uygulandı. Bu raporda yer verilen tüm farklılıklar, istatistiksel olarak  $p < 0.05$  düzeyinde değerliydi ve tüm istatistiksel değer, bu düzeyde değerlendirildi.

Niteliksel analiz sürecine çevrimiçi tehlikeler, ifade özgürlüğü ve teknoloji politikaları alanlarında çalışmış uzmanların (raporun Teşekkür kısmında bu uzmanların isimlerini görebilirsiniz) yer aldığı bir ekip oluşturarak başladık.

Ekibin görevi, veri setindeki bildirilmiş videoları izleyerek temaları tespit etmektir. Bu ekip üç ay boyunca, YouTube Topluluk Kuralları'nı baz alarak bazı videoları sınıflandırmaya yarayacak kavramsal bir çerçeve geliştirdi. Niteliksel analize kılavuzluk etmesi için YouTube Topluluk Kuralları'nın seçilme sebebi, hem sorunlu video içeriklerine dair faydalı bir sınıflama sağlaması hem de YouTube'un ne türden içeriklerin platformda yer alacağına dair vaadini temsil etmesiydi.

Exeter Üniversitesi tarafından seçilen (ve isimleri bu raporun Teşekkür kısmında yer alan) 41 araştırma görevlisinden oluşan ekip, bildirilen videoları tek tek değerlendirmek için bu kavramsal çerçeveden faydalandı. Bu çerçeveye göre araştırma görevlilerinin aşağıdaki gibi soruları yanıtlaması gerekiyordu:

- Sizce bu video YouTube'da olmalı mı?
- Sizce bu video Şimdi İzle kısmında veya YouTube ana sayfasında önerilmeli mi?

- Eğer bu sorulardan birine HAYIR cevabını verdiyseniz: Sizce bu video Topluluk Kuralları kategorilerinden hangisi ihlal ediyor olabilir?

Araştırma görevlileri ayrıca (mümkün olduğunda) pişmanlık uyandırıcı videoya uzanan öneri rotasının niteliksel özelliklerini de analiz etti ve videonun birincil dili veya Covid-19 ile ilgili olup olmadığı gibi ekstra sorulara da cevap vererek analizimizin diğer unsurlarına katkı sağladı.

Araştırma görevlileri 7-15 Haziran tarihleri arasında bin 141 bildirim, yani toplam 3 bin 362 bildirimden yüzde 33,9'unu analiz etti.

Bildirim verileri, önce belirli bir beta grubuyla 22 Temmuz 2020'de toplanmaya başladı, daha sonra Eylül 2020 itibarıyla de herkese açık hale getirildi. Analiz için veri toplamaya 31 Mayıs 2021'de son verildi. Veri temizleme için iki prosedür izlendi:

**Eylemsizlik kriteri.** Kimi gönüllüler, uzantı hâlâ yüklü olmasına rağmen bildirim özelliğini kullanmadı. Uzantıyı yüklemenin amacı bildirim yapmak olduğundan, son bildirim yaptıkları veya hiç bildirim yapılmadıysa uzantıyı indirdikleri zamandan sonraki 56 gün içinde bildirim yapmadıysa bu gönüllüleri eylemsiz kabul ediyoruz. Bu 56 günlük süre bittikten sonra gönüllünün kullanım verileri sayılmıyor. Gönüllü eylemsiz olduktan sonra yeniden bir bildirimde bulunduğu zaman yeniden aktive edilip tüm kullanım verileri sayılmaya başlanıyor.

Gönüllülerin yüzde 91'i RegretsReporter'ı indirdikten sonraki ilk 56 gün bildirim yaptığından bu sonlandırma süresinin uygun olduğunu düşünüyoruz.

**Sapma kriteri.** İki gönüllü uzantıyı, genel gönüllü nüfusumuzla tutarlı olmayan biçimde kullandı.

Bu iki gönüllü 231 ve 109 bildirimde bulundu; onlardan sonraki en aktif gönüllülerin bildirim sayısı ise 65, 53 ve 37'ydi. Bu gönüllülerin verilerini eklersek sonuçlarımızın genelleştirilebilme niteliğinin zayıflayacağını düşündüğümüzden bu verileri analize dahil etmedik.

## • Açıklama

Mozilla karbonsuz olma konusunda kararlı ve sera gazı ayak izimizi de Paris İklim Anlaşması'ndaki salım dengesine uygun olarak her yıl azaltacağız (bu anlaşmanın gereklerinin ötesine geçecek hedeflerimiz de var). RegretsReport'u saklamak ve sorguları yürütmek için Google Cloud Platform kullanıyoruz. GCP yüzde 100 karbonsuz ve RegretsReporter da birincil olarak Oregon bazlı GCP altyapısını kullanıyor ki bu da yüzde 89 karbonsuz enerji kullanıyor. RegretsReporter ile bağlantılı karbon salımının tam kapsamı hesaplanacak ve bu bilgiler Mozilla'nın 2022'de yayınlanacak bir sonraki Sera Gazı Envanteri'nde paylaşılacak.

Bu proje için toplanan ve işlenen tüm veriler, ya Mozilla çalışanları ya da sözleşmeyle Mozilla'nın gizlilik, güvenlik, veri işleme ve mahremiyet politikalarından sorumlu araştırmacılar tarafından yönetildi.

İçerik moderasyonu ve bununla alakalı işlerin psikolojik etkilerine dair araştırmalardan haberdarız. Analizimizi desteklemek için bildirilen videoların büyük çoğunluğunu izleyen Exeter Üniversitesi araştırma görevlileri de dahil, bu araştırmada çalışan ekiplere araştırma süresince psikolojik destek sağlandı.

Google, dünyanın pek çok yerinde Firefox'taki varsayılan arama motoru. Bu ilişkiye rağmen, Mozilla yıllardır YouTube'dan ve diğer çevrimiçi platformlardan daha kapsamlı bir şeffaflık ve denetime açıklık talep ediyor.

# Kaynakça

- » Adamczyk, Roman. "What's the Hold-up? How YouTube's inaction allowed the spread of a major French COVID-19 conspiracy documentary." EU DisinfoLab, 12 Mayıs 2021, <https://www.disinfo.eu/publications/whats-the-hold-up%3F-how-youtubes-inaction-allowed-the-spread-of-a-major-french-covid-19-conspiracy-documentary/>
- » Alexander, Julia. "YouTube claims its crackdown on borderline content is actually working." The Verge, 3 Aralık 2019,
- » <https://www.theverge.com/2019/12/3/20992018/youtube-borderline-content-recommendation-algorithm-news-authoritative-sources>
- » Bergen, Mark. "YouTube Executives Ignored Warnings, Letting Toxic Videos Run Rampant." Bloomberg, 2 Nisan 2019,
- » <https://www.bloomberg.com/news/features/2019-04-02/youtube-executives-ignored-warnings-letting-toxic-videos-run-rampant>
- » Boyd, Ashley. "Senate Hearing Confirms YouTube Won't Fully Release Recommendations Data Without More Pressure from Public and Congress." Mozilla, 28 Nisan 2021, <https://foundation.mozilla.org/en/blog/senate-hearing-confirms-youtube-wont-fully-release-recommendations-data-without-more-pressure-from-public-and-congress/>
- » Bridle, James. "Something is wrong on the internet." Medium, 6 Kasım 2017,
- » <https://medium.com/@jamesbridle/something-is-wrong-on-the-internet-c39c471271d2>
- » Campbell, Eliza ve Spandana Singh. "The flaws in the content moderation system: The Middle East case study." Middle East Institute, 17 Kasım 2020,
- » <https://www.mei.edu/publications/flaws-content-moderation-system-middle-east-case-study>
- » Chen, Annie Y., Brendan Nyhan, Jason Reifler, Ronald E. Robertson ve Christo Wilson. "Exposure to Alternative & Extremist Content on YouTube" ADL, Şubat 2021,
- » <https://www.adl.org/resources/reports/exposure-to-alternative-extremist-content-on-youtube>
- » Cobbe, Jennifer ve Jatinder Singh. "Regulating Recommending: Motivations, Considerations, and Principles." European Journal of Law and Technology (EJLT) 10(3), 30 Aralık 2019, <https://ejlt.org/index.php/ejlt/article/view/686>
- » Cook, Jesselyn ve Sebastian Murdock. "YouTube Is A Pedophile's Paradise." HuffPost, 20 Mart 2020,
- » [https://www.huffpost.com/entry/youtube-pedophile-paradise\\_n\\_5e5d79d1c5b6732f50e6b4db](https://www.huffpost.com/entry/youtube-pedophile-paradise_n_5e5d79d1c5b6732f50e6b4db)
- » Córdova, Yasodara, Adrian Rauchfleisch ve Jonas Kaiser. "The implications of venturing down the rabbit hole." Internet Policy Review, 27 Haziran 2019,

- » <https://policyreview.info/articles/news/implications-venturing-down-rabbit-hole/1406>
- » Fisher, Max ve Amanda Taub. "How YouTube Radicalized Brazil." The New York Times, 11 Ağustos 2019,
- » <https://www.nytimes.com/2019/08/11/world/americas/youtube-brazil.html>
- » Fisher, Max ve Amanda Taub. "On YouTube's Digital Playground, an Open Gate for Pedophiles." The New York Times, 3 Haziran 2019,
- » <https://www.nytimes.com/2019/06/03/world/americas/youtube-pedophiles.html>
- » Geurkink, Brandi. "Congratulations, YouTube... Now Show Your Work." Mozilla, 5 Aralık 2019,
- » <https://foundation.mozilla.org/en/blog/congratulations-youtube-now-show-your-work/>
- » Geurkink, Brandi. "Our recommendation to YouTube." Mozilla, 14 Ekim 2019, <https://foundation.mozilla.org/en/blog/our-recommendation-youtube/>
- » Leerssen, Paddy. "The Soap Box as a Black Box: Regulating Transparency in Social Media Recommender Systems." European Journal of Law and Technology (EJLT) 11(2), 31 Ekim 2020, <http://www.ejlt.org/index.php/ejlt/article/view/786>
- » Lewis, Rebecca. "Alternative Influence: Broadcasting the Reactionary Right on YouTube." Data & Society, Eylül 2018,
- » [https://datasociety.net/wp-content/uploads/2018/09/DS\\_Alternative\\_Influence.pdf](https://datasociety.net/wp-content/uploads/2018/09/DS_Alternative_Influence.pdf).
- » Maréchal, Nathalie ve Ellery Roberts Biddle. "It's Not Just the Content, It's the Business Model: Democracy's Online Speech Challenge." Ranking Digital Rights, 17 Mart 2021, <https://rankingdigitalrights.org/its-the-business-model/>
- » Molter, Vanessa. "Platforms of Babel: Inconsistent misinformation support in non-English languages." Election Integrity Partnership, 21 Ekim 2020,
- » <https://www.eipartnership.net/policy-analysis/inconsistent-efforts-against-us-election-misinformation-in-non-english>
- » Nicas, Jack. "How YouTube Drives People to the Internet's Darkest Corners." The Wall Street Journal, 7 Şubat 2018,
- » <https://www.wsj.com/articles/how-youtube-drives-viewers-to-the-internets-darkest-corners-1518020478>
- » Ohlheiser, Abby. "They turn to Facebook and YouTube to find a cure for cancer — and get sucked into a world of bogus medicine." The Washington Post, 25 Haziran 2019, <https://www.washingtonpost.com/lifestyle/style/they-turn-to-facebook-and-youtube-to-find-a-cure-for-cancer--and-get-sucked-into-a-world-of-bogus-medicine>
- » Ribeiro, Manoel Horta, Raphael Ottoni, Robert West, Virgílio A. F. Almeida ve Wagner Meira. "Auditing radicalization pathways on YouTube." In Proceedings of the 2020 Conference on Fairness, Accountability, and Transparency (FAT\* '20) (s. 131-141), 27 Ocak 2020, <https://doi.org/10.1145/3351095.3372879>
- » Roberts, Sarah T. "Behind the Screen: Content Moderation in the Shadows of Social Media." Yale University Press, 2019, <https://doi.org/10.2307/j.ctvhrcz0v>
- » Roth, Camille, Antoine Mazières ve Telmo Menezes. "Tubes and bubbles topological confinement of YouTube recommendations." PLoS ONE 15(4), 21 Nisan 2020, <https://doi.org/10.1371/journal.pone.0231703>

- » Sanna, Leonardo, Salvatore Romano, Giulia Corona ve Claudio Agosti. "YTTREX: Crowdsourced Analysis of YouTube's Recommender System During COVID-19 Pandemic." *Information Management and Big Data* (s.107-121), Mayıs 2021, [https://doi.org/10.1007/978-3-030-76228-5\\_8](https://doi.org/10.1007/978-3-030-76228-5_8)
- » Singh, Spandana. "Why Am I Seeing This? How Video and E-Commerce Platforms Use Recommendation Systems to Shape User Experiences." Open Technology Institute, 25 Mart 2020, <https://www.newamerica.org/oti/reports/why-am-i-seeing-this/>
- » Singh, Spandana. "Everything in Moderation: An Analysis of How Internet Platforms Are Using Artificial Intelligence to Moderate User-Generated Content." Open Technology Institute, Temmuz 2019, <https://www.newamerica.org/oti/reports/everything-moderation-analysis-how-internet-platforms-are-using-artificial-intelligence-moderate-user-generated-content/>
- » Tufekci, Zeynep. "YouTube, the Great Radicalizer." *The New York Times*, 10 Mart 2018, <https://www.nytimes.com/2018/03/10/opinion/sunday/youtube-politics-radical.html>
- » Vermeulen, Mathias. "The keys to the kingdom. Overcoming GDPR-concerns to unlock access to platform data for independent researchers". Knight First Amendment Institute Draft paper, 27 Kasım 2020, <https://doi.org/10.31219/osf.io/vnswz>
- » "Answers to Your Questions About the Dark Side of the Internet." Mozilla, 3 Eylül 2019, <https://foundation.mozilla.org/en/blog/answers-your-questions-about-dark-side-internet/>
- » "Continuing our work to improve recommendations on YouTube." YouTube Official Blog, 25 Ocak 2019, <https://blog.youtube/news-and-events/continuing-our-work-to-improve>
- » "How YouTube Works." YouTube, 2021. <https://www.youtube.com/intl/howyoutubeworks/>
- » "2020 Ranking Digital Rights Corporate Accountability Index." Ranking Digital Rights, 2020, <https://ranking-digitalrights.org/index2020/explore-indicators>
- » "The Four Rs of Responsibility, Part 2: Raising authoritative content and reducing borderline content and harmful misinformation." YouTube Official Blog, 3 Aralık 2019, <https://blog.youtube/inside-youtube/the-four-rs-of-responsibility-raise-and-reduce>
- » "What Happened After My 13-Year-Old Son Joined the Alt-Right." *Washingtonian*, 5 Mayıs 2019, <https://www.washingtonian.com/2019/05/05/what-happened-after-my-13-year-old-son-joined-the-alt-right>

# Teşekkür

Bu rapor, McCrosky ve Brandi Geurkink tarafından yazıldı.

Katkı sunan diğer yazarlar:

Kevin Zawacki, Anna Jay, Carys Afoko, Maximilian Gahntz ve Owen Bennett.

Analiz çerçevemizi geliştirmemiz için uzmanlıklarını bizimle paylaşan ekip üyelerine teşekkür ederiz: Gabrielle Guillemain, Dia Kayyali, Chico Camargo, Udbhav Tiwari, Jason Chuang, and Amber Sinha. The views in this paper do not necessarily reflect theirs, nor their employers. We also want to thank the research assistants at the University of Exeter, supervised by Dr. Chico Camargo, for their diligence and hard work analyzing the reported videos: Julia Dominika Burzyk, Maria Campbell, Giulia Catelani, Hoi Ying Chang, Sharon Choi, Hannah Cox, Isabel Dally, Ben Entwisle, Alice Gallagher Boyden, Laura Garratt, Adriano Giunta, Lisa Greggi, Rosemary Griggs, Matthew Gurney, Connie Hitchin, Olliver Hopkins, Oana Ionescu, Elliot Jones, Ritvika Kedia, Ruslan Kudryashov, Smriti Lakhotia, Michael Lewis, William Lewis, Mitran Malarvannan, Lois Mander, Zachary Marre, Alina McGregor, Inês Mendes de Souza, Ayodele Ogunyemi, Henry Payne, Sadaf Sahel, Matej Svoboda, Jia Tang Zhi, Martina Toneva, Olivia Warnes, Katherine Williams, Ching Yin Yan, Suchitra Bansode, Mingting Hong, Ayush Shrivastav, Feng Xu.

Nathalie Maréchal, Spandana Singh ve Mathias Vermeulen de dahil, bu rapora görüşleri, yönlendirmeleri ve yorumlarıyla katkı sunan sivil toplumdan meslektaşlarımıza minnettarız. Bu rapordaki görüşler onların veya çalışanlarının görüşlerini her zaman yansıtmıyor olabilir. RegretsReporter'ı geliştiren Fred Wollsen'a ve bu çalışmayı finanse eden Reset Tech ekibine teşekkür ederiz.

Son olarak, RegretsReporter aracılığıyla bize veri sağlayan 37 bin 380 gönüllümüze teşekkür ederiz. Yalın veri politikamız gereğince bu insanların kim olduğunu bilmiyoruz (dolayısıyla kendilerine isimleriyle teşekkür edemiyoruz), fakat onların katkısı olmasaydı bu araştırma mümkün bile olmazdı.

Ekler: [Kategorilere göre YouTube Pişmanlık örnekleri \[İngilizce\]](#)

teyit 

 DEPİM