# Arama Motorlarının Geleceği

Dr. Tuba Sarıgül [ *TÜBİTAK Bilim ve Teknik Dergisi*

**İlk web sitesinin 6 Ağustos 1991'de genel kullanıma sunulmasından sonra web sitelerinin sayısı her geçen gün katlanarak arttı ve aradan geçen 30 yılı aşkın sürede 1,2 milyara ulaştı. İnternetteki bilgi miktarının artmasıyla arama motorları, kullanıcıların istedikleri bilgiye ulaşabilmesi için kritik öneme sahip araçlar hâline geldi.**

**Klasik arama motorlarında kullanıcıların girdiği anahtar kelimelere göre en uygun sonuçlar listeleniyor, kullanıcıların ise arama sonuçları arasında aradıkları bilgiyle en fazla ilişkili olan bağlantıyı seçmesi gerekiyor. Ancak günümüzde kullanıcılar daha karmaşık sorgular yapıyor ve arama amaçlarının bağlamına daha uygun cevaplar bekliyor. Geleneksel arama motorları bu ihtiyaçları karşılamak için yetersiz kalırken doğal dil işleme ve büyük dil modelleri gibi yapay zekâya dayalı sistemler, yapılan sorguların bağlamına uygun anlamlı cevaplar verilmesine olanak tanıyor.**

**Peki, bu köklü değişim bilgiye ulaşma alışkanlıklarımızı nasıl etkileyecek?**

da-kuk / Getty

# Arama Motorlarının Gelişimi

İlk arama motoru olan Archie, 1990 yılında McGill Üniversitesinde lisansüstü eğitimine devam eden ve aynı zamanda üniversitenin bilgi teknolojileri bölümünde sistem yöneticisi olan Alan Emtage tarafından geliştirildi. Emtage'ın görevi, öğrencilerin ve öğretim üyelerinin ihtiyaç duyabileceği yazılımları ve dosyaları bulmak için internet üzerindeki dosya transfer protokolü (FTP) sunucularını taramaktı. Manuel olarak yapılan bu görev son derece zaman alıcı olduğundan Emtage, FTP sunucularındaki dosya adlarını otomatik olarak tarayan bir kod geliştirdi. Archie olarak isimlendirilen sistem sadece dosya isimlerinde arama yapıyor, dosya içeriğini analiz etmiyordu. Ancak Archie'nin ardındaki fikir, 1990'lı yılların sonu ve 2000'li yılların başında internetin yaygınlaşmasıyla Google dâhil olmak üzere bugün kullandığımız arama motorlarının geliştirilmesine öncülük etti.

Massachusetts Teknoloji Enstitüsünden (MIT) Matthew Gray'in 1993 yılında web sitelerinin sayısını ve büyüme hızını takip etmek için geliştirdiği World Wide Web Wanderer adlı sistem ise web sitelerinin URL adreslerini tarıyordu. Sistem daha sonra Wandex isimli veri tabanında saklanan bu adreslerde arama yapmaya imkân verecek şekilde geliştirildi. Böylece Wandex, ilk web arama motoru olarak tanımlanan sistem oldu.

İnternetin giderek yaygınlaşması, kullanıcıların internette aradıkları bilgilere ulaşmasını zorlaştırıyordu. Stanford Üniversitesinde doktora öğrencileri olan David Filo ve Jerry Yang, 1994 yılında ilgilerini çeken web sitelerini kategorilere ayıran bir sistem üzerinde çalışırken internette bir gezinme aracına olan ihtiyacı fark etti. Bunun üzerine projelerini geliştirerek web sitelerini bilim, spor, haber gibi farklı kategorilere göre gruplandıran bir dizin olan Yahoo'yu kurdular. Yahoo'nun sisteminde web siteleri manuel olarak kategorilere ayrılsa da bu dizinleme sistemi, web sitelerindeki içeriklerin otomatik olarak taranıp analiz edildiği modern arama motorlarında kullanılan indeksleme sisteminin geliştirilmesine ilham kaynağı oldu.

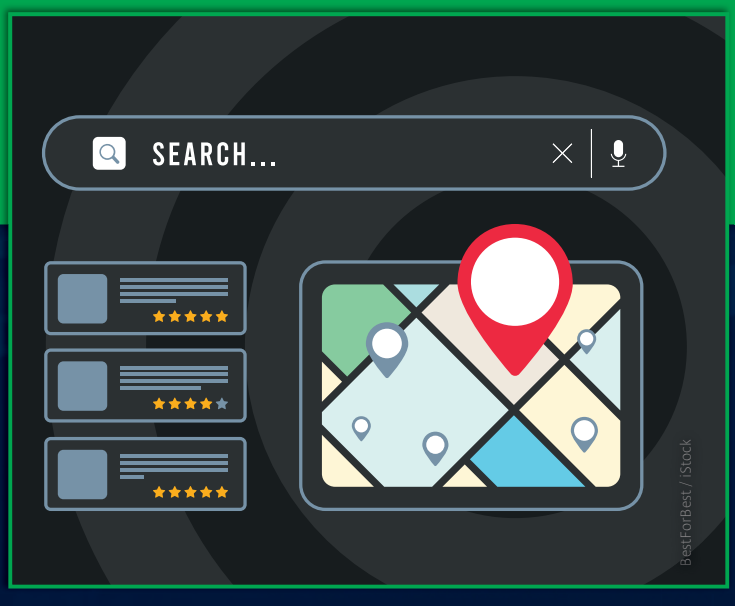
Standford Üniversitesinde doktora araştırmalarına devam eden Larry Page ve Sergey Brin tarafından 1998 yılında kurulan Google ise arama motorlarının gelişim sürecinde devrim niteliğinde bir değişime öncülük etti. Larry Page ve Sergey Brin, o dönemde var olan sistemlerden daha etkili arama sonuçları üretmek için web sayfaları arasındaki bağlantıları analiz eden, PageRank adını verdikleri bir algoritma geliştirdi. Bu sistemde, web sayfalarının birbirine verdiği bağlantılar analiz edilerek sayfanın güvenilirliği ve önem derecesi belirleniyordu. Bir web sayfasına başka web sayfalarından ne kadar çok bağlantı veriliyorsa sayfa o kadar önemli ve güvenilir kabul ediliyordu. Bu sayede arama sonuçları listelenirken sadece aranan

anahtar kelimedeki eşleşme değil, aynı zamanda web sayfasının başka sayfalardan aldığı bağlantılar da değerlendiriliyordu. Bu sayede Google, kullanıcılara beklentileriyle son derece ilişkili sonuçlar gösterebildi. Bu özelliği ve kullanıcı dostu arayüzü sayesinde Google'ın popülaritesi hızla arttı ve Google, arama motoru teknolojisini yönlendiren bir platform ve bu alanda ölçüt kabul edilen bir standart hâline geldi.

Sonraki yıllarda sadece anahtar kelimelerdeki eşleşmeye odaklanmayan, aynı zamanda kelimelerin arasındaki ilişkiden hangi bağlamda kullanıldıklarını anlamaya çalışan arama algoritmaları kullanıma sunuldu. Örneğin "Apple" ve "fiyat" kelimeleriyle yapılan bir sorguda anahtar kelime odaklı algoritmalar elma fiyatlarını listelerken, bağlam temelli aramanın sonuçlarında -kullanıcı daha önce bilgisayar, tablet gibi teknoloji ürünlerine yönelik arama yaptıysa- Apple markalı elektronik cihazların fiyatları gösterilebiliyor.

Ayrıca sonuçlar listelenirken metin, görsel, video, harita, haber gibi farklı içerik türlerinin listelendiği çok modlu arama sistemleri geliştirildi. Kullanıcıların arama geçmişleri ve davranışları analiz edilerek kişiselleştirilmiş sonuçlar sunulmaya başlandı. Bu sayede artık "en yakın ATM" şeklinde yapılan bir aramada, kullanıcının konumu algılanarak en yakınında bulunan bankalar ve ATM cihazları sonuçlarda listelenebiliyor. Ayrıca son yıllarda başlık, kısa özet ve URL adresi şeklinde sıralanan sonuçların yanı sıra kullanıcıların aradıkları bilgiyi doğrudan gösteren bilgi kartları, öne çıkan kısa bölüm gibi özellikler kullanıma sunuldu. Örneğin kullanıcı "Ankara hava durumu" şeklinde bir arama yaptığında, hava durumu verilerinin olduğu bir web sayfasına tıklamasına gerek kalmadan güncel hava durumu verileri arama sonucunda doğrudan görüntülenebiliyor.

Arama motorlarında kaydedilen bu büyük ilerlemelere rağmen yapay zekâ teknolojilerinde son yıllarda yaşanan gelişmeler, internetteki bilgilere ulaşma alışkanlıklarımızda köklü değişimlere yol açacak bir dönüm noktasında olduğumuzu gösteriyor.




Son yıllarda arama sorgularında başlık, kısa özet ve URL adresi şeklinde sıralanan sonuçların yanı sıra bilgi kartları ve öne çıkan kısa bölüm gibi alanlarda kullanıcıların aradıkları bilgiyi doğrudan gösteriliyor.

# Arama Motorları, Yapay Zekâ Desteği ile Arama Robotlarına Dönüşebilir

tondone / iStock

Yapay zekâ teknolojilerinin internetteki bilgiye ulaşma konusunda sunacağı belki de en büyük yenilik, sohbet şeklinde bir etkileşime imkân vermesi. Geleneksel arama motorlarında kullanıcı anahtar kelimeler aracılığıyla bir sorgu yapar ve arama motoru bu sorguyla en ilişkili sonuçları listeler. Yani bu sistemlerde tek yönlü bir etkileşim vardır. Yapay zekâya dayalı arama teknolojilerindeyse kullanıcılar anahtar kelimeyle sorgu yapmak yerine günlük konuşmadakine benzer sorular sorabilir. Bu sistemler, konuşmanın bağlamını değerlendirerek kullanıcıların sorularının daha iyi anlaşılmasını ve bu sayede daha doğru cevaplar üretilmesini sağlar. Kullanıcılar, verilen yanıtları iyileştirmek için ek sorular sorabilir ya da yeni bilgiler vererek sorularını netleştirebilir. Yani yapay zekâya dayalı sistemlerde sohbet tarzında karşılıklı bir etkileşim söz konusudur.

Yapay zekâya dayalı arama sistemleri, sorguları anlamak ve cevap oluşturmak için büyük dil modellerini kullanır. Büyük dil modelleri insan dilini anlamak, analiz etmek ve dil ile ilişkili farklı görevleri (örneğin metin üretme, çeviri, özetleme) gerçekleştirmek için büyük veri setleriyle eğitilen sistemlerdir. İnsan dilini anlamak, yorumlamak ve buna göre metin ve konuşma üretmek için doğal dil işleme (NLP) algoritmalarını kullanırlar. Günümüzde yaygın olarak kullandığımız ChatGPT, Gemini, DeepSeek gibi üretken yapay zekâ teknolojileri büyük dil modeli uygulamalarının en bilinen örnekleridir.

NurPhoto / Getty

Büyük dil modellerinin arama motoru sistemleriyle birlikte kullanılması iki şekilde gerçekleşebilir. Birincisi, büyük dil modellerinin arama motorlarının sahip olduğu büyük ve nitelikli veri setleriyle eğitilmesi. Normalde büyük dil modelleri internetteki geniş veri setleriyle eğitilir. Ancak bu veri setlerinde yanlış, eksik ya da önyargılı cevaplar oluşturulmasına yol açabilecek yanlı bilgiler bulunabilir. Arama motorlarıysa yapılan sorguyla en ilişkili sonuçları vermek için internetteki içerikleri önemine ve niteliğine göre sıralar. Bu veriler ile eğitilen büyük dil modelleri, kullanıcıların internette arama yaparkenki temel amaçlarını daha iyi anlayabilir ve yapılan sorgunun bağlamına uygun, doğru, güvenilir ve tutarlı cevaplar üretebilir. Arama motorlarının büyük dil modellerine en önemli katkılarından biri de internetten gerçek zamanlı bir şekilde veri alınmasını

sağlayarak en güncel cevapların üretilmesini mümkün hâle getirmesidir.

İkinci yaklaşım ise arama motorlarının özelliklerinin büyük dil modelleriyle iyileştirilmesi. Büyük dil modelleri, internetteki bilgiyi etkili bir şekilde özetleyerek arama motorlarının bu bilgileri daha başarılı bir şekilde indekslemesine imkân verir. Böylece web sayfalarındaki arama sorgularıyla ilgili en uygun içerikler bulunur.

Günümüzde Google, OpenAI, Microsoft gibi teknoloji şirketleri büyük dil modelleri ile arama motorlarının birlikte çalıştığı hizmetleri kullanıma sundu. Örneğin Microsoft'un arama motoru Bing, yapay zekâ desteğiyle zenginleştirilmiş sonuçlar veriyor. En bilinen büyük dil modeli uygulaması olan ChatGPT'nin bazı sürümlerinde web sayfalarına gerçek zamanlı erişim bulunuyor. Google, arama işlevi için özelleştirdiği Gemini isimli büyük dil modelini kullanarak arama sonuçlarında yapay zekâ destekli hızlı cevaplar gösterebiliyor. Bir startup şirketi tarafından geliştirilen Perplexity ise kullanıcıların sorgularına büyük dil modellerini kullanarak doğrudan cevap üreten bir uygulama yani bir cevap motoru.

Birbirlerinin özelliklerini geliştirseler de arama motoru teknolojilerinin ve büyük dil modellerinin bir arada kullanılması durumunda karşı karşıya kalınan bazı zorluklar ve sorunlar bulunuyor.

## Büyük Dil Modellerinin Ürettiği İçeriklerin Sahibi Kim?

Klasik arama motorları, yapılan sorguyla ilişkili cevapların bulunduğu web sayfalarını, ilgi düzeyi en yüksek olanı en üst sırada görünecek şekilde arama sonuçlarında listeler. Kullanıcılar bu web sayfalarından seçtiklerine tıklayarak orijinal kaynağa ulaşır. Büyük dil modeli uygulamaları ise web sayfalarındaki içerikleri kullanarak arama sorgusuna yönelik doğrudan cevaplar verir. Cevabı oluştururken kullanılan web sayfalarının listesi sonuçlarla birlikte sunulsa da çoğunlukla kullanıcılar orijinal kaynağa tıklamaz. Web sitesi trafiğinde azalmaya neden olan bu durum web içerik üreticileri için son derece önemli.

Özellikle özgün ve nitelikli içeriklerin büyük dil modellerinin eğitilmesinde kullanılması telif hakkı açısından sorun olabilir. Çünkü yararlanılan içerikler kaynak olarak gösterilse de özgün bir içeriğin tamamının izinsiz kullanılması telif hakkı ihlali

riskine yol açabilir. Örneğin yakın zamanda bir haber sitesinin abonelere özel yayınladığı içeriklerin büyük bir dil modeli uygulamasında neredeyse aynı şekilde kullanılması nedeniyle haber sitesi, telif hakkı ihlaline yönelik yasal bir süreç başlattı.

Ayrıca web sitesi trafiğine dayanan dijital ekonomik pazar bu dönüşümden belirgin şekilde etkilenebilir. Örneğin web reklamları genellikle kazandırdığı tıklama sayısına göre ücretlendirilir. Günümüzde dijital ekonomide önemli bir endüstriye dönüşmüş olan SEO uygulamalarının da yapay zekâya dayalı arama

sistemlerinin web trafiğinde neden olduğu değişime uyum sağlaması gerekiyor.

Geleneksel arama motorlarında web sitelerinin arama sonuçlarında üst sıralarda listelenmesi, tıklanma oranlarını önemli derecede etkiliyor. Bu nedenle web içeriklerini, arama motoru algoritmalarının önceliklendireceği şekilde hazırlamak son derece önemli. Arama motoru optimizasyonu yani SEO olarak isimlendirilen bu uygulamalar, dijital ekonomide büyük bir endüstriye dönüşmüş durumda. Dolayısıyla yapay zekâya dayalı arama sistemlerinde, arama sonuçlarında üst sıralarda listelenmenin öneminin azalması SEO uygulamalarının da bu değişime uyum sağlamasını zorunlu hâle getiriyor.

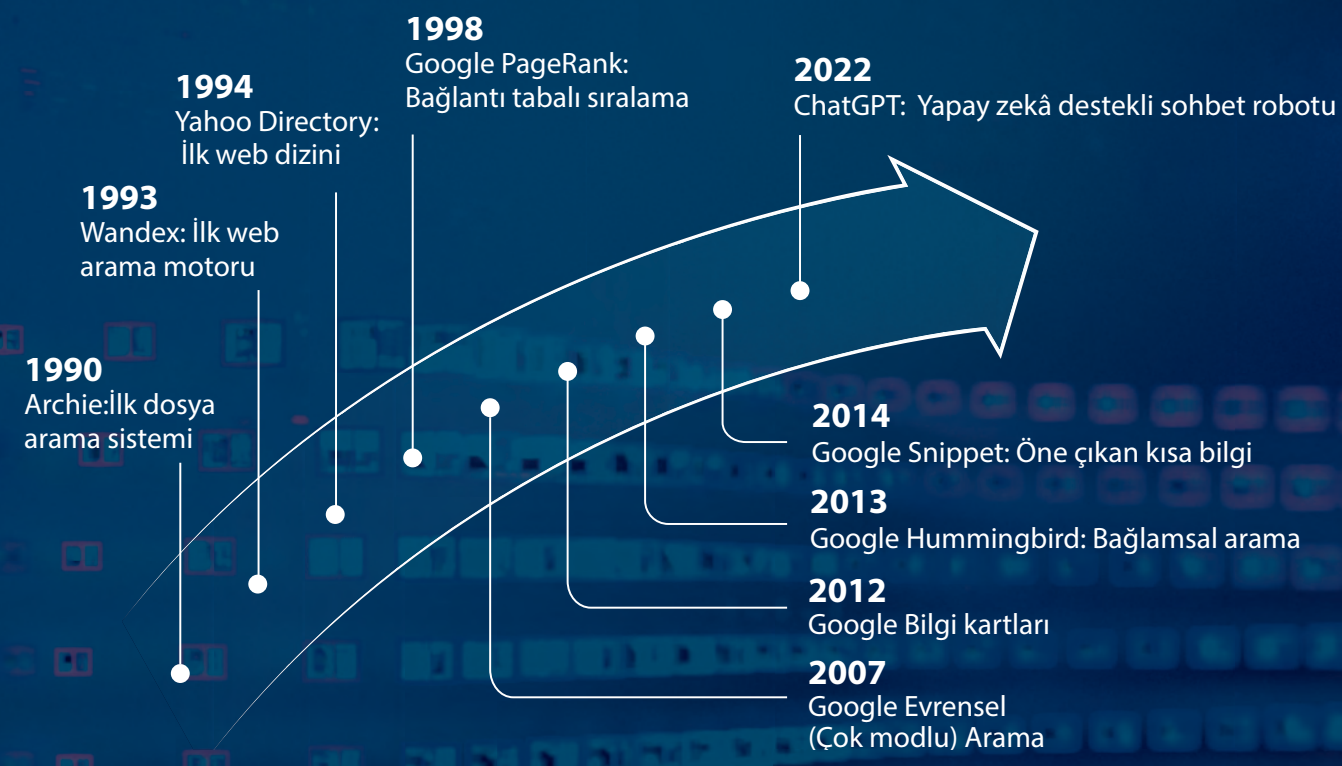
Ayrıca yapay zekâ destekli arama işlevinde kişiselleştirilmiş sonuçlar sunulabilmesi için kullanıcılara ait çok miktarda veri toplanıyor, saklanıyor ve analiz ediliyor. Bu durum ise kişisel verilerin gizliliği konusunda endişelere neden oluyor. Kişisel verilerin kötüye kullanılması ya da üçüncü taraflarla paylaşılması riski nedeniyle verilerin kullanımına yönelik politikaların şeffaflaştırılması gerekiyor.

Üretken yapay zekâ teknolojileri, internet üzerinden bir konuyla ilgili bilgi almak istediğimizde uzmanıyla sohbet eder gibi iletişim kurmamıza imkân veriyor. İnternetteki arama alışkanlıklarımızı kökünden değiştiren bu yeni dönemin tam anlamıyla başarıya ulaşması ise veri kaynağının şeffaflaştırılması, telif hakkı ve kişisel verilerin kullanımına yönelik ihlallerin önüne geçilmesi, yanlı ve hatalı sonuçların ayıklanması gibi endişelerin giderilmesine bağlı. ■

Mongkol Chuewong / Getty

Arama sonuçlarının kişiselleştirebilmesi için konum, arama geçmişi, cihaz bilgileri gibi çok miktarda kişisel veri kullanılabilir.

# Arama Motorlarının Gelişimindeki Dönüm Noktaları




**Kaynaklar**


Brin, S., Page, L., "The Anatomy of a Large-Scale Hypertextual Web Search Engine", *Computer Networks and ISDN Systems*, Cilt 30, s. 107- 117, 1998.

https://home.cern/science/computing/birth-web/short-history-web

https://www.captechu.edu/blog/alan-emtage-creator-of-archie-worlds-first-search-engine

https://ahrefs.com/blog/google-pagerank/

https://blog.google/products/search/about-knowledge-graph-and-knowledge-panels/

https://blog.google/products/search/generative-ai-google-search-may-2024/

https://intellectual-property-helpdesk.ec.europa.eu/news-events/news/intellectual-property-chatgpt-2023-02-20_en