

# BÜYÜK VERİ

Büyük veri, kısaca farklı kaynaklardan gelen, hacmi büyük, dağınık, görüntüde ilişkili gözükmeyen veri yığınına ifade ediyor.

Bu kavram, son dönemde dijital dünyayla ilgilenen herkesin en önemli gündem maddelerinden biri. Büyük Veri'nin ne kadar kıymetli olduğu üzerine yapılan değerlendirmeler bir yana, hangi verinin Büyük Veri olduğu üzerine ciddi bir tartışma başlamış durumda. Şu ana kadar oluşan konsensüs sosyal medya, web logları gibi verilerin sabit disklerde işgal ettikleri yerler nedeniyle Büyük Veri olduğu yönünde. Büyük Veri meselesini sadece internet düzleminde görmeyenler bu listeye, hızlıca hava durumu tahmini, araç trafiği veya CERN'deki Hadron Çarpıştırıcısında toplanan verileri de ekliyorlar. Tıp alanında çalışan bilim insanlarının genom araştırmaları da yine Büyük Veri düzlemine dahil edilebilen konular. Liste bu şekilde uzuyor gidiyor.

Sorun şu ki bu veriler eskiden de vardı. En az yüz yıldır hava tahmini amaçlı olarak toplanan veriler, gen haritasının belirlenmesi ile yapılan çalışmalar, aynı web logları gibi 15 yıldan fazla süredir hayatımızın içinde. Peki ne değişti?

Değişen sadece şu: veriye bakış şeklimiz. Eskiden bir sorun olduğunda dönüp bakılan veriler, artık değer yaratacak birer kaynak olarak görülüyor. Yeter ki toprağın altında yatan bir maden gibi çıkarmasını bilelim. Büyük Veri bir iş yapış felsefesi. Sadece verileri kullanarak dahi işimizi, daha verimli ve karlı yapabiliriz.

Evet, internet; hayatımızın önemli bir kısmının dijitalleşmesine, dolayısıyla da veri olarak kolayca toplanabilmesine ve analiz edilebilmesine olanak sağladı. Evet, disklerimiz ve disklerdeki verileri işleyebilecek bilgisayarlarımız çok daha güçlü. Evet, artık hazır teknolojiler, veri madenciliği ve istatistik gibi yöntemlerin daha kolay uygulanabilmesini mümkün kılıyor. Peki bütün bunlara sahip olmamıza rağmen neden hala Büyük Veriyi kullanırken zorlanıyoruz?

Çünkü az önce de bahsettiğimiz gibi, Büyük Veri bir yaklaşım şeklidir. Siz sabit disklerinizde duran o verilerin bir fayda üreteceğine inanmıyorsanız veya inansanız bile o verileri hangi senaryo ile fayda üreteceğinizi bilmiyorsanız Büyük Veriden yararlanamazsınız.

Peki Büyük Veri'yi Ne Amaçla Kullanacağız?

Toplayacağımız veriye karar verirken, kullanım amacımızı da en başta netleştirmemiz gerekiyor. Henüz hiçbir teknoloji bu tür stratejik bir kararı sizin yerinize sağlıklı şekilde veremez.

Örneğin bir hastaneyi ele alalım. "Asansörün en çok hangi katlarda durduğu" verisi ne işimize yarayabilir? Eğer ekibimizde bir analist var ise şunu söyleyebilir: "ben asansörlerin takvim günleri ile ilişkilerini analiz etmek istiyorum, böylece hangi servisin daha yoğun olduğunu bulacağım" ancak bu bilgiye kolayca hastanenin bilgi sisteminden ulaşılabilir. Diğer yandan böyle bir bilgi katlara servis yapmak ile görevli kantincinin ilgilenebileceği bir bilgidir, yeter ki böyle bir bilgiden haberi olsun.

Bu örnekteki kantinci gibi, Büyük Veriden en büyük faydayı elde etmeye katkı sağlayacak tarafların da sürece dahil edilmesi büyük önem taşır. Önemli olan; toplayabildiğimiz kadar veri toplayalım

yaklaşımının ötesinde, daha çalışmanın başında veriyi nasıl faydaya dönüştüreceğimizin mümkün mertebe netleştirilmesidir.

## **Büyük Veri'ye Nereden Başlayalım?**

Bu konuya ilk önce “vaka çalışmaları” ile başlamanızı öneririm. Sadece kendi sektörünüzdeki değil tamamen farklı alanlardaki Büyük Veri hikayelerini okuyun çünkü büyük veri, mühendislik gibi tamamen analitik düzlemde ilerleyen bir disiplin değildir. Ne kadar farklı çalışma okursanız vizyonunuz o derece genişler ve yenilikçi bir Büyük Veri projesi geliştirme ihtimaliniz artar. Çalışmanın daha başında hangi veriyi ne için ve nasıl kullanacağınızı belirlemenin, başarıya giden yolda kritik bir faktör olduğunu yeniden vurgulamak istiyorum.

## **Büyük Veri Hangi Bileşenlerden Oluşur?**

Büyük Veri 5 ana kavram ile ifade ediliyor. 5V olarak da bilinen bu kavramlar:

- Variety yani Çeşitlilik
- Velocity yani Hız
- Volume yani Veri Büyüklüğü
- Veracity yani gerçeğe uygunluk
- ve son olarak da Value yani Değeri

Şimdi bu 5V'yi kısaca açıklayalım:

- **Çeşitlilik anlamındaki Variety;** üretilen verilerin genelde yapısal olmadığını ve farklı ortamlardan gelen veri formatında olduğunu ifade eder. Sağlıklı bir çalışma için, bu çeşitli formattaki verilerin, birbirine dönüştürülebilir olmaları gerekir.
- **Hız anlamına gelen Velocity,** büyük veri üretiminin her geçen gün artan hızını ifade eder. Hızla büyüyen veri ile beraber, veriyi işlemek için gerekli işlem sayısı ve çeşitliliği de aynı oranda artar. Bu yükü kaldırabilmek için hem yazılım hem de donanım altyapımızın yeterli olmasına dikkat etmeliyiz.
- **Veri Büyüklüğü yani Volume,** hızla birikmeye başlayan veri yığınları ile nasıl başa çıkacağımızın önemini vurgulamaktadır.
- **Verinin gerçeğe uygunluk derecesini belirten Veracity;** veriler içerisinde anlamsız kayıtların olabileceğini ve bu kayıtların sağlıklı sonuç alabilmek adına temizlenmesi gerektiğini ifade etmektedir. Mesela akan trafikte araba hızlarını değerlendirirken, bazı araçların ortalama hızlarının eksi değerinde geldiğini görürsek, o araçların hız sensörlerinin arızalı olduğunu düşünerek bu kayıtları devre dışı bırakabiliriz.
- **Değer yani Value;** büyük verinin üretimi ve işlenmesi ile elde edilen verilerin şirketimiz için artı değer sağlıyor olmasının gerekliliğini vurgulamaktadır.

Büyük Verinin kullanımı ile ilgili bazı somut örnekler verelim:

- Kişilerin sosyal medya paylaşımları, kişilerin tüketim alışkanlıkları, dünya görüşleri, yaşam tarzları hakkında önemli ipucu verir.
- Standart bir tüketicinin internette yaptığı tüm haberleşme, arama, satın alma ve paylaşım türü işlemleri, hizmetleri aldığı platformlar tarafından saklanır.
- Sağlık Kuruluşları; hastalarına kişiye özel sağlık hizmeti sunabilmek için, tüm bireysel veriler sayısal ortamlarda depolanır ve değerlendirilir.
- Bankalar, müşteri verisini kullanarak, internet bankacılığında kişiye özel deneyim yaşatabilirler.
- Arama motorlarında gelen sonuçlar, kişisel verilerin analizi sonucu en uygun şekilde gösterilir.

### **Şimdi gelin, bir KOBİ olarak büyük veri ile nasıl çalışabileceğimizi adım adım değerlendirelim.**

1. **Hedef belirlemek:** Yeni müşteriler çekmek, mevcut müşterileri elde tutmak, sosyal medya üzerinden müşterilerle etkileşimi artırmak gibi hedeflere ulaşmak için, toplamamız gereken veri türlerini netleştirmemiz gerekiyor.
2. **Veri haritası çıkarmak:** Hangi veriyi nereden toplayacağımızı bu adımda netleştiriyoruz. Bu kısmın iyi planlanması çok önemlidir.
3. **Tüm veriler için ortak bir platform oluşturma.** Temel veri kaynaklarınıza ek, sosyal medya ve Nesnelerin İnterneti gibi yeni teknolojilerle elde edilen bilgiler bu platformda toplanmalı. Burada toplanan tüm verilerin analizi, şirketlere sağlıklı bir 'pusula' gibi yol gösterebilir.
4. **Doğru işgücüne yatırım:** En iyi teknolojiler ve en gelişmiş analitik araçlar, onları nasıl kullanacağını bilemeyenler için anlam taşımaz. Eğer Büyük Veriden yararlanmayı düşünüyorsanız, gerekli becerilere sahip yeni çalışanlarla ekibinizi güçlendirmeniz gerekir.
5. **Analitik bir bakış açısı:** Analiz sonuçlarını, organizasyonun her noktasında kullanmak, başarıyı da beraberinde getirecektir.
6. **Kurum kültürü:** Müşteri odaklılığını hedefleyen şirketler, bunu kurum kültürü haline getirmelidir. Bunu yapabilmenin en iyi yolu, kurum yatırımlarını da yönlendirecek, müşteri deneyimi konusunda tecrübeli yöneticilerin istihdam edilmesi ve gerekli yetkilerin verilmesidir.
7. **Daha iyi hizmet için tüm bilgileri kullanın:** Ürünleri ve hizmetleri geliştirirken tüm bilgilerden faydalanmak gerekiyor. Bu bilgileri zenginleştirmek adına da web, mobil ve sosyal medya teknolojilerine eğilmek gerekiyor. Şirketler sundukları ürünleri ve hizmetleri özelleştirmek adına müşterilerin hakkındaki görüşlerinden faydalanabilir.
8. **Hedefe bağlı kalma:** İş araçları ve iş süreçleri farklı olabilir. Ancak hedef hep aynı: Veriyi iyi kullanarak müşteriye daha iyi anlamak ve memnuniyet sağlamak. Şirketler, hedefleri doğrultusunda çok basit bir analitik raporlama da karmaşık bilişsel teknolojiler de kullanabilir. Tek bir çözümün olmadığı bu yolculukta şirketlerin kendi müşterilerinin memnuniyeti için en doğru rotayı çizmesi gerekiyor.
9. **Sonuçların ölçümü:** Veri analitiği kullanan şirketler ileriye dönük olarak müşterisini daha iyi anlayıp anlamadığını çözmek isteyecektir. Müşteri sayısı arttı mı, müşteri başına gelir artıyor mu, müşteri sadakati geliyor mu, sosyal medyada şirket daha olumlu etkileşim alıyor mu ve tabii ki tüm bu faktörler gelirleri ne yönde etkiliyor gibi sorular yanıtız kalmamalıdır.

Sonu olarak;

Byk verilerin depolanması ve iřlenmesinde kritik unsurlardan birisi kiřisel verilerin korunmasıdır. Devletimizin de konuya gsterdiđi hassasiyet neticesinde, veriyi toplayan ve iřleyen kurumlara ciddi sorumluluk yklenmiřtir.

Byk veri alıřmaları ile aslında hedeflenen, kullanabildiđimiz tm verilerden faydalanarak mřteri memnuniyetini arttırmak ve bunun neticesinde de řirketimize deđer ve para kazandırmaktır.

Byk Veri ile hayal gcnz eř deđer tutabilirsiniz nk bu kadar byk veri kmeleri ile yapabileceđin her řey hayal gcnz ile sınırlıdır. Sadece ne yapmak ve neyi bilmek istediđimize iyi karar vermemiz gerekmektedir.