

Büyük Veri: Fırsatlar ve Tehditler

"Hakkında çok fazla şey bilinen bir yaşam, dayanılmaz bir yaşamdır."

(Elias Canetti)

ALİ ÖZCAN

Öz

Sanayi toplumunda siyasal, ekonomik, sosyal ve kültürel alanda kitle olan insanlar, bilgi toplumuna geçişle birlikte iletişim teknolojilerinin sunduğu hizmetlerle birer birey olarak konumlandırılmıştır. Bu bireyler teknolojiyle aracılı bir şekilde sürdürdükleri ilişkiyle gündelik yaşamlarında sürekli olarak dijital izler bırakmakta ve veriler üretmektedirler. Kaydedilen bu veriler her geçen gün büyük bir artış göstermekte, yazılımlar sayesinde işlenen bu verilerden soyut ve somut değerler üretilmektedir. Büyük veri (big data) olarak kavramsallaştırılan bu durumun karakteristik özelliği, verinin yapılandırılmış formdan yoksun olması ve sürekli bir akışın (veri boyutu) söz konusu olmasıdır. Veri miktarının yoğunluğunun önem arz etmesine rağmen asıl nokta, büyük veriyi yönetmek ve ondan faydalanmak için analiz ederek değere dönüştürmekte yatmaktadır. Bu çalışma kapsamında yeni iletişim teknolojilerinin birey ve toplum hayatında yaygınlaşmasıyla birlikte ortaya çıkan büyük veri olgusuna ilişkin betimleyici bir yöntem kullanılacaktır. Büyük veri olgusuna ilişkin bakış açıları, büyük verinin karakteristik özellikleri ve yeni medya ortamında birer kullanıcı olan bireylerin büyük veriyle ilişkisel düzeyi fırsatlar ve tehditler boyutuyla tartışılacaktır.

Anahtar Kelimeler: Bilgi ve İletişim Teknolojileri, Büyük Veri, Gözetim, Panoptikon

Derleme

Geliş Tarihi: 30.10.2020

Kabul Tarihi: 09.01.2021

ORCID ID: 0000-0002-4773-9764 DOI:10.37679/trta.818569

E-mail: aliozcan29@gmail.com

Big Data: Opportunities And Threats

"Intorelable, a life one knows too much about."

(Elias Canetti)

ALİ ÖZCAN

Abstract

Humans are positioned as 'individuals' through the transition to information society accompanied by the freshness and newness of communication technologies, while they were named solely as 'masses' in economic, social and cultural fields of industrial society. Daily lives mediated by these technologies, individuals leave digital marks behind, as well as producing data. This stored data is getting bigger gradually enabling tangible and intangible material production through the use of algorithms. Conceptualised as 'big data', this phenomenon characteristically is lack of a structured form while there is a constant data flow (data size). Management and analysis of this big data is the core of the matter here, in order to make use of it as data transformed into values, in contrast with the importance of the density of data amount or size. This study through employing a descriptive method aims to give insight to big data analysis and big data phenomenon, which has made its way to daily lives of individuals and society. Various approaches toward big data, together with its characteristics, and the relation between big data and individuals as users in the new media environment will be discussed from both positive and negative aspects.

Keywords: Information and Communication Technologies, Big Data, Surveillance, Panopticon

Composition

Recieved: 30.10.2020

Accepted: 09.01.2021

ORCID ID: 0000-0002-4773-9764 DOI:10.37679/trta.818569

E-mail: aliozcan29@gmail.com

1. Giriş

Sanayi toplumunda siyasal, ekonomik, sosyal ve kültürel alanda kitle olan insanlar, bilgi toplumuna geçişle birlikte iletişim teknolojilerinin sunduğu hizmetlerle birer birey olarak konumlandırılmıştır. Kitleleşme-öz iletişim (Castells, 2013: 20-21) olarak ifade bulan yeni iletişim ortamında bireyler teknolojiyle aracılı bir şekilde sürdürdükleri gündelik yaşamlarında bilinçli ya da bilinçsiz bir şekilde dijital izler bırakmakta ve sürekli olarak veri üretmektedir. Gündelik hayatın her bir anının bilgi ve iletişim teknolojileri aracılı olarak gerçekleşmesi -sağlık, alışveriş, kamu hizmetleri, sosyal hayatın internet dolayımıyla yaşanması- tüketiciler olarak vatandaşların davranışlarının birer veri olarak kaydedilmesini beraberinde getirmektedir. Kaydedilen bu veriler her geçen gün nicel ve nitel olarak büyük bir artış göstermekte, yazılımlar sayesinde işlenen bu verilerden soyut ve somut değerler üretilmektedir. Bu verilerin üreticileri, toplayıcıları ve kullananları olmak üzere üç paydaşı bulunmaktadır. Ayrıca veriler; yapılandırılmış, yarı yapılandırılmış ve yapılandırılmamış olarak bulunmaktadır. Diğer yandan bu veriler insan etkileşimleri ve makine etkileşimleri olmak üzere iki şekilde üretilmektedir. Büyük veri (big data) olarak kavramsallaştırılan bu durumun karakteristik özelliği verinin üretiminde sürekli bir akışın (veri boyutu) söz konusu olmasıdır. Veri miktarının yoğunluğunun önem arz etmesine rağmen asıl nokta büyük veriyi yönetmek ve ondan faydalanmak için analiz ederek değere dönüştürmekte yatmaktadır. Çünkü büyük veriyle birlikte geçmişe dönük analizler yapılmasının yanında ileriye dönük tahminlerde de bulunmaktadır (Öcal, 2019: 102). Bilgisayar bilimcileri, fizikçiler, iktisatçılar, matematikçiler, siyaset bilimciler, biyo-enformatikçiler, sosyologlar ve diğer akademisyenler, insanlar, nesnelere ve onların etkileşimleri tarafından ve bunlar hakkında üretilen büyük miktardaki bilgiye erişim konusunda çalışmalar yürütmektedir. Genetik dizileri, sosyal medya etkileşimleri, sağlık kayıtları, telefon kayıtları, devlet kayıtları ve insanların bıraktığı diğer dijital izleri analiz etmenin potansiyel faydaları ve maliyetleri hakkında çeşitli çalışmalar yapılmaktadır. Büyük verinin, "iyi araç ve hizmetler oluşturmamıza mı yoksa gizlilik ihlali ve istilacı pazarlama dalgasına mı; çevrim içi topluluk ve siyasi hareketleri anlamamızı mı yoksa protestocuları izlemek ve bastırmak için yararlanılacağı" (Boyd ve Crawford, 2012: 662) önemli sorular arasında yer almaktadır. Bu yüzden sosyo-teknik bir fenomen olarak büyük verinin yükselişinin eleştirel bir sorgulamaya tabi tutulması gerekmektedir.

2. Bilgi-Güç ve İktidar İlişkisi

Tarih boyunca gücü elinde bulunduranlar (efendiler, soylular, dini yöneticiler, imparatorlar) toplum üzerindeki hâkimiyetini sağlamak adına her zaman iletişim araçlarını ve ulaşım imkânlarını kullanmıştır. Doğasında insanlar arası ilişkinin tesis edil-

mesi ve devamlılığının sağlanması adına gerekli olan iletişim olgusu, gücün tahkimi için kullanılmış ve iletişim teknolojilerindeki her yeni gelişme de bu gücün pekiştirilmesinde iletişim araçlarının kullanım düzeyini artırmıştır. Bu noktadan hareketle iletişim teknolojilerindeki gelişmelerin motivasyon değerinde en önemli ilkenin bilgi ve haberleşme üzerindeki gücün kontrolü/denetimi olduğunu söylemek yerinde olacaktır. İletişim sürecinin bir güç tahkim etme biçimi olarak tasarlanması İlk Çağ uygarlıklarından bugüne kadar gelişerek çeşitlenmiş ve zenginleşmiştir. Aristoteles, Retorik (1995: 43) kitabında, iletişim sürecini bir tasarım olarak ilk kez konuşmacı, konu ve seslenen kişi olarak üç ögeye ayırmış; karar verme durumunda bir yargıç olarak seslenen kişi/dinleyicinin konuşmanın amaç ve hedefini belirlediğini ifade etmiştir. Aristoteles'in günümüzde de hâlâ en etkili iletişim biçimi olarak sözel iletişimi bir tasarım olarak görmesi nihayetinde konuşmacının dinleyiciler üzerinde sözünün etkisini artırmayı, onlar üzerinde gücünü meşrulaştırmayı getirmektedir. İletişim sürecini eşitler arası anlam paylaşımı olarak gören genel iletişim anlayışına karşı İrfan Erdoğan da (1997: 20-21) tam tersi eleştirel bir bakış açısı benimsemiştir. Erdoğan, ortam ve koşullar altında oluşup devam eden iletişimin bir güç uygulaması olduğu gerçeğinden hareketle, üretim ilişkilerinin biçimlenmiş pratiği olarak iletişimin egemenliğin mücadele alanı olduğu gerçeğine işaret etmektedir. Hitabet bu yönüyle bir söz söyleme sanatı olarak iletişimde egemenliğin inşa edilmesi ya da gücün mücadelesinin sunulduğu bir alan olmuştur. Soylular, efendiler, krallar ve dini otoriteler sözün büyüğünden faydalanıp toplumları kontrol altına almayı başarmıştır. Tarihsel süreçte sözle birlikte başlayan iletişim tasarımı ile gücü kontrol etme anlayışı her yeni iletişim aracı ile çeşitlenmiş ve etki alanı da genişlemiş, kitlelerin kontrolü ve denetimi ise daha da kolaylaşmıştır.

Yüzyıllar boyunca iletişimi/haberleşmeyi sağlayan en önemli araçlardan biri olarak ulaşım da bilgiyi bir güç olarak kullanmanın yöntemlerinden olmuştur. Dünya tarihinin en büyük imparatorluklarından olan Roma İmparatorluğu; oluşturduğu mükemmel yol ve haberleşme ağıyla askeri gücünü tahkim etmiş, egemenliğini genişletmiş, bu sistemi ticaret ve ulaşım amaçlı kullanmıştır. Aynı şekilde İpek ve Baharat yollarının kontrolü beraberinde haberleşmeyi ve bilginin dağıtımını kontrol etmeyi getirmiş, bu sayede İpek ve Baharat yollarının geçiş güzergâhına hâkim devletler diğer devletler üzerinde güç unsurunu kullanmıştır. Benzer bir süreç günümüzde daha keskin ve hızlı bir şekilde sürmekte, iletişim ve ulaşım ağlarına sahip yapılar bu gücü merkezileştirmekte askeri, siyasal ve ekonomik amaçlar için kullanmaktadır. Roma İmparatorluğu'ndan günümüze ulaşım ve iletişim unsurları güç ve egemenlik bağlamında ilerlemektedir. Roma İmparatorluğu mükemmel yol ve posta sistemini askerî amaçlarla oluşturmuş ancak bu yapıları

sosyal ve ekonomik amaçlar için de kullanılmıştır. Günümüzün iletişim ve küreselleşme fenomeni internet başlarda ABD tarafından askerî amaçlar için tasarlanmış fakat daha sonra ekonomik ve sosyal alanın bütününe etki edecek ölçüde genişleyip fonksiyonları çeşitlenmiştir. ABD bilgi ve iletişim teknolojilerinin en büyük üreticisidir. Aynı zamanda internet gibi bilgi ve iletişim teknolojisine dayalı mecralarda bilgi üretmekte ve bu bilgiyi küresel ölçekte yaymaktadır. Bu durumda ABD'nin ekonomik ve siyasal gücünün küresel düzlemde derinleşip yayılmasına katkı sunmaktadır (Güneş, 2015: 276). Bu gerçekten hareketle Arthur'un (2017: 44) "Google'da veri güç, tıklama da para demektir." ifadesi de günümüz iletişim çağında üretilen enformasyonun "ekonomik yapının yapıtaşına" (Öcal, 2019: 102) dönüşerek bilgi, güç ve iktidar ilişkisine göndermede bulunmaktadır.

3. İletişim Teknolojileri ve Büyük Veri

"Söz uçar, yazı kalır." ifadesi; Batılı toplumların yazılı, Doğulu toplumların ise sözlü gelenekten geldiğine işaret eden deyişlerden biridir. Yazıyla birlikte tarih başladığı için yazının keşfinden öncesi tarih öncesi olarak adlandırılmaktadır. Tarih öncesi aşırı enformasyon yüküne getirilen Taş Devri çözümü olarak yazı, kolaylıklar sağlamış olsa da nihayetinde enformasyon artışındaki hızı artırmıştır (Mul, 2008: 269). Kültürel enformasyonun artışını sağlayan yazıyla birlikte basılı metinlerde yer bulan belge ve bilgiler, iletişim teknolojilerinin sunduğu imkânlarla elektronik ortama, bir başka deyişle sayısal ortama aktarılmaya ve kaydedilmeye başlanmıştır. Birer veri özelliği taşıyan bu bilgi ve belgeler binlerce yıldır insanlar tarafından analiz edilmiştir. Analog Çağ olarak ifade edilen dönemde, miktarı çok az olsa da verilerin toplanıp analiz edilmesi maliyet ve zaman açısından sıkıntıları beraberinde getirmiştir. Bilgi toplumu olarak kavramsallaştırılan günümüzde bilgi ve iletişim teknolojilerindeki hem nitelik hem de niceliksel açıdan üstel bir gelişme söz konusudur. Birey ve toplum hayatında yaygınlaşan bu teknolojilerin kullanıcılar üzerindeki etkisi de doğal olarak artış göstermektedir. Bilgi ve iletişim teknolojileri ile insan arasında vazgeçilmez bir ilişki kurulmuştur. Bu ilişkinin sosyal, kültürel, siyasal ve ekonomik boyutları toplum bilimciler tarafından incelenmekte, insan ve iletişim teknolojisi arasındaki bağın ortaya çıkardığı yeni toplumsal olgular tartışılmaktadır. İnsanın/toplumun iletişim, eğlence, iş ve sosyal hayatı kapsayan gündelik pratiklerinin bilgi ve iletişim teknolojileri aracılı olarak dönüşmesi yeni toplumsal çıktılara neden olmaktadır. 2000'lerden günümüze başta bilgisayarlar olmak üzere cep telefonlarındaki GPS sistemleri, MOBESE kameraları, internete otomatik olarak bağlı olan pek çok cihaz ve günlük hayatta kullanılan sayısız ekipman giderek akıllanmakta, internete bağlanarak kendi aralarında ağlar oluşturmaktadır. Özellikle de internet uygulamaları ve bugünün yaygın kullanım ağına

sahip sosyal paylaşım platformları bireyin gündelik hayatının önemli bir bölümünü çevrim içi/dijital ortamda geçirmesine neden olmaktadır. Sanal bir gerçekliğin inşa edildiği çevrim içi platformlarda birey, gerçek hayatının uzantısı ya da yeni bir benlik sunumu ile varlığını inşa edebilmektedir. Kullanıcı olarak bireyin gerçekliği inşa etmesi kendi elinde olmasına rağmen, kullanılan dijital dil sayesinde kullanıcıların ürettikleri yeni dünyalar her anı ile kayıt altına alınmaktadır.

Teknolojik alanda yaşanan üstel gelişmeleri açıklamak için Moore Yasası kullanılmaktadır. Bir önceki yıla göre teknolojik gelişmelerin ikiye katlanmasının grafiksel eğrisini ifade eden Moore Yasası (Brynjolfsson ve McAfee, 2015: 60) çevrim içi ortamda üretilen veri miktarının yıllar itibariyle artışını açıklamada yetersiz kalmaktadır. Çünkü özellikle de son on yıllık süreçte dijital ortamlara bağlanan kullanıcı ve araç sayısındaki artış doğal olarak veri miktarını şaşırtıcı şekilde çoğaltmaya devam etmektedir. Belgeler, dosyalar, videolar, fotoğraflar, haritalar, müzikler, kişisel durum güncellemeleri, sosyal ağlar, haberler, bilgi talepleri ve bilgi taleplerine verilen cevaplar ile her türlü sensör, dijitalleşmeyle birlikte veri oluşturmaktadır. Sanayi toplumunu birinci makine çağı, bilgi toplumunu ise ikinci makine çağı olarak ifade eden bu bakış açısı, dijitalleşmenin yaygınlık ve hız kazanmasını belirleyici gösterge olarak görmektedir. Cisco Systems tarafından açıklanan verilere göre tüm dünyada internet trafiği 2006-2011 döneminde 12 kat artarak aylık 23,9 eksabit düzeyine ulaşmış bulunmaktadır. Teknolojik trendleri araştıran IDC firmasının tahminlerine göre de 2011'deki veri miktarı yüzde 50'lik bir artış göstererek 2012'de dünyada 2,7 zetabit, bir başka deyişle 2,7 sektrilyon bit dijital veri oluşmuştur (Brynjolfsson ve McAfee, 2015: 81). Algılayıcı veriler (sensör verileri/dataları) her geçen gün artış göstermekte ve kalıcı olmaktadır. Çünkü gündelik hayatta bilgi ve enformasyonun işlendiği süreçler hızla dijitalleşirken süreçlerdeki özne, nesne, olgu ve olayların sayısal karşılıkları kod olarak kaydedilmeye başlanmıştır (Arslantaş vd., 2012: 5). 2011 yılında ağa bağlı cihazların sayısı insan nüfusunu geçmişken, 2025 yılı için 50 milyar algılayıcının internete bağlı olacağı ve büyük veri yığınları oluşturacağı öngörüsünde bulunmaktadır. Bu verilerin sadece yüzde 0,5'inin herhangi bir şekilde analiz edildiği belirtilmektedir (Lokke, 2018: 61). Veriyi yapılandırmak, analiz etmek ve ondan değer elde etmek (yüzde 25'inin potansiyel değeri olduğu tahmin ediliyor) en önemli amaç olarak görülmektedir (Davenport, 2014: 7-8). Her geçen saniye dahi artış gösteren ve tamamen işlenmemiş ham veri olarak ifade edilen bu kodlar Dijcks (2013: 2)'e göre, yapısal olmayan veya bugünkü tabirle herhangi bir ilişkisel veri tabanına aktarılamayan devasa verilerdir. Bu kadar büyük ve ciddi farklılıklar içeren verilerin kullanılmak üzere işlenmesi sürecini ifade etmek için kullanılan büyük veri (big data) kavramı,

bilgi ekonomisinin petrolü (Mayer-Schönberger ve Cukier, 2013: 22) olarak ifade edilmektedir. Bu tarz veri setlerinin işlenmesine olan ihtiyaç, beraberinde ağın, yapıların ve programlama dillerinin gelişmesine yol açmıştır. Burada büyük veri gibi verilerin çokluğundan/miktarından ziyade işlenmesi/analiz edilmesi ve değer elde edilmesi büyük önem taşımaktadır.

4. Büyük Verinin Paydaşları, Boyutları ve Özellikleri

Büyük verinin paydaşları; toplayanlar, kullananlar ve üretenlerdir. Büyük veriyi toplayanlar, verinin hangi parametreler doğrultusunda toplanıp kullanılacağına karar vermektedir. Kullananlar ise amaç doğrultusunda verileri kullanmaktadır. Son olarak üretenler ise gönüllü ya da gönülsüz bir şekilde veri üreten aktörlerdir. Üretilen veriler ise yapılandırılmış, yarı yapılandırılmış ve yapılandırılmamış olmak üzere üç şekilde bulunmaktadır. Üzerinde işlem yapmaya olanak sağlayan verilere yapısal veri denilmektedir. Her türlü düzen ve yapı içinde bulunabilen verilere ise yapısal olmayan veri denilmektedir. İnternet ve sosyal medya üzerinden elde edilen veriler yapısal olmayan veri özelliği taşımaktadır. Yarı yapısal veriler ise kendine has düzen ve yapı içerisinde bulunmaktadır. XML dosyaları, RSS verileri, işlem tabloları gibi dosyalardan gelen veriler yarı yapısal verilerdir (Holmes, 2017: 6).

Büyük veri "insanlar tarafından oluşturulan dijital ayak izleri" ve "makine verileri" olmak üzere iki geniş kategoride sınıflandırılmaktadır. İnsanlar tarafından oluşturulan dijital ayak izleri, internet üzerindeki etkileşimlerin büyümesiyle ortaya çıkan verilerden oluşmaktadır. Ayrıca bireyler, dijital sistemlerde günlük olarak gerçekleştirilen tıklamalar ve girilen etkileşimin geride bıraktığı bilginin farkında değildir. İnsanlar tarafından oluşturulan dijital ayak izleri büyüyerek devasa boyutlara ulaşmaktadır. İkinci büyük veri kategorisi ise makine verileridir. Bireylerin dijital ayak izlerini destekleyen çok sayıda güvenlik duvarı, yük dengeleyici, anahtar ve bilgisayar bulunmaktadır. Bunlar, güvenlik ve denetleme günlük dosyalarından, alışveriş sitesinde sepete konulduktan sonra geri bırakılan ürünlerin bilgisi olmak üzere geniş yelpazede bilgileri kaydetmektedir (Zadrozny ve Kodali, 2013: 2-3).

Büyük verinin kabul gören özelliklerinin başında veri hacmi (volume), veri hızı (velocity) ve veri çeşitliliği (variety) gelmektedir. Bu özellikleri ilk kez 2001 yılında Loug Laney "3V" olarak ifade etmiş, farklı yazarlar da 3V'ye doğruluk (verification) ile değer (value) özelliklerini de ekleyerek büyük verinin 5V üzerinden tanımlamasına katkıda bulunmuştur. ABD teknoloji şirketi SAS Enstitüsü ise hacim, hız ve çeşitliliğe iki boyut daha eklemiştir: değişkenlik ve karmaşıklık. Değişkenlik veri yığınlarının belli olaylar ve zaman dilimleri olarak nasıl izole edilebileceğini tanımlamaktadır. Karmaşıklık ise kompleks bağlamları işleme kapasitesini ve verile-

ri farklı sistemler karşısında dönüştürürken aynı zamanda bağlantıları da görmeyi ifade etmektedir (Lokke, 2018: 60-61). Büyük veri; yüksek hacim, hız ve çeşitlilik ile karakterize edilen bilgiyi temsil ederken değere dönüştürülmesi için özel teknoloji ve analitik yöntemler gerektirmektedir (Mauro vd., 2015: 97-100). Davenport (2014: 13) ise büyük veriye yönelik problemleri bir yaklaşımın sonucu olarak ortaya çıkan "çıkarıcılık" (venality) kavramına da göndermede bulunmaktadır.

Hacim (Volume): Büyük veri kavramının ilk kelimesi verinin hacmine göndermede bulunmaktadır (Zadrozny ve Kodali, 2013: 1). Bilgi ve iletişim teknolojileri aracılığıyla üretilen ham verilerin hacimleri geleneksel veri üreten sistemlerden çok daha fazlasını kapsamaktadır. Her iki yılda bir, insanlık tarihi boyunca üretilen tüm veri ve içerik kadar yeni veri ve içerik üretilmektedir. Veri üretim hızının ivmelenecek artışı beraberinde yeni hacim ölçülerini getirmiştir. Bit, byte, kilobyte, megabyte, gigabyte ile günümüzde terabyte gibi veri hacminin ölçümü için kullanılan kavramlara yenileri eklenmektedir. Petabyte (1000 terabyte), exabyte (1000 petabyte), zettabyte (1000 exabyte) ve son olarak yottabyte (1000 zettabyte) dijital ortamda üretilen verinin hacmini belirlemenin zorluğunu göstermektedir.

Hız (Velocity): Büyük verinin üretim hızında dikkat çekici bir artış yaşanmaktadır. Hızla artış yaşanan veri, o veriye muhtaç olan işlem sayısının ve çeşitliliğinin de aynı hızda artmasına neden olmaktadır (Ünal, 2015: 9) Kesintisiz bir şekilde oluşmakta olan veri ve içerikler aynı anda ilişkilendirilmekte, internet üzerinden erişilir kılınmaktadır. Sosyal paylaşım ağları aracılığıyla gerçekleşen veri akışları, kullanıcıların en küçük bir olaya anlık cevap vermelerini tetiklemektedir. Veri tarihi incelendiği takdirde, insanlığın hiçbir zaman diliminde bu kadar yoğun bir veri üretiminde bulunmadığı görülecektir. 2003 yılına kadar 5 trilyon gigabayt veri üretilirken 2011 yılından sonra her iki günde bir 5 trilyon gigabayt veri, 2015 yılından sonra ise 10 saniyede bir 5 trilyon gigabayt veri üretilmiştir (Saka ve Sayan, 2016: 90). Hız, çoğunlukla gerçek zamanlı diye bilinen süre içerisinde (yani olay olurken) veri yığınlarının ne kadar hızlı işlenebildiğine göndermede bulunmaktadır (Lokke, 2018: 60).

Çeşitlilik (Variety): Çeşitlilik hem yapılandırılmış hem yapılandırılmamış farklı veri formatlarını, e-posta, video ve metin dosyalarını işleme kapasitesini tanımlamaktadır (Lokke, 2018: 60). Veri çeşitliliği, verilerin farklılık ve zenginliğinin bir ölçüsü olarak ifade edilmektedir. Telefonlar, tabletler veya bütünleşik devrelerden gelen, sosyal medya, sensör verileri, CRM dosyaları, dokümanlar, imajlar, videolar, e-postalar vb. akla gelebilecek tüm veriler, veri kaynak ve tipleri, her türlü çeşitlilikte veri tipi ile uğraşılması gerekmektedir. Geleneksel platformlar çeşitlilik arz eden verileri depolama ve analiz etmekten yoksun olduğu için yoğun veri kümeleriyle baş edememektedir. Verilerin; yapılandırılmış, yapılandırılmamış

ve yarı yapılandırılmış bir şekilde sayı, metin, video, fotoğraf, ses ve diğer farklı formatlarda bulunması çeşitliliğin göstergesidir. Yeni algılayıcı ve servislerin eklenmesiyle birlikte geleneksel yoldan elde edilen verilerin aksine veri kaynakları ciddi çeşitlilik göstermektedir. Örneklem olarak alınan verinin çeşitlilik arz etmesi büyük zorlukları barındırmamasına rağmen doğru analizler yapma fırsatını da beraberinde getirmektedir.

Değer (Value): Karmaşık bir yapı izlenimi verse de büyük veri gerçeklik hakkında çok şey söylemektedir. Büyük verinin artı değer yaratıyor olması gerekmektedir. Çünkü bu veriler gerçekliğin bir yansıması olma özelliği taşımaktadır. Büyük verinin en önemli özelliği yukarıdaki özelliklerin genel bir sonucu olarak analiz ve anlam çıkarmayı mümkün kılmasıdır. Günümüz veri ortamının geçmişe dönük en büyük farklılığı da verilerin analiz edilerek anlamlı hâle dönüştürülebilmesi olmuştur. Daha net bir ifadeyle 'değer, kaostan ayıklanmıştır' (Gantz ve Reinsel, 2011). Bu tarz geleneksel olmayan veriler, pazarlama dünyası için yeni pazarlar bulma yolunda ciddi yönlendiricilerdir, bu sebeple de çok değerlidirler (Dijcks, 2013: 3-4). Değer kavramıyla "veriden anlam çıkarma"ya vurgu yapılmakta olup beraberinde verim ve faydayı sağlaması beklenmektedir. Bilinçli ya da bilinçsiz üretilen dijital enformasyon kaynaklarının teknolojik aygıt ve yazılımlarla analiz edilmesine olanak sağlayan büyük veri, ekonominin odak noktası olmasının yanında artık siyasal ve ekonomik boyutları içinde barındırmaktadır (Saka ve Sayan, 2016: 88-89). Geçmişte veriler toplandıktan sonra analiz edilirken teknolojik gelişmelerle birlikte verinin toplanmasıyla analiz edilmesi arasındaki süre çok kısalmış ve gerçek zamanlı bir şekilde analiz gerçekleşmektedir (Gürsakil, 2014: 32)

5. İki Ucu Keskin Kılıç: Büyük Veri

Günümüz bilgi ve iletişim teknolojileri çevrim içi ortamdaki bütün hareketlerimizi birer "davranışsal kalıntı" olarak kaydetmektedir. (Tong vd., 2008: 534). Dijital birer iz olarak bu davranışsal kalıntılar yine bilgi ve iletişim teknolojilerinin sunmuş olduğu imkânlarla analiz edilmekte, kalıntıyı bırakan kişilere ilişkin bilgi üretilmektedir. Önemli bir konu hâline gelen büyük veri, iki ucu keskin bir kılıç olarak görülmektedir. İnsanlara kolaylık getirdiği kadar belli riskler de taşımaktadır. Veri toplama, saklama ve kullanma sürecinde kişisel bilgilerin kolayca sızması en önemli tehlike olarak gösterilmektedir. Büyük veri güvenliği ve mahremiyetin korunması önemli konulardan birisidir. İnternetin hızla gelişmesiyle birlikte insanlar internette çok fazla veri izi bırakmaktadır. Büyük veri çağının ortaya çıkışı sosyal ilerleme için önemli fırsatlar sağlamanın yanında toplum için birçok bilgi güvenliği tehdidini de beraberinde getirmiştir. Kişisel veri gizliliği en önemli endişe kaynağı olarak görülmektedir (Zhang, 2018:

275-277). Günümüzde çevrim içi ortam kullanıcıların tüm bilgilerinin toplandığı ve sınıflandırılarak depolandırıldığı büyük veri havuzuna dönüşmüştür. Kullanıcılar bilgisi olmadan takip edilmekte, çevrim içi tüm bilgileri kaydedilip reklam endüstrisi şirketleriyle paylaşılmaktadır (Uluk, 2018: 43-44). Ayrıca çevrim içi bağlantıya sahip cihazların kendi aralarındaki iletişiminin yanı sıra bireylerle kurdukları etkileşim de kullanıcılar hakkında bilgi elde etmeye imkân sunmaktadır. Önümüzdeki yıllarda etkileşimin ve internetin hemen her nesnenin içine gireceği ve farklı nesnelere ortak amaçlar için hareket edeceği bir döneme girileceği öngörülmektedir (Şeker, 2018: 164).

Büyük veriyi bir serebreskopa benzeten Stephens-Davidowitz (2018: 9-10), kullanıcıların ürettiği verilerin insan ruhuna dair eşi bulunmaz bir içgörü sunduğunu ifade etmektedir. Steven Pinker da kullanıcıların karşılarında tepki verecek bir insan olmadığına bazı istek ya da korkularının yükünden kurtularak klavyelerinin mahremiyetinde en tuhaf şeyleri itiraf ettiklerine dikkat çekmektedir. Dijital çağdan önce insanların utanç verici düşüncelerini diğerlerinden sakladıklarını, dijital çağda da saklamaya devam ettiklerini ama internetten ve özellikle de anonim kalmalarını sağlayan Google gibi sitelerden saklamadıklarını anlatan Stephens-Davidowitz (2018: 49), bu sitelerin bir tür "dijital hakikat serumu" işlevi gördüğünü ifade etmektedir: "Büyük veri, bize insanların gerçekten ne istediklerini ve ne yaptıklarını, ne istediklerini söylediklerini ve ne yaptıklarını söylediklerini nihayet görme imkânı tanıyor. Dürüst veriler sağlaması büyük verinin en güçlü yanlarından biridir. Ayrıca yeni türde veriler sunması da güçlü yanıdır." Bu yönüyle büyük verinin gerçek hayattaki sezgi, içgörü ve altıncı hissin bilimsel bir karşılığı olabilmesi söz konusudur. İnsanların artık çevrim içi ortamda aradıkları daha değerli hâle gelirken, John Battelle'in Google'ı "niyetlerin veri tabanı" (Arthur, 2017, s. 63) olarak adlandırması da yerinde bir tespittir.

Büyük veriyle birlikte ortaya çıkan en önemli sorunlar şu başlıklar altında sıralanmaktadır:

- **Bağımlılık:** Düşünme işi çok daha kolay ve hızlı oldukları için yazılım ve algoritmalara bırakılmaktadır.
- **Kafa Karışıklığı:** Kullanıcının aldığı kararın kendisi mi yoksa dijital asistan tarafından mı yönlendirildiğini bilmemesidir.
- **Kontrol Kaybı:** Mantiğini takip edemediğimiz ve işleyişini kavrayamadığımız yapay zekânın sezindiği şeyin doğru olup olmadığını bilememektir.
- **Feragat:** Kişisel takvim eş güdümü, randevu alınması ya da basit e-postalara cevap yazılması gibi görevleri, bunları yerimize hâlledecek sistemlere bırakmak. Olumsuz sonuçlar ortaya çıktığında suçu buluta, bota veya yapay zekâyâ atmaktır (Leonhard, 2018: 68).

Taylor Armerding de büyük veriyle bağlantılı ayrımcılık, güvenlik ihlali, anonimliğe veda, düzenlenmemiş bilgi işlem ve devlete verilen geniş yetkiler başlıkları altında sorunlara neden olduğunu ifade etmektedir (Lokke, 2018: 64-65). Arre Zuurmond da bilgi teknolojileri sayesinde ortaya çıkan ağ örgütlenmesinin dostça görünüşünün arkasında ekonomik süper güçlerin gizlendiğini ve yeni bir enfok-rasinin (enfomasyon bürokrasisi) ortaya çıktığını söylemektedir (Mul, 2008: 59).

Büyük veri beraberinde etik sorunlara da neden olmaktadır. Hem bireyler hem de kuruluşlar için büyük veri etiği noktasında neyin çerçeve olarak kabul edilebileceğine ilişkin dört ortak unsur bulunmaktadır (Davis ve Patterson, 2012: 2-3).

- **Kimlik:** Çevrim dışı ve çevrim içi kimlik arasındaki ilişki nedir?
- **Gizlilik:** Verilere erişimi kim kontrol etmelidir?
- **Mülkiyet:** Verilere kimin sahip olduğu, aktarılmasındaki haklar ve kişilerin yükümlülükleri nelerdir?
- **İtibar:** Hangi verilerin güvenilir olduğunu nasıl belirleyebiliriz?

6. Büyük Veri ve Gözetim İlişkisi

Dijital ortamların en temel özelliği olarak kayıt olgusu beraberinde önemli sorunları da getirmektedir. Sanayi devriminden sonra ekonomide verimliliği artırmak için fabrika ve mahkûmları denetlemek adına hapishanelerde bir önlem olarak yaygınlaştırılan gözetim (panoptikon) olgusu bilgi toplumunda teknolojik gelişmelerle eş değer olarak dönüşmüş, birey ve toplum hayatındaki yaygınlığı genişlemiştir. Sanayi toplumunda zaman ve mekânın kısıtlamalarından dolayı sınırlı kalan gözetim, bilgi ve iletişim teknolojileriyle gelen küresel iletişim ortamında her bir bireyin gündelik hayatının önemli bir kısmının denetime açık hâle gelmesini mümkün hâle getirmiştir. Gözetimin yaygınlaşması sanayi dönemi için kullanılan panoptikon kavramını da dönüştürmüş, bilgi ve iletişim teknolojileriyle yaygınlaşan gözetim için süperpanoptikon/sinoptikon kavramları tercih edilmiştir.

Sosyal medya platformlarının yaygınlaşmasıyla birlikte büyük verinin kullanım alanı yaygınlaşmıştır. Bireyler gündelik yaşamlarının bir parçası hâline gelen sosyal medya platformlarında kendileri hakkında önemli bilgileri paylaşmaktadır. Bu platformlardan elde edilen veriler işlenerek kullanıcılara yönelik olarak reklam ve pazarlama amaçlı olarak kullanılmaktadır. Bu durum beraberinde gizlilik sorunsalına neden olmaktadır. Sosyal medya platformlarında gizliliğin yanı sıra ticari ve resmî gözetim de söz konusu olmakta ve sıkça tartışılmaktadır. Helland (2011: 1), büyük verinin varlığını, gerçekleştirmiş olduğu işlemleri ve geleceğe yönelik ortaya koyduğu çıkarsama ve varsayımları bilerek artık temiz bir dünyada yaşıyormuş gibi yapmamızın mümkün olmadığını ifade etmektedir.

Günümüzde gözetim teknolojilerinin yanı sıra bireyler kullandıkları teknolojik ekipmanlarla gönüllü olarak gözetlenmeye tabi olmaktadır. Büyük verinin gözetimde oynadığı rol üzerine yapılan çalışmalar yetersiz kalmaktadır. Bilgi ve iletişim teknolojilerinde güvenlik ve gizlilik her zaman önemli iki temel kavramdır. Büyük veri çağında veri hacmi hızla büyüdükçe, daha ciddi güvenlik riskleri ortaya çıkarken geleneksel veri koruma yöntemleri büyük verileri korumada yetersiz kalmıştır. Veri toplama sırasında kişisel mahremiyetin korunmasının yanı sıra kullanıcının izni ile elde edilmiş olsa bile depolama, iletim ve kullanım sırasında kişisel gizlilik sağlanmamaktadır (Chen vd., 2014: 203). Sosyal medya kullanıcıları gizlilik ayarlarına güvendikleri dijital platformlarda paylaşım yapma konusunda gönüllü davranmaktadır. Ancak son dönemde dijital platformlardaki gizlilik kavramı dönüşmüştür. Dijital platformların gözetleme ve dikizlemeye açık olması bu durumun bir göstergesidir. Büyük veri, insanların günlük yaşamının birçok yönünü bilgisayarlı verilere dönüştürmede kritik bir rol oynamaktadır ve böylece yeterli kaynaklara sahip aktörlerin benzeri görülmemiş bir ölçekte gözetleme yapmalarına olanak sağlamaktadır (Kubitschko, 2015: 79). Teknolojinin kendi başına gözetleme yapmayacağını ifade eden Kubitschko (2015: 85), “Teknolojileri ve gözetimi koşullandıran - yasal çerçeveyi belirleyen politikaları kullanan -bireysel, kolektif, örgütsel, kurumsal, aktördür. Bilgi ve iletişim teknolojilerinin ve altyapılarının neredeyse her yerde mevcut yayılmasının ardından, gözetim, çok çeşitli kuruluşlar tarafından hayata geçirilen sıradan bir uygulamaya dönüşmüştür.” ifadelerini kullanmıştır.

Azınlık Raporu (Minority Report) filminde psikologlar/kâhinler suçları gerçekleşmeden önce durdurmak amacıyla polisle ortaklaşa hareket etmektedir. Bu noktada büyük verinin de suçlar gerçekleşmeden önce harekete geçmek konusunda kullanılması tartışılmaktadır. Algoritmaları temel alan suç önleme için sosyal medya, cep telefonu konumları ve trafik verileri kullanılmaktadır. Bazı vakalarda sosyal hizmet uzmanı ya da polis memuru, suç işlemesi muhtemel kişileri tek tek ziyaret etmektedir (Leonhard, 2018: 62). Suç faaliyetleriyle alakalı arama motorlarında yapılan aramalar ile suç faaliyetleri arasında bağıntı olduğuna dair kanıtların giderek arttığına göndermede bulunan Stephens-Davidowitz (2018: 199-200), bir suç işlenmeden bireylerin peşine düşmenin özel hayatın ihlali olduğunu söylemektedir: “Hükûmetin binlerce, yüzbinlerce insanın arama verilerine sahip olması ile bir bireyin arama verilerinin polisin elinde olması arasında etik açıdan devasa bir uçurum vardır.” Örneğin Çin Halk Cumhuriyeti’nde 2020’den itibaren bütün vatandaşlara bir güven puanı verilmesi planlanmaktadır. Mobil uygulamalarla izlenen her kişi A, B, C, D gibi puanlanacak; A notunu alan makbul vatan-

daşlar en iyi okullara gidebilecek, kolay iş bulabilecek, kolay kredi alacak, devlet katındaki her işlerinde ihtimam görecek, D ile damgalanmış olanların trene ve uçağa bile binemeyecekleri iddia edilmektedir. Bankaların kredi notu olarak aşına olunan bu durum ekonomiden çıkıp hayatın her alanına yayılmaktadır. Böylece mahremiyetin tanımı değişmektedir. Devamında giyilebilir elektronik ürünlerden Google gözlüğüne, şoförsüz otomobillerden akıllı buzdolabına kadar elektronik bütün araç gerecin akıllanmaya başlaması gelmektedir. Nesnelerin internetiyle birlikte akıllı gereçlerin birbirleriyle iletişim kurmaları yaygınlaşacak, akıllı fabrikalar, akıllı şehirler gibi büyük ölçekli endüstriyel uygulamalarla devam edecektir. Arslantunalı (2019: 147-149), akıllı ev kapıyı açar açmaz ısınma, aydınlatma gibi işlevleri yönetmeye başlıyorsa ne zaman evde olduğumuzu/olacağımızı da bileceğine dikkat çekmekte ve böylece verilerin işlenmiş olacağına vurgu yapmaktadır. Akıllı cihazların kullanıcıların sunduğu veriler sayesinde akıllı hâle geldiğine işaret eden Arslantunalı, veri akışının ve izleme tekniklerinin merkezietçi olmayışının içimizi ferahlatamayacağını söylemektedir: “Çünkü bütün kişisel bilgilerimize erişebilecek durumdaki hükümetler, açık açık ya mahremiyet ya güvenlik açığı şantajını yapıyorlar, daha çok bilgi istiyorlar... Pek yakında nesnelerin interneti ile yüz milyonlarca -kısa sürede milyarlarca- akıllı alet edevat, bizi gözetliyor olacak. Panoptikon yerini elektronik süperpanoptikona mı bıraktı, gözetim faaliyetlerini sürdüren çok sayıda iktidar merciinin olduğu polioptikona mı, yoksa herkesin birbirini gözetlediği omnioptikona mı?” Büyük veri, gözetimi sağlamasının yanında vahim sonuçlar doğurabilecek tehlikeleri de barındırmaktadır. Kişisel verileri ele geçirmenin yanında bu verilerin istenilen şekilde değiştirilebilmesi de yakın gelecek için tehdit olarak görülmektedir. Bu durumun kullanıcıları savunmasız bırakacağı ifade edilmektedir (Goodman, 2016: 180).

George Orwell’in “1984” kitabı gözetim kavramı ekseninde yoğun olarak değerlendirilse de yazılı ve sözlü dilin üzerindeki kontrolün beraberinde alternatif fikirleri ve algılamayı güçleştireceğine işaret etmektedir. Orwell, (1960: 39) "yeni dil" (newspeak) adında, düşünce ve ifade özgürlüğünü ciddi şekilde kısıtlayan bir kurgusal dilden söz ederken, günümüzde internetin milyonlarca kanal sunması bu kâbustan uzak olduğumuzu düşünmemize sebep olmaktadır ancak giderek daha büyük sıklıkla kararları bizim adımıza makinelerin verdiği gerçeğiyle yüzleşmekteyiz. Bu süreçte daha yumuşatılmış bir kontrolün gerçekleştiğini "yeni Orwellci toplum" olarak karakterize eden Markoff (2017: 18-19), şu anda attığımız her adımın, aldığımız her nefesin, büyük birader tarafından olmasa da giderek artan sayıdaki ticari küçük biraderler tarafından takip ve kaydedildiğine işaret etmekte, bu durumun sözde bizlere yardım etmek amacıyla rızamızla gerçekleştiğini ifa-

de etmektedir. Markoff, böyle bir dünyanın özgür olabileceğine şüpheyle yaklaşmaktadır. Verilerin depolanması, analiz edilmesi ve kullanılması arasındaki ilişki depolamayı bir sorun hâline getirmektedir. Bu noktada depolama ve analiz arasındaki ilişki, silahlar üzerine yapılan tartışmalara benzetilmektedir. Silahın kimseyi öldürmediği, tetiği çekecek bir parmağa ihtiyaç duyulduğuna göndermede bulunulmakta ancak silah sayısının artışının beraberinde tetik çekilmesine imkân sunduğu göz ardı edilmektedir. Aynı şekilde ne kadar çok verinin depolanmasına izin verilirse özel hayat için o denli fazla sorun ortaya çıkmaktadır. Çünkü verilerin analizi sırasında kötüye kullanım ihtimali artacaktır (Lokke, 2018: 64).

Yukarıda ifade edilen eleştirilere yönelik olarak Google CEO'su Eric Schmidt'in, "Başkalarının bilmesini istemediğiniz bir şeyiniz varsa belki de hiç yapmamanız lazımdır." şeklinde verdiği cevap ise tutarsızlık göstermektedir. Saklayacak bir şeyin olmaması gerektiği yanılıgısı içeren bu ifade ya tamamen gözetimi kabul etmeyi ya da şüphelenilen suçlular kategorisine girmeyi içermektedir (Goodman, 2016: 108-109). Gündelik hayatın bu denli teknolojiyle bütünleşmesi "dijital kirliliğe" de neden olmaktadır. Tıpkı 20. yüzyılın başında endüstriyel atıkların nehirlere boşaltılması ve sokağa çöp atılmasının normal olarak görülmesi gibi günümüzde de dijital faaliyetlerin uzun vadede ortaya çıkarabileceği sonuçları algılama konusunda sıkıntılar yaşanmaktadır (Goodman, 2016: 113).

7. Gözetimin Kapitalistleşmesi: Büyük Veri

Bilgi ve iletişim teknolojilerinin sağladığı ve yaygınlaştırdığı kayıt imkânı, gözetim olgusunun yanında ekonomik amaçlı olarak da kullanılmaktadır. Bilgi ve iletişim teknolojilerinin aracılığıyla üretilen veriler, bilgi toplumunun petrolü olarak görülmektedir. Bazı yazarlar da verileri ham petrol olarak görmektedir, değerinin ortaya çıkması için tasnif edilerek analize tabi tutulması gerektiğini vurgulamakta, yoksa kullanımının mümkün olmadığını söylemektedir (Palmer, 2006). Günümüzde verileri organizasyonel süreçlerde kullananların rakiplerine karşı rekabette öne geçtiği ifade edilmektedir. Teknoloji bilgi ile yaratılırken bilgiyi elde etmenin yolu da verilerin işlenmesinden geçmektedir. Ayrıca dijital ortamda bireylerin bütün verilerinin kaydedilmesi bu verilerin reklam amaçlı kullanımına imkân sunmaktadır. Çünkü çevrim içi ortamda yer alan veriler kişilere ilişkin bilgi sunmanın ötesinde, bu bilgilerin kişilere dönük reklam ve pazarlama stratejisinde kullanılmasına olanak sağlamaktadır. Bireylerin çevrim içi ortamda/dijital dünyada kaydedilen verilerinin analiz edilerek değer elde edilmesi işlemleri/süreci günümüz işletmeleri için büyük avantajlar sunmaktadır. Günümüzde kuruluşlar daha önce sahip olduklarından daha fazla veriye sahiptirler. Teknolojideki gelişmeler, büyük hacimli

veri üretmede kritik bir rol oynamaktadır. Veri tabanları sadece büyümekle kalmıyor, aynı zamanda gerçek zamanlı hâle gelmektedir. Gelişen teknolojiler ve işlem gücü, kuruluşlara bir kişinin çevrim içi ve çevrim dışı davranışlarına dayalı olarak daha karmaşık ve derinlemesine bireysel profiller oluşturma yeteneği sağlamaktadır (Al-Khoury, 2012: 1). Buna karşın hükûmetlerin teknoloji geliştirme hızına ayak uydurması zor olacak ve vatandaşlar gerçekten de zor anlar yaşayacaktır (Al-Khoury, 2012: 4).

İstihbarat servislerinin kullanmış olduğu en gelişmiş gözetleme ve veri analizi biçimleri, büyük işletmelerin pazarlama stratejilerinin de ayrılmaz bir parçası olmuştur. Göz hareketlerini, görsel ilginin nelerde ve ne kadar süreyle yoğunlaştığını takip eden ekranlar ve diğer gösterim biçimleri artık yaygın olarak kullanılmaktadır Bir kişinin tek bir web sayfasına şöyle bir göz atması kılı kırk yararcasına analiz edilmekte; gözün bu sayfayı taraması, durması, hareket etmesi ve bazı alanlara ötekilerden öncelik verip daha fazla dikkat etmesi bakımından birtakım ölçümlere tabi tutulmaktadır (Crary, 2015: 54-55).

Büyük verilerin korunup korunamayacağına ilişkin mülkiyet hakkı sorunu ortaya çıkmaktadır. Verinin sahibi kim, verilere sahip olunabilir mi soruları hâlâ cevap beklemektedir (Hoeren, 2014: 751). Otomatik olarak oluşturulan veriler, veriyi üreten kişi değil de makine ise belirsizlik daha da büyük bir soruna neden olmaktadır (Hoeren, 2014: 754). Büyük verinin toplandığı ve depolandığı dijital ortamların tümü küresel şirketlerin denetiminde bulunmaktadır. Bu noktadan hareketle büyük veri üzerine tüm sosyal süreci anlamlandırmak adına yeni bir çerçeveden bakmak gerekmektedir. Bu çerçeve, sömürgeciliktir ve burada salt bir metafor olarak da tarihsel sömürgecilik biçimlerinin yankısı veya basit bir devamından öte 21. yüzyılın yeni bir sömürgecilik biçimine atıfta bulunmak için "veri sömürgeciliği" kavramı kullanılmaktadır. Veri sömürgeciliği tarihsel sömürgeciliğin yağmacı uygulamalarını soyut hesaplama yöntemleriyle birleştirmektedir. Big data ile birlikte tüm dünyadaki sosyal yaşam bir şekilde sermaye için açık bir kaynak hâline gelmektedir. Küresel veri akışları tarihsel sömürgeciliğin toprak, kaynak ve bedenlere el koyması kadar kapsamlı gerçekleşmektedir. Veri sömürgeciliği hem harici olarak küresel ölçekte hem de içsel olarak çalışmaktadır. (Couldry ve Mejias, 2019: 336-337). Veri sömürgeciliğinin yanı sıra veri tekelliliği de tartışılmaktadır. Büyük verinin sınırlı sayıda şirket tarafından kendi lehlerine kullanmaları verinin gücünün suistimaline neden olmaktadır. Goodman'ın (2016: 100) "veri simsarları" olarak ifade ettiği şirketler hızla büyüyen veri gözetleme sektörünün önemli isimleri olarak dikkat çekiyorlar. Acxiom, Epsilon, Datalogix, RapLeaf, Reed Elsevier, BlueKai, Spokeo ve Flurry gibi isimleri duyulmamış bu şirketler Goodman'a göre yıllık 156

milyar dolar veri geliri elde etmektedir: “Veri simsarları aradıkları bilgileri internet servis sağlayıcılarımızdan, kredi kartı şirketlerinden, operatörlerden, bankalardan, kredi ofislerinden, eczanelerden, motorlu taşıt dairelerinden, bakkallardan ve artan online aktivitelerimizden buluyorlar. Kullandığımız sosyal ağlara her gün ücretsiz bir şekilde verdiğimiz tüm veriler –tüm like’lar, dürtmeler ve tweetler–daha sonra reklamcılar ile pazarlamacılara satılmak üzere etiketleniyor, konumu belirleniyor ve sıralanıyor. Eski pazarlamacılar bile artık müşteri verilerini ikincil bir gelir kapısı olarak kullanabileceklerini fark etmiş durumda. Hatta kimileri, bunun sattığı mallardan bile değerli olabileceğini görüyor. Bu yüzden, şirketler bu yeni gelir kaynağından bir an önce kâr sağlamak için büyük bir yarışa giriyor ve daha önce maliyet merkezi olan veri altyapılarını birer kâr merkezine dönüştürüyorlar.” (Goodman, 2016: 100). Sadece Acxiom Corporation, 23 binden fazla bilgisayar sunucusu çalıştırarak, her gün 50 trilyondan fazla eşsiz veri işlemini toplayarak düzenleyip analiz etmektedir. Bu veri bankaları üzerinden ABD’deki hanelerin yüzde 96’lık bir bölümüne ulaşabilmektedir. 700 milyon tüketicinin verilerinin yer aldığı Acxiom’un arşivinde her bir insan için oluşturulan ayrı ayrı profillerde ırk, cinsiyet, telefon numarası, kullanılan otomobil türü, eğitim seviyesi, çocuk sayısı, evinin metrekaresi, mevkisi, son alışverişleri, yaşı, boyu, kilosu, medeni hâli, siyasi görüşü, sağlık sorunları, mesleği ve evcil hayvan sahipliği ile hayvanın türü gibi yüz elliden fazla özel başlık bulunmaktadır. Acxiom ve diğer veri simsarı olarak faaliyet gösteren şirketlerin amacı alternatif şekillerde "davranışsal saptama", "tahmini saptama" veya "kişiye özel davranışsal analiz" gibi yöntemlerle kullanıcıya ve hayatına dair her şeyi elde etmektir. Bu yöntemlerle kullanıcıların net bir şekilde tanınmasına olanak veren bilgilerin toplanması karar alım sürecindeki şirket, reklam ve pazarlamacılar arasında yüksek fiyatı verene satışını sağlamaktadır. Veri simsarları, topladıkları verilerin değerini olabildiğince arttırabilmek için kullanıcıları daha spesifik gruplara ve profillere ayırmaktadır (Goodman, 2016: 101).

8. Büyük Veriyi Kamuoyu Yararına Kullanmak

Toplumların dijitalleşme pratikleri beraberinde daha fazla verinin elektronik ortamda oluşmasına neden olmaktadır. Bilgi ve iletişim teknolojilerinin sunduğu devasa veriler hükümetler tarafından gözetim amaçlı, sermaye grupları tarafından reklam ve pazarlama amaçlı kullanılırken, aynı zamanda gazeteciler tarafından da kamu yararına konular hakkında rapor oluşturmak, iktidarlara hesap sormak için araştırma amacıyla da kullanılmaktadır.

Verilerin kamu yararına kullanımının ilk örneklerinden biri 1854 yılında Londra’da ortaya çıkan kolera salgınında yaşanmıştır. Doktor John Snow, hastalığın baş gös-

terdiği yerlerin bir taramasını yaptıktan sonra salgının çıktığı noktanın su pompası olduğunu bulmuştur. Bu detaylı haritalama sayesinde Doktor Snow hem kolera salgının nedenini ortaya çıkarmış hem de salgının yayılmasını önlemiştir. Kullanılan yaklaşım, veri işleme ve büyük veri uygulamalarının öncüsü olarak görülmektedir (Lokke, 2018: 59).

Dijital veriler hesap verilebilirlik, halkı bilgilendirme ve katılımını sağlamada kilit bir araç olarak görülmektedir. Dijital çağda gazeteciliğin kamu yararını gözeten yayıncılık anlayışını geliştirmesi için veri gazeteciliği gibi yeni yöntemleri kullanması gerekmektedir. Veriyi anlamlandırarak içinde saklanan bilgiyi kamu yararına ortaya çıkaracak "dijital bekçi köpekleri" olarak veri gazetecilerine ihtiyaç söz konusudur. Bu yöntemle aynı zamanda okuyucuların da ilgisini çekmek mümkündür (Felle, 2016: 85, 93). Verinin gazetecilik bağlamında ele alınması 2010 yılında ortaya çıkan Wikileaks sızıntılarıyla başlamıştır. Veriyi ayıklama, düzenleme, işleme ve ilişkilendirme aşamalarını kapsayan bu yeni pratik, veri gazeteciliği olarak ifade edilmektedir. Gizli verilerden hareket ederek "gölgelerde saklı olanı" arayan, veriyi çözümlenip içinde saklı ilişkileri bulmayı içeren veri gazeteciliği, demokrasinin önemli bir aracı olarak görülmektedir (Erkmen, 2020: 65).

9. Sonuç

Tarihi boyunca sürekli veri üretmiş olan insanoğlu, kendisi ya da gelecek nesillerin kullanımı için bunları kayıt altına almıştır. İlk kayıtlar mağara duvarlarına çizilen şekillerden oluşurken zamanla yazı, matbaa ve günümüzde de dijital teknolojiler birer kaydetme biçimi olarak görülmüştür. Büyük veri, kullanım alanı giderek artan, neredeyse her sistem ya da organizasyon içerisinde kullanılan ve elde edilen verilerin çeşitli teknolojiler ile saklandığı bir sistem olarak değerlendirilmektedir. Hızla gelişen büyük veri kavramıyla sadece veri oluşturma, depolama, analiz etme süreçleri dikkate değer olarak incelenmiş olsa da giderek artan önemi ile farklı araştırmalara konu olacağını göstermektedir. Büyük veri; ekonomik, sosyal, kültürel ve teknolojik bir fenomendir. Büyük veri, genel olarak kamusal yaşam ve medya bağlamında incelenmeye ve teorileştirilmeye çalışılmaktadır. Büyük veri, aralarında iş dünyası, siyaset, tıp ve neredeyse tüm doğa bilimleri ve sosyal bilimler de dâhil olmak üzere pek çok alanda devrim yaratmaktadır. Birbirleriyle iletişimde bulunan insanların yanı sıra akıllı cihazların da sürece dâhil olmasıyla herkes bilerek ya da farkında olmadan veri üretmektedir. Reklamcılık ve pazarlamadan, eğitim ve sağlığa kadar hayatın her alanında kullanılan büyük verinin oluşumuna en fazla katkıyı sağlayan araç sosyal medya platformlarıdır. Kullanıcıların kişisel bilgilerinin en yoğun olarak

bulunduğu sosyal medya platformları, bu verileri başta reklam ve pazarlama amaçlı olmak üzere kişi ve kurumlarla paylaşmaktadır. Bu sorunun en somut örneklerinden biri, sosyal paylaşım platformu Facebook'un yaklaşık 50 ile 100 milyon arasındaki kullanıcı bilgisini Cambridge Analytica şirketiyle paylaşmasıdır. Büyük veri teknolojilerindeki gelişmelerle eş değer etik açıdan ortaya çıkan sorunlarda artış yaşanmıştır. Bazı şirketler her bir çalışanın bastığı her bir tuşu kaydediyorlar. E-postalar, telefon kayıtları, veri tabanı sorguları ve dosya erişimleri, şirkete girişler ve çıkışlar, internet aramaları ve daha nice veri türü -bazen işçilerin bilgisi dâhilinde bazen haricinde- toplanabiliyor. Şimdilik veri toplamanın amacı daha etkili bir yönetim sergilemek ve çalışanların performanslarını daha sağlıklı değerlendirmek olsa da ileride bu verilerin başka amaçlar için de kullanılması söz konusudur (Ford, 2018: 116). Büyük veri sadece ekonomik boyutu olan bir kavram olmaktan çıkmıştır. Ekonomik boyutuyla büyük veriyi kullanan küresel şirketlerin yanında devletler/yönetimler de büyük veriyi yönetim/denetim zihniyetine dönüştürmektedir. İnsan teknolojisinin yerine insansız teknolojinin geçirilmesiyle denetim daha da yaygınlaşacaktır (Ritzer, 2011: 107-116). Günümüz insanı çevresindeki tüm nesnelerin, hizmetlerin, maddi ürünlerin kuşatmasıyla karşı karşıya kalmıştır. Baudrillard (2004: 15) bu gerçekten hareketle şu ifadeleri kullanmaktadır: “Bu insanların gündelik alışverişi artık başka insanlarla değil; daha çok istatistiksel olarak yükselen bir eğriye göre mal ve iletilerin edinilmesi, algılanması ve güdümlenmesi biçimini taşıyor.”

Büyük veri beraberinde bir kısmı yeni olmayan etik sorunları da getirmiştir. Ancak büyük verinin kullanım alanları ve sağladığı avantajlara yoğunlaşan çalışmalar bu etik sorunların çok fazla tartışılmamasına neden olmuştur (Zwitter, 2014: 1). Büyük veri kaynaklı etik sorunların başında kullanıcı gizliliği ve bilgi güvenliği sorunları gelmektedir. Bu sorunlar aynı zamanda hukuki boyutu olan etik sorunlar olarak görülmektedir. Büyük verinin üreticileri olarak kullanıcılar, eğer sorulması durumunda verilerinin başka yerde kullanılmasına izin vermeyecektir. Sosyal medya platformlarını kullanan bireyler takip edilmekte ve tüketilebilir, pazarlanabilir bir ürün hâline gelmektedir. Her bir kullanıcının davranışları parça parça alınarak satılmaktadır. Bu ifade, büyük verinin hem hukuki hem de etik sorunları beraberinde getirdiğini göstermektedir.

Büyük veri diğer yandan kullanıldıkça kendisini besleyen bir sistem hâline gelmiştir. Veri üretenler de veri ürettiklerinin bilincinde olmadan sürecin en önemli bileşeni görevini yerine getirmektedir. Google yöneticisinin kullanım trafiğinin artışıyla birlikte sistemin doğruluğunun artacağı (Gürsakar, 2014: 6) yönünde değerlendirmesi de bu gerçeğe işaret etmektedir. Özellikle de Kovid-19 salgını

sürecinde büyük bir dijital dönüşüm gerçekleşmekte, bu süreçte büyük veri "dijital diktatör" (Tarhan, 2020: 127) olarak görülmektedir. Nevzat Tarhan, büyük veriyi eline alanın dünyaya istediği gibi hükmetmeye başlayacağına işaret ederek bireylerin özel alanlarını koruması gerektiğini ifade etmektedir.

Kaynakça

- Al-Khouri, A. M. (2012). Data Ownership: Who Owns "My Data"? *International Journal of Management & Information Technology*, 2 (1), 1-8.
- Aristoteles. (1995). *Retorik*. (M. H. Doğan, Çev.) İstanbul: Yapı Kredi Yayınları.
- Arslantunali, M. (2019). *Teknopolis - Akıllı Makineler Dağılık Zihinler*. İstanbul: İletişim Yayınları.
- Arthur, C. (2017). *Dijital Savaşlar: Apple, Google, Microsoft ve İnternet Savaşı*. (A. Önsan, Çev.) İstanbul: Türkiye İş Bankası Kültür Yayınları.
- Baudrillard, J. (2004). *Tüketim Toplumu* (2. b.). (H. Deliceçaylı, & F. Keskin, Çev.) İstanbul: Ayrıntı Yayınları.
- Bilbilik, E. (2015). *Big Data Age*. İstanbul: Nergiz Yayınları.
- Boyd, D., & Crawford, K. (2012). Critical Questions For Big Data. *Information, Communication & Society*, 15(5), 662-679.
- Brynjolfsson, E., & McAfee, A. (2015). *The Second Machine Age*. (L. Göktem, Çev.) İstanbul: Türk Hava Yolları Yayınları.
- Canetti, E. (2017). *Saatın Gizli Yüreği* (2. b.). (A. Cemal, Çev.) İstanbul: Sel Yayınları.
- Castells, M. (2013). *İsyan ve Umut Ağları İnternet Çağında Toplumsal Hareketler*. (E. Kılıç, Çev.) İstanbul: Koç Üniversitesi Yayınları.
- Chen, M., Mao, S., & Liu, Y. (2014). Big Data: A Survey. *Mobile Netw Appl*(19), 171-209.
- Couldry, N., & Mejias, U. (2019). Data Colonialism: Rethinking Big Data's Relation to The Contemporary Subject. *Television and New Media*, 20(4), 336-349.
- Crary, J. (2015). *7/24 Geç Kapitalizm ve Uykuların Sonu*. (N. Çatlı, Çev.) İstanbul: Metis Yayınları.
- Davenport, T. (2014). *Big Data @ Work*. (M. Çavdar, Çev.) İstanbul: Türk Hava Yolları Yayınları.
- Davis, K., & Patterson, D. (2012). *Ethics of Big Data*. United States of America: O'Reilly.
- Dijcks, J.-P. (2013). *Oracle: Big Data for the Enterprise*. <http://www.oracle.com/us/products/database/big-data-for-enterprise-519135.pdf>
- Erdoğan, İ. (1997). *İletişim Egemenlik ve Mücadeleye Giriş*. Ankara: İmge Kitabevi.
- Erkmen, Ö. (2020). Türkiye Örneğinde Veri Gazeteciliği Uygulamaları ve Demokrasi İlişkisi Üzerine Bir Değerlendirme. *Connectist: Istanbul University Journal of Communication Sciences*(58), 65-103.
- Felle, T. (2016). Digital Watchdogs? Data Reporting and The Traditional 'Fourth Estate' Role For Journalists. *Journalism Theory, Practice and Criticism*, 17(1), 85-96.
- Ford, M. (2018). *Robotların Yükselişin* (5. b.). (C. Duran, Çev.) İstanbul: Kronik Kitap.
- Gantz, J., & Reinsel, D. (2011). *Extracting Value from Chaos*. emc.com. <http://www.emc.com/collateral/analyst-reports/idc-extracting-value-from-chaos-ar.pdf>
- Goodman, M. (2016). *Geleceğin Suçları*. İstanbul : Timaş Yayınları.

- Günek, A. (2015). Roma'dan Pentagon'a İletişimin Merkezi Rolü ve İktidar. VI.Türkiye Lisansüstü Çalışmalar Kongresi (s. 267-276). Kütahya: İلمي Etüdler Derneği.
- Gürsakal, N. (2014). Büyük Veri (3. b.). Bursa: Dora Yayınevi.
- Helland, P. (2011). If You Have Too Much Data, then "Good Enough" Is Good Enough. *Ac-mqueue*, 1-11. <https://queue.acm.org/detail.cfm?id=1988603> adresinden alındı.
- Hoeren, T. (2014). Big Data and The Ownership in Data: Recent Developments in Europe. *European Intellectual Property Review*, 36(12), 751-754.
- Holmes, D. E. (2017). *Big Data: A Very Short Introduction*. Oxford: Oxford University Press.
- Kubitschko, S. (2015). The Role of Hackers in Countering Surveillance and Promoting Democracy. *Media and Communication*, 3(2), 77-87.
- Leonhard, G. (2018). *Teknolojiye Karşı İnsanlık*. (C. Akkartal, & İ. Akkartal, Çev.) İstanbul: Siyah Kitap.
- Lokke, E. (2018). *Mahremiyet: Dijital Toplumda Özel Hayat*. (D. Başak, Çev.) İstanbul: Koç Üniversitesi Yayınları.
- Markoff, J. (2017). *Sevgi Dolu Makineler*. (T. Gezer, Çev.) İstanbul: Gedik Üniversitesi Yayınları.
- Mauro, A. D., Greco, M., & Grimaldi, M. (2015). What is Big Data? A Consensual Definition and a Review of Key Research Topics. *AIP Conference Proceedings* (s. 97-104). AIP Publishing.
- McAfee, A., & Brynjolfsson, E. (2018). *Makine, Platform, Kitle*. (T. Gezer, Çev.) İstanbul: Optimist Yayınları.
- Mul, J. D. (2008). *Siberuzayda Macera Dolu Bir Yolculuk*. (A. Özdamar, Çev.) İstanbul: Kitap Yayınevi.
- Narin, B., Ayaz, B., Fırat, F., & Fırat, D. (2017). Büyük Veri ve Gazetecilik İlişkisi Bağlamında Veri Gazeteciliği. *AJIT-E*, 8(30), 215-235.
- Orwell, G. (1960). 1984. (H. Derin, Çev.) Ankara: Maarif Basımevi.
- Öcal, D. (2019). *Teknolojik Yeniliklerin Yönetimi ve Tüketen Bireyin Dönüşümü*. Karadeniz İletişim Araştırmaları Dergisi, 97-115.
- Palmer, M. (2006). Data is The New Oil. [ana.blogs.com: https://ana.blogs.com/maestros/2006/11/data_is_the_new.html](https://ana.blogs.com/maestros/2006/11/data_is_the_new.html) adresinden alındı.
- Ritzer, G. (2011). *Büyüsü Bozulmuş Dünyayı Büyülemek* (2. Baskı b.). (Ş. S. Kaya, Çev.) İstanbul: Ayrıntı Yayınları.
- Saka, E., & Sayan, A. (2016). *Büyük Veri Nedir? Olanaklar ve Sınırlar*. N. Timisi içinde, *Dijital: Kavramlar, Olanaklar, Deneyimler* (s. 87-108). İstanbul: Kalkedon Yayınları.
- Stephens-Davidowitz, S. (2018). *Bana Yalan Söylediler*. (F. B. Aydar, Çev.) İstanbul: Koç Üniversitesi Yayınları.
- Şeker, S. (2018). *5G Nesnelerin İnterneti ve Sağlığımız*. İstanbul: Hayykitap.
- Tarhan, N. (2020). Geleceğin İhtiyaçlarına Odaklanmalıyız. 122-129. (T. Akademi, Röportaj Yapan) <https://dergipark.org.tr/tr/download/article-file/1312430> adresinden alındı.
- Tong, S. T., Heide, B. V., Langwell, L., & Walther, J. B. (2008). Too Much of a Good Thing? The Relationship Between Number of Friends and Interpersonal Impressions on Facebook. *Journal of Computer-Mediated Communication*, 13(3), 531-549.

- Uluk, M. (2018). Hakikat Sonrası Çağda Yeni Medya & Yalan Haber. Eskişehir: Dorlion Yayınları.
- Ünal, F. (2015). Büyük Veri ve Semantik. İstanbul: Abaküs Kitap.
- Zadrozny, P., & Kodali, R. (2013). Big Data Analytics Using Splunk. Berkeley: Apress.
- Zhang, D. (2018). Big Data Security and Privacy Protection. 8th International Conference on Management and Computer Science , (s. 275-278).
- Zwitter, A. (2014). Big Data ethics. Big Data & Society, 1-6.



Büyük verinin iş hayatındaki etkisi, Freepik