

## Sosyal Bilimlerde Büyük Veri Üzerine Yapılan Akademik Çalışmaların Analizi: Yök Tez Örneği (2010 - 2020)

### Analysis of Academic Studies on Big Data in Social Sciences: Higher Education Council Thesis Example (2010 - 2020)

Kenan BÖLÜKBAŞ<sup>1</sup>

#### Öz

Günümüzde, teknolojinin gelişmesiyle birlikte bilime entegre olan sistemlerin karmaşık, dijital ve çok yönlü bir yapıda olduğunu söylemek mümkündür. Bu sistemlerden biri de son dönemde daha da popüler hale gelen “büyük veri” (big data) kavramıdır. Özellikle sosyal medya kullanımının artmasıyla devasa bir bilgi havuzu ile karşımıza çıkmakta olan büyük veri kavramına, birçok kurum ve kuruluşun yanı sıra akademi dünyasının da ilgi gösterdiği söylenebilir. Bu çalışmada, büyük verinin akademik yazındaki durumu, ele alınış biçimleri, sosyal ve beşeri bilimler öncelenecek değerlendirilmiştir. Bu bağlamda sosyal bilimler alanında 2010 – 2020 yılları arasında yazılmış YÖK Tez doktora tezleri, örneklem olarak belirlenmiştir. Araştırmada, yöntem olarak nitel ve nicel içerik analizi kullanılmıştır. Elde edilen bulgular; yıllara, enstitülere ve anabilim dallarına göre tasnif edilmiş ve oransal dağılımlar sonucunda büyük veri ile ilgili tezler incelenmiştir. Sonuç olarak, literatürde büyük veri kavramının daha çok fen bilimleri alanında ele alındığı, sosyal bilimlerin ise ikinci planda kaldığı görülmüştür.

**Anahtar Kelimeler:** Büyük veri, Ulusal Tez Merkezi (YÖK), doktora tezleri.

#### Abstract

Nowadays, with the development of technology, we can say that systems integrated into science are complex, digital and versatile. One of these systems is the concept of “big data”, which has become more popular recently. It can be said that many institutions and organizations, as well as the academy world, have shown interest in the concept of big data, which is emerging with a huge pool of information, especially with the increase in the use of social media. In this study, the situation of big data in academic literature, the way it is handled, social sciences and humanities were evaluated with priority. In this context, the Higher Education Institution Thesis doctoral dissertations in the field of social sciences between 2010 and 2020 were determined as a sample. In the research, qualitative and quantitative content analysis was performed as a method. Findings obtained; They were classified according to years, institutes and departments and as a result of proportional distributions, the theses about big data were examined. As a result, it has been observed that the concept of big data is mostly discussed in the field of science, while social sciences remain in the background in the literature.

**Keywords:** Big data, National Thesis Center(Council of Higher Education), doctoral theses.

Araştırma Makalesi (Research Article)

Gönderim Tarihi (Received): 25.02.2021

Kabul Tarihi (Accepted): 31.05.2021

Atıf (cite as): Bölükbaş, K. (2021). “Sosyal Bilimlerde Büyük Veri Üzerine Yapılan Akademik Çalışmaların Analizi: YÖK Tez Örneği (2010 - 2020)”. Akdeniz Üniversitesi İletişim Fakültesi Dergisi, (AKİL) Haziran (35), s. 158-173, DOI: 10.31123/akil.886708

## Giriş

Teknolojinin ilerlemesiyle; sosyal, siyasal, kültürel, ekonomik, ticari ve daha pek çok alanda popüler ve yeni kavramlar karşımıza çıkmaya devam etmektedir. Bu durum, akademik alanyazın için de geçerlidir. Her geçen gün yeni kavramlar ve bileşenler disiplinlerarası bir değişim/dönüşüm ya da gelişim göstererek farklı bir araştırma alanı oluşturmaktadır. Sözel kavramlar birtakım olumlu ve/ veya olumsuz sonuçları - tartışmaları da beraberinde getirmektedir. Bu kavramlardan birisi de literatürde “dev veri” ya da “big data” şeklinde yer bulan ve karmaşık süreçleri beraberinde getiren “büyük veri” kavramıdır.

Bugün bilim ve teknolojiye yaşanan gelişmeler, internet teknolojilerindeki kullanımın yaygınlaşması ve internet erişiminin mobile kadar indirgenmesiyle birçok kaynaktan sürekli veri akışı mümkün hale gelmiştir. Toplanan bu enformasyon, birçok farklı kurum tarafından kurumsal politikalar ve çıkarlar doğrultusunda farklı alanlarda kullanılmaktadır. Dolayısıyla, bu veri yığınının işlenmesi ve anlamlandırılması büyük önem taşımaktadır. Günümüzde, kurum ve kuruluşlarının duyarlılığı tartışmalı olsa da özellikle bilim, bilişim ve teknoloji ile ilgili çevrelerin büyük veri üzerine çok daha fazla yoğunlaştıkları söylenebilir.

Büyük veri, bilgi – enformasyon ve veri arasındaki döngüde, bilginin oluşumuyla ilgili genel kabul gören hiyerarşik süreçleri içermektedir. Bu etkileşimli süreçlerin temelinde ise “insan” faktörü bulunmaktadır (Bayrakçı, 2015, s. 18 – 19). Dolayısıyla, sosyal bilimlerin bu enformasyon açısından ayrı düşünülmemesi gerekmektedir. Bu yüzden Işıklı’ya (2014) göre büyük verinin, yalnızca teknik bir yaklaşımla değil genel beşeri bilimlerin anlaşılması için de ele alınması gerekmektedir. Bu noktada büyük verinin sıradan kullanıcılar ile anlaşılan süreçler dizisi olduğunu söylemek mümkündür. Örneğin, alışveriş merkezinden giriş yaptığımız anda sensörler bizi kaydetmeye başlarken, toplu taşıma kullanırken okutulan manyetik kartlar ise ulaşım ile ilgili verilerimizi içermektedir. Aynı şekilde Twitter’da paylaştığımız siyasi görüş, takip ettiğimiz kişiler, beğeniler, Swarm aracılığıyla yaptığımız check-in, Foursquare aracılığıyla verdiğimiz tavsiye ile hakkımızda detaylı bir veri birikimi sağlanmaktadır. Bu durumda hem şirketler hem de sosyal medya platformları biriken verilerin depolandığı yerler olarak karşımıza çıkabilmektedir. Dolayısıyla tüketim ve kullanım alışkanlıklarımızla bıraktığımız sayısal yani dijital izlerle (Akcan, 2016, s. 312 - 328) yalnızca teknoloji alanında değil, düşünme ve algılama biçimlerinde, araştırma yöntemlerinde kurum ve kuruluşların yanı sıra bireylerin de dışında kalamayacağı bir sistemden bahsetmek mümkündür (Doğan ve Arslantekin, 2016, s. 15 - 34). Bu noktada başlıca büyük veri kaynaklarını; fotoğraf, ses, video belgeleri, sosyal medya platformları, internet istatistikleri, bloglar, mobil cihazlar, sensörler, GSM operatörleri, mobese kayıtları, e – ticaret uygulamaları, bankalar, akademik çalışmalar şeklinde sıralamak mümkündür. Aynı şekilde kamudaki veri kaynakları da bu doğrultuda şekillenmektedir. Seçimler, nüfus hareketlilikleri, vergi, seyahat, eğitim gibi aktiviteler de büyük veriye kaynak oluşturmaktadır (Aktan, 2018, s. 6 - 7; Altunışık, 2015, s. 55 - 57; Erkmen, 2018, s. 325; Işıklı, 2014, s. 93; Milli, Şentürk vd., 2016; Onay, 2020, s. 131; Özbilgin, 2015, s. 1 - 8; Uğur, 2020, s. 4 - 6).

Böylesi bir sistemin temelini oluşturan “veri” aynı zamanda büyük verinin de ana gövdesini oluşturmaktadır. Veri, bahsi geçen kaynaklardan ve birtakım araştırma ve gözlemsel süreçlerden elde edilen genel bir terimi ifade etmektedir (Düşünen Beyinler, 2020). Latince tekil “datum” sözcüğünün çoğulu olan veri (data); bir araştırmanın, bir tartışmanın, bir muhakemenin temeli olan ana öge olarak tanımlanmaktadır (Düşünen Beyinler, 2020; Türk Dil Kurumu [TDK], 2020).

Bir başka tanıma göre veri; kaydedilen, yeniden düzenlenebilen ve analiz edilebilen bilgi (information) birimidir (Işıklı, 2014, s. 92). Seker'e (2014, s. 22) göre veri henüz işlenmemiş kayıt ya da bilgi olarak ele alınmaktadır. Veri, "çağımızın en değerli hammaddesi"dir (Ege, 2013, s. 23 – 24). Dolayısıyla gelecekte enformasyona sahip olanlar ve bu enformasyonu iş süreçlerine dâhil edip daha iyi kullananlar rakiplerine oranla çok daha değerli bir kimlik kazanabilecektir (Doğan vd., 2016, s. 16 - 19).

Bu gerçek, bugün milyonlarca kullanıcısı olan Google, Twitter, Instagram ve Facebook gibi bilişim şirketlerince fark edilmiş durumdadır. Öyle ki "Visual Capitalist"ın düzenlediği raporda Domo'nun (2020) verilerine göre sosyal medya ile ilgili ilginç verilerle karşılaşılmaktadır. Buna göre internette bir dakikada:

- Facebook'a 147 bin fotoğraf yüklenmekte,
- YouTube'a 500 saatlik video yüklenmekte,
- Netflix'e 404 bin 444 saatlik video akışı olmakta,
- Instagram'da 347 bin 222 hikâye paylaşılmakta,
- Twitter'a 319 yeni kullanıcı katılmakta,
- Spotify uygulama kitaplığına 28 yeni parça eklenmekte,
- LinkedIn'de 69 bin 444 iş başvurusu yapılmakta,
- Amazon 6 bin 659 paket sevkiyatı yapmakta,
- Zoom uygulamasındaki toplantılara 208 bin 333 kişi katılmaktadır (Visual Capitalist, 2020; TRT Haber - Bilim & Teknoloji, 2020).

Bu verilerden hareketle internet kullanımının artmasıyla kullanıcıların kullanım biçimleri ve eğilimleri, içerik üretme şekilleri ve sayıları artmaktadır. Bu suretle ortaya çıkan veri yığını ya da veri havuzu ciddi bir potansiyel taşımaktadır. Ancak, söz konusu veri, çoğu zaman homojen bir şekilde dağılmamaktadır. Dolayısıyla bu heterojen yapıdaki verinin tasnif edilmesi yönünde bir gereklilik ortaya çıkmaktadır (Doğu, 2015, s. 181 - 183). Böylelikle büyük veri çalışmaları ile çok sayıda veri hızlı bir şekilde işlenip, analiz edilebilmektedir. Büyük veri analizlerinde verilerin toplanıp, dağınık haldeki bilgilerin düzenlenip/ seçilerek nitelikli bilgiye erişim sağlanmaktadır. Bu noktadan sonra ise geliştirilmiş algoritmalar yardımıyla hedef kitleye yönelik ölçümlenmelerle veriler anlamlandırılmakta ve bir "değer" kazanmaktadır. (Gülle, 2013, s. 1 - 2; Seker, 2014, s. 21 - 25).

Buradan hareketle büyük verinin elde edilen enformasyonun dağıtımı, yönetimi ve analizi olduğunu söylemek mümkündür (Mollaoğulları ve Özdoğan, 2018, s. 630 - 631). Büyük veri; geleneksel ve mevcut veri işleme uygulamalarının yetersiz kaldığı oldukça büyük ve karmaşık veri kümelerini tanımlamak için kullanılmaktadır (Çevik, Öztürkcan ve Kasap, 2015; Çifçi, 2017). Işıklı'ya (2014, s. 93 - 94) göre büyük veri, klasik veri tabanlarının kaldıramayacağı büyüklükteki verileri ifade etmektedir. Dolayısıyla teknoloji ve bilişim alanında zengin bir bilgi toplumu ait çıktılarının "büyük veri" olarak nitelendirilmesi mümkündür.

Büyük veri ile ilgili çalışmalar incelendiğinde, kavram üzerinde uzlaşılmış kesin bir tanıma rastlanılmamıştır. Bu nedenle literatürde birçok tanımlamaya rastlamak mümkündür. Seker'e (2015, s. 10 - 11)'e göre büyük veri, sürekli büyüyen ve bilgisayarların işleyemeyeceği kadar büyük olan verilerdir. Yaşantımızın her alanına nüfuz eden internetin yoğun kullanımı ile birikmeye başlayan

veriler günümüzde büyük veri olarak adlandırılmaktadır. Metinler, videolar ve görseller her gün sanal dünyaya yüklenmekte, dolayısıyla büyük veri her geçen daha da büyüyerek inanılmaz boyutlara ulaşmaktadır (Çelik, 2019). Dolayısıyla büyük veri, farklı kaynaklardan gelen farklı veri tiplerini içeren büyük boyutlu verilerdir (Milli, Şentürk vd., 2016). Günümüzde büyük veriler, teknolojik bir fenomen olarak karşımıza çıkmaktadır. Büyük veri sosyal medya mecralarında kullanılan emojilerden, sosyal medyanın etkileşimli yapısını oluşturan beğen tuşlarına; insansız hava araçlarının topladığı görüntülere kadar uzanan çeşitlilikteki dağınık ve karmaşık enformasyon birimlerinden oluşmaktadır (Narin, Ayaz vd. 2017, s. 217 - 219). En genel anlamıyla büyük veri, çok yüksek miktarlardaki verilerin çeşitli yöntemler ile gerçek zamanlı veya gecikmeli bir biçimde elde edilmesi, depolanması, saklanması, analiz edilmesi ve yararlı bilgilere dönüştürme süreci olarak tanımlanmaktadır (Kırcı Çevik ve Özdemir, 2018, s. 30 - 31).

Oldukça karmaşık fakat bir o kadar da sistematik olan büyük veri sistemi kısa zaman aralıklarında dahi kendini güncelleme gereksinimi duymaktadır. Her geçen gün gelişen teknoloji, bilişim ve yazılım sistemlerindeki değişimler, farklılaşan bireysel eğilimler neticesinde büyük verinin temelde sahip olduğu özelliklere yenilerinin eklendiği ve eklenmeye devam ettiği görülmektedir. Karakteristik özelliklerdeki gelişimi, şu şekilde incelemek mümkündür:

Temelde Meta Group isimli şirketin, 2001 yılında büyük veriyi hacim, hız ve çeşitlilik ile nitelediği bilinmektedir. Bu yüzden büyük veri kavramının tanımlanmasında genellikle verinin hacmi, hızı ve çeşitliliğini ifade eden “3V” (volume, velocity, variety) modülasyonu yaygın bir şekilde kullanılmaktadır (bkz. Tablo: 1).

Verinin hacmi, verinin büyüklüğünü, boyutunu, miktarını ya da kapladığı yeri ifade etmektedir. Büyük verinin büyüklüğü, hacminden çok verilerin işleme süreçleri ile ilgilidir (Doğan vd., 2016, s. 23). Bugün sosyal medya mecralarına her an çok yüksek miktarda veri akmaktadır. Bu paylaşılan bilgiler o kadar yoğun akmaktadır ki, bazı konulardaki bir saatlik paylaşımları okumak için dahi bir insan ömrü yetmemektedir. Örneğin, sadece 2012 yılında sosyal ağlarda ve webde üretilen yazılı içerik, yazının icadından sonra 6000 yıl boyunca yazılmış bütün içerikten daha fazladır (Seker, 2014, s. 21 - 25). Dolayısıyla büyük veri içindeki “büyük” adlandırması hacimsel vurgusundan ziyade nasıl anlamlandırıldığı ve işe koşulduğu ile önem kazanmaktadır (Akcan, 2016, s. 312 - 328).

Verinin hızı, verinin dinamikliği ile ilgilidir. Örneğin Twitter’da her gün çok sayıda hashtag paylaşılmaktadır. Oluşan bu veri havuzundaki verilerin işlenebilmesi için belirli bir işlem hızının yakalanması gerekmektedir. Verinin çeşitliliği ise değişik kaynaklardan değişik içeriklerden değişik özelliklere sahip verilerin gelmesidir. Verilerin çeşitliliği uydu verileri, metin verileri, akıllı telefon verileri, sosyal ağ verileri gibi kaynak temelli şekillenmektedir.

**Tablo 1.** Büyük Verinin “3V” Boyutu

<b>Hacim (Volume)</b> (Boyut/ büyüklük)	<b>Çeşitlilik (Variety)</b> (Değişik içerikler)
<b>Büyük Verinin Karakteristiği</b>	
<b>Hız (Velocity)</b> (Dinamiklik/ hareket)	

Büyük verinin karakteristik özellikleri günümüzde değişebilmekte ve kendini güncellemektedir. Literatürde 3V'ye "verinin değerini" (value) ekleyerek 4V ile tanımlayanlar da bulunmaktadır. Bu dördüncü boyut, verinin kalite unsurunu ifade etmektedir (bkz. Tablo: 2).

**Tablo 2.** Büyük Verinin "4V" Boyutu

Hacim (Volume)	Çeşitlilik (Variety)
<b>Büyük Verinin Karakteristiği</b>	
Hız (Velocity)	<b>Verinin Değeri (Value) (Kalite)</b>

Bir verinin büyüklüğüne, hızına ve çeşitliliğine göre bir değerden bahsetmek mümkündür. Bu noktada ilgili verinin güvenilirliği, kaynağa ne kadar itimat edileceği önem kazanmaktadır. Dolayısıyla verinin güvenilirliği (veracity) beşinci boyut olarak karşımıza çıkmaktadır (bkz. Tablo: 3).

**Tablo 3.** Büyük Verinin "5V" Boyutu

<b>Verinin Güvenilirliği (Veracity) (Veri kaynağına itimat)</b>	
Hacim (Volume)	Çeşitlilik (Variety)
<b>Büyük Verinin Karakteristiği</b>	
Hız (Velocity)	Verinin Değeri (Value)

Büyük veri kavramı ve büyük veri çalışmaları, veriyi daha değerli hale getirmek için ilk 4V problemlerine odaklanmakta ve çözmeye çalışmaktadır. Büyük veri çalışmalarının temel problemi, büyük ölçekteki verileri, hızlı değişen, çeşitli kaynaklardan gelen ve farklı güven değerlerine sahip olan verileri yönetebilmektir. Günümüzde gelişen teknoloji, yazılımlar ya da beklentiler sonucunda büyük veri boyutları kendini yenilemektedir. Bugün "5V" boyuta yeni olgular eklenmiştir (bkz. Tablo: 4). Bunlar veri doğrulama (verification) ve verinin değişkenliği (variability)'dir. Dolayısıyla bu noktada "7V"den söz etmek mümkündür (Aktan, 2018, s. 4 - 6; Altunışık, 2015, s. 48 - 55; Bayrakçı ve Albayrak, 2019, s. 75 - 78; Cackett 14 - 15'den aktaran Doğan vd., 2016, s. 23 - 24; Çifçi, 2017; Kırıcı Çevik vd., 2018, s. 31 - 33; Mollaoğulları vd., 2018, s. 630 - 631; Narin, Ayaz vd. 2017, s. 219; Onay, 2020, s. 132 - 136; Özbilgin, 2015, s. 1 - 8; Öztürk, 2017, s. 221; Seker, 2015, s. 10 - 11).

**Tablo 4.** Büyük Verinin "7V" Boyutu

<b>Verinin Güvenilirliği (Veracity)</b>	
Hacim (Volume)	Çeşitlilik (Variety)
<b>Büyük Verinin Karakteristiği</b>	
Hız (Velocity)	Verinin Değeri (Value)
<b>Verinin Geçerliliği (Validity)</b>	
<b>Verinin Değişkenliği (Variability)</b>	

Görüldüğü üzere karakteristik özellikleri gelişim/değişim ve yeniliklere açık olan büyük veri günümüzde “bilgi çöplüğünden hazine çıkarmaya” denk bir nitelik taşımaktadır (Çifçi, 2017; Vikipedi, 2020). Bunu bir kazanca dönüştürmek isteyen kurum ve kuruluşlara göre veri, “veri = para” denklemini sağlamaktadır. Dolayısıyla kurumların rekabet avantajı elde etmek için çağın petrolü veya para birimi olarak nitelenen büyük veriye vereceği değer ve öncelik önem kazanmaktadır (Ketizmen ve Kart, 2019, s. 66 - 68). Birçok farklı kaynaktan kartopu etkisiyle çoğalan bu bilgi (Gülle, 2013, s. 1 -2), yeni metot ve süreçlere dahil olarak karmaşık yapıların anlamlandırma sürecine katkı sağlayabilmektedir. Fakat bu noktada büyük ölçekli araştırma verilerinden kamu yararı oluşturma beklentilerinin yanı sıra yeni bir gizlilik ihlaline sebep olunabileceği üzerinde de durulmaktadır. Bu tartışma konusu etik unsurları sorgular hale getirmektedir. Bu noktada oluşabilecek veri sızmalarının, kişisel mahremiyeti olumsuz anlamda etkileyeceği açıktır (Boyd, Crawford ve Savcı, 2015, s. 199 – 211; Bozkanat ve Çömlekçi, 2020, s. 47 - 48).

Söz konusu etkilerde kurumsal etmenler etkili olabileceği gibi kişisel unsurlar da etkin rol alabilmektedir. Kişiler yankı fanusları ile ötekini ve/ veya farklı olanı reddetme eğilimine yönelmekte ve eko fanuslara çarpan bilgiler analiz edildiğinde siyasiler için politik, şirketler için ekonomik güç sağlanabilmektedir. Bu noktada “büyük veriyi” “büyük paraya” dönüştürmeye çalışanlara (Bozkanat vd., 2020, s. 55) özellikle sosyal medyada ve dijital ortamlarda bırakılan izler, hakkımızdaki “gerekli verileri” “ilgili birimlere” ulaştırmaktadır. Dijital izlerle oluşan algoritmalar doğrultusunda toplanan verilerin hacmi, gün geçtikçe artmakta ve bu durum büyük verinin özellikle analiz ve sonuçlar noktasında etik bir sorunu olarak karşımıza çıkmaktadır. Bu açıdan bakıldığında büyük veri, sadece matematik ve fen alanında değil sosyolojik ve kişisel bağlamda da oldukça önemli bir kavramdır. Bu aşamada veri analizleri ile bir anlamda “gözetim/ denetim toplumunun” temelini de atıldığından söz etmek mümkündür.

Böylesine çok yönlü bir kavramın literatür taraması yapıldığında çalışmaların genel olarak fen bilimleri alanında yoğunlaştığı görülmektedir. Bu araştırma, büyük verinin literatürde sosyal bilimler alanındaki genel kabullerini ve ele alınış biçimlerini, 2010 – 2020 yılları arasında Yükseköğretim Kurulu Başkanlığı Tez Merkezi’ndeki (YÖK Tez) doktora tezleri üzerinden değerlendirmektedir. Bu bağlamda YÖK Tez’deki tezlerin yani bir anlamda veri havuzunun analizi ile alanyazına katkı sunulacağı düşünülmektedir.

## Çalışmanın Amacı

Akademik kaynaklar, büyük veri ile ilgili çalışmaların toplanmasında başvurulacak kaynaklardan biridir. Bu çalışmada, 2010 – 2020 yılları arasında büyük veri kavramının sosyal bilimler alanındaki akademik çalışmalarda nasıl yer bulduğunun ve nasıl işlendiğinin ortaya konulması amaçlanmaktadır. Bu amaçla ilgili yıllarda, büyük veri ile ilgili YÖK Tez’in arama sayfasından erişilebilen doktora tezleri irdelenerek, alanyazındaki araştırmacıların büyük veri kavramını nasıl değerlendirdikleri ortaya konulacaktır.

Böylece “Sosyal Bilimlerde Büyük Veri Üzerine Yapılan Akademik Çalışmalar: YÖK Tez Örneği (2010 – 2020)” araştırmanın problem cümlesini oluşturmaktadır. Bu bağlamda araştırma sorusu; “Sosyal bilimler alanında ‘büyük veri’ kavramını konu edinen veya bu kavrama yer veren tezler bulunmakta mıdır?” şeklinde karşımıza çıkmaktadır. Benzer şekilde aşağıdaki sorulara da cevap aranmaktadır:

- 2010 – 2020 yılları arasında tez sayılarındaki oransal dağılım nasıl olmuştur?

- Enstitülere göre tez sayılarında nasıl bir oransal dağılım vardır?
- Sosyal bilimler enstitülerinin anabilim dallarına göre tez sayılarındaki dağılım nasıldır?
- Sosyal bilimler enstitülerinin yayımladığı tezlerin anahtar kelimelerinde, içindekiler bölümünde ve içeriğinde büyük veri kavramı yer almakta mıdır?
- Sosyal bilimler alanındaki araştırmalarda büyük veri ile ilgili en çok hangi kavramlar ve sorunsallar üzerinde durulmuştur?

Büyük veri günümüzde birçok bilginin kaynağını oluşturmaktadır. Dolayısıyla disiplinlerarası bir kavram olarak bilimsel çalışmalara ve araştırmalara dâhil olmaktadır (Bayrakçı vd., 2019, s. 74 – 75; Yılmazel, 2019: 227 - 228). Bu noktada büyük veri ile ilgili yapılan akademik çalışmalara dair 10 yıllık bir özetin sunulması oldukça önem taşımaktadır.

## Yöntem

Çalışmada 2010–2020 zaman aralığında sosyal bilimler alanında büyük veri ile ilgili yapılan akademik çalışmaların ortaya konulması amacıyla YÖK Tez'in gelişmiş tarama sayfasına 2010 – 2020 yıllarını kapsayacak şekilde “büyük veri” veya “big data” yazılarak doktora tez türü kapsamında tarama yapılmıştır. Tarama sonucunda verilerin karşılaştırılması amacıyla grup sekmesinde -fen, sosyal, tıp gibi- herhangi bir sınırlamaya gidilmemiştir. Tarama araştırmasının ardından “büyük veri” veya “big data” kelime grubu için toplamda 150 doktora tezine ulaşılmıştır. Bu tezlerin 37 tanesi doğrudan sosyal bilimler enstitüleri ile bağlantılıdır. Araştırmanın bu kısmında ise sosyal bilimler alanı ile ilgili olan tezlerin nitel ve nicel içerik çözümlemesi yapılmıştır. İçerik analizi, özellikle sosyal bilimlerde sık başvurulan araştırma yöntemlerinden biridir. Bir gözlemden daha fazla bir çözümleme yöntemi olarak değerlendirilebilecek olan içerik analizi, mevcut verileri belli bir standarda kavuşturan, bir biri ile mukayese eden ya da mevcut halinden farklı bir forma dönüştüren yöntemdir.

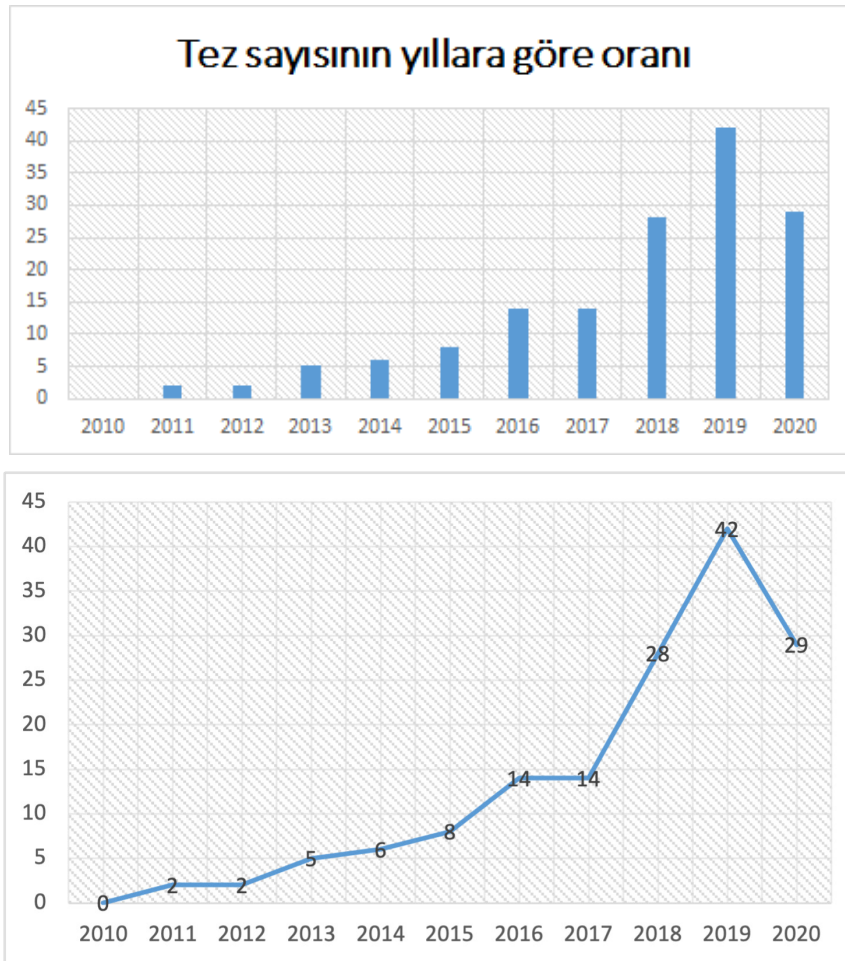
Bu yöntem araştırma bulgularının kodlama ve kategorileştirme yoluyla düzenli bir şekilde analiz edilmesine imkân tanımaktadır. Birçok araştırmada içerik analizi ile vurgulanan temel nokta “tarafsız ve sistematik” bulgular sunabilmesi olarak karşımıza çıkmaktadır (Koçak ve Arun, 2006, s. 22,23). Dolayısıyla çalışmada ölçümlenen/ ölçümlenemeyen veriler, nicel ve nitel içerik analizi ile anlamlandırılmaya çalışılmıştır.

## Bulgular

Gelişmiş tarama sonucunda, enstitü ve program kapsamında herhangi bir sınırlamaya gidilmeden ulaşılan 150 araştırma sonucunun yıllara göre dağılımı şu şekilde olmuştur:

**Tablo 5.** 2010 - 2020 yılları arasında büyük veri içerikli yayımlanan doktora tezi sayıları

Yıllar	Tez Sayıları
2010	0
2011	2
2012	2
2013	5
2014	6
2015	8
2016	14
2017	14
2018	28
2019	42
2020	29

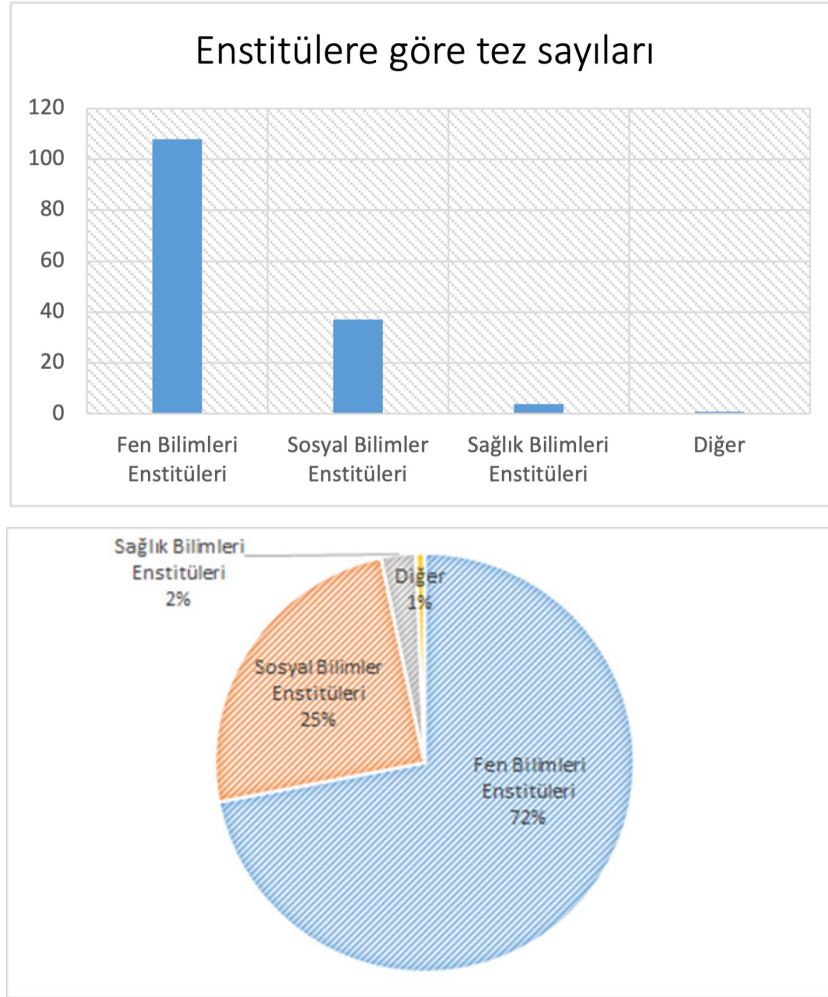
**Şekil 1.** Tablo 5'in çizgisel dağılımı

Tablo 5 ve şekil 1 incelendiğinde büyük veri ile ilgili akademik çalışmaların % 28'lik bir oranla 2019 yılında yoğunlaştığı görülmektedir. Bunu % 19 ile 2018 ve 2020, %10 ile 2016 ve 2017, % 5 ile 2015, % 4 ile 2014, % 3 ile 2013, % 1 ile 2011 ve 2012 yılları izlemektedir. 2010 yılında bu anlamda herhangi bir çalışmaya rastlanılmamıştır. Tez sayılarının yıllara göre yüzdesel dağılımından da anlaşılacağı üzere bu alanda yapılan akademik çalışmalar 2010 yılından 2019 yılına doğru periyodik bir artış göstermektedir. 2020 yılında ise 2019 yılına göre % 9'luk bir azalma söz konusudur.



**Tablo 6.** Enstitülere göre tezlerin toplam sayısı (2010 – 2020)

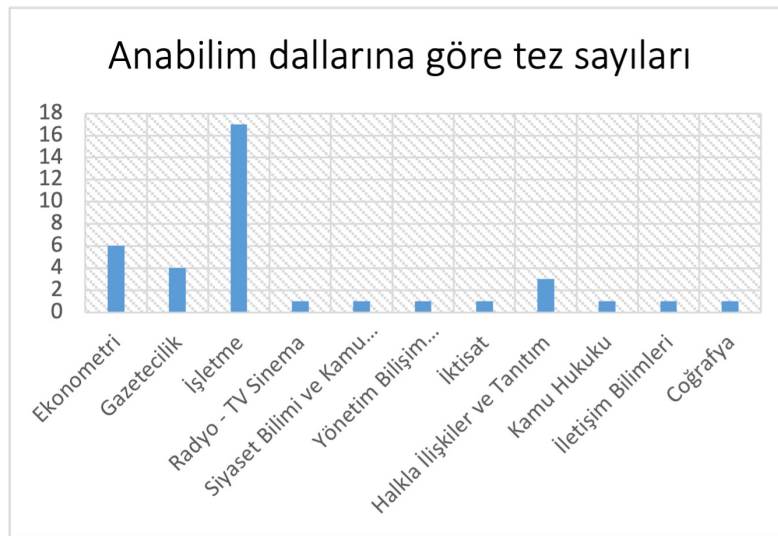
Enstitüler	Tez Sayıları
Fen Bilimleri Enstitüleri	108
Sosyal Bilimler Enstitüleri	37
Sağlık Bilimleri Enstitüleri	4
Diğer	1

**Şekil 2.** Tablo 6'nın çizgisel dağılımı ve yüzdesel dağılımı

Tablo 6 ve şekil 2 incelendiğinde tezlerin % 72'lik bir oranla “fen bilimleri enstitüleri” alanında yoğunlaştığı görülmektedir. Bunu % 25'lik bir oranla “sosyal bilimler enstitüleri, % 2 ile “sağlık bilimleri enstitüleri” ve % 1 ile “diğer enstitüler” izlemektedir. Buna göre fen bilimleri enstitülerinin büyük veri alanındaki akademik çalışmalarda diğer enstitülere göre daha ön planda olduğunu söylemek mümkündür. Bu noktada ulaşılan 150 tezin içinde sosyal bilimler enstitülerinin diğer enstitülere oranı ve tez sayısı oldukça önem kazanmaktadır. Buna göre 150 tezin 37 tanesine % 25'lik bir oranla tekabül eden sosyal bilimler enstitülerinin 2010 - 2020 yılları arasında yaptığı büyük veri ile ilgili çalışmaların anabilim dallarına göre sayısı ve oranları şöyledir:

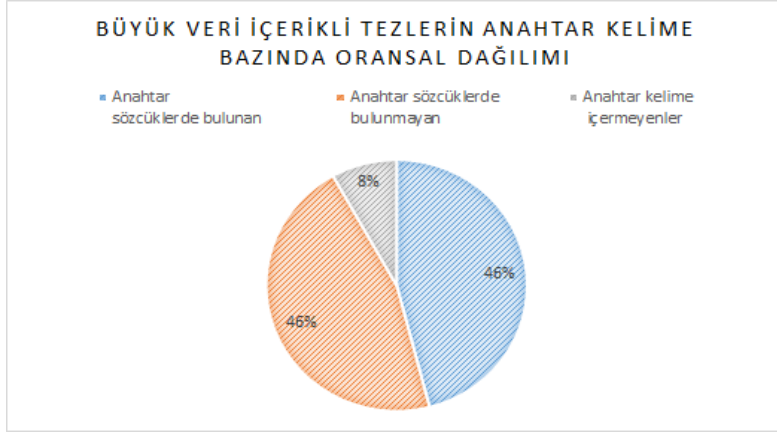
**Tablo 7.** Sosyal bilimler enstitülerinde büyük veri ile ilgili yapılan akademik çalışmaların anabilim dallarına göre sayısı

Sosyal Bilimler Enstitülerindeki Anabilim Dalları	Tez Sayıları
Ekonometri	6
Gazetecilik	4
İşletme	17
Radyo - TV Sinema	1
Siyaset Bilimi ve Kamu Yönetimi	1
Yönetim Bilişim Sistemleri	1
İktisat	1
Halkla İlişkiler ve Tanıtım	3
Kamu Hukuku	1
İletişim Bilimleri	1
Coğrafya	1

**Şekil 3.** Tablo 7'nin çizgisel dağılımı

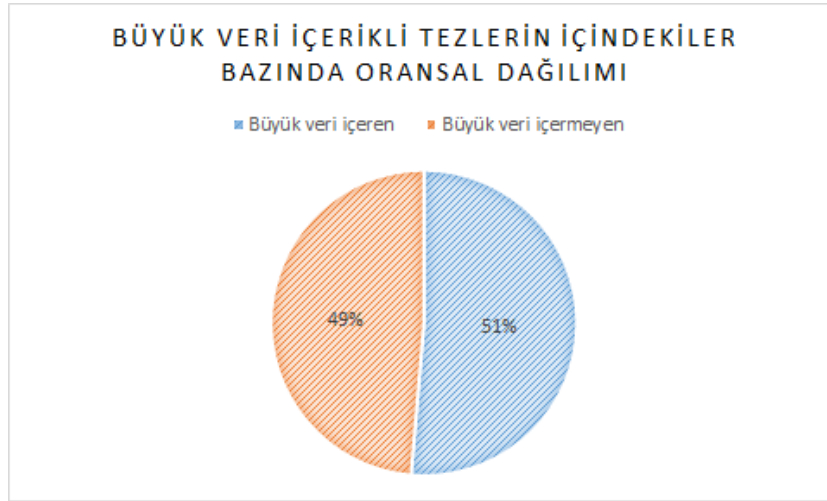
Tablo 7 ve şekil 3 incelendiğinde sosyal bilimler enstitülerinde büyük veri ile ilgili doktora tezlerinin % 46'lık bir oranla "işletme" alanında öne çıktığı görülmektedir. Bunu % 16'lık bir oranla ekonometri, % 11'lik bir oranla gazetecilik, % 8'lik bir oranla halkla ilişkiler ve tanıtım, % 3 ve % 2'lik oranlarla diğer anabilim dalları izlemektedir. Şekil 3'ün karşılıklı geldiği yüzdesel dağılımda 2010 – 2020 yılları arasında sosyal bilimler enstitülerinde yayımlanan doktora tezlerinde işletme anabilim dalının öne çıktığı görülmektedir. Bunda kurumsal markaların giderek artan rekabet ortamında daha çok kazanç sağlamak amacıyla büyük veri datalarından yararlanma eğilimi göstermelerinin etkili olduğu söylenebilir.

Çalışmanın ana konusunu oluşturan sosyal bilimler enstitülerinin, 2010 – 2020 periyodunda büyük veri bağlantılı çalışmalarını irdelemek bu noktada önem kazanmaktadır.



**Şekil 4.** Büyük veri içerikli tezlerin anahtar kelimeler bazında oransal dağılımı

Şekil 4’de sosyal bilimler enstitüleri tarafından 2010 – 2020 yılları arasında büyük veri ile ilgili yayımlanan doktora tezlerinin anahtar kelimelerine göre oransal bir dağılım oluşturulmuştur. Buna göre incelenen tez içeriklerinde, “büyük veri” kelime grubunu içeren anahtar sözcüklerin oranı ile büyük veri kelime grubunu içermeyen çalışmaların oranı % 46’lık bir oranla eşit bir dağılım göstermiştir. Erişim kısıtlaması bulunan ya da özetlerinde anahtar kelime bulunmayan tezlerin oranı ise % 8 olarak belirlenmiştir.



**Şekil 5.** Büyük veri içerikli tezlerin içindekiler bazında oransal dağılımı

Şekil 5’e göre tezlerin içindekiler bölümünde ise “büyük veri”ye dair bir içerik, bölüm ya da başlık içeren tezlerin oranı ise % 51 olarak belirlenmiştir. Bu noktada bazı tezlerin anahtar kelimelerinde “büyük veri” kelime grubuna yer verilse dahi içindekiler bölümünde bu kavrama rastlanılmadığı görülmüştür. Tezlerin içindekiler bölümünde büyük veriye başlık olarak yer vermeyen çalışmaların oranı ise % 49 olarak belirlenmiştir. Bu noktada oranların birbirine yakın seyrettiğini söylemek mümkündür.

Sosyal bilimler alanında bu kavrama yönelik yazılan tezlerin genel olarak veri setleri, yazılımlar, algoritmalar ve çeşitli programlar ile özellikle işletmelerin kamuya açık verileri kullanılmakta ve analiz edilmektedir. Yani büyük veri, amaca ulaşmak için bir araç olarak yer almaktadır. Doğrudan büyük veri ile ilgili sosyal bilimler kapsamında bir araştırma tespit edilememiştir.

## Sonuç

Büyük veri, günümüzün en popüler kavramlarından biridir. Birçok boyutuyla gelişmekte olan, kendini yenileyen disiplinlerarası bir süreçtir. Özellikle teknoloji ve bilişim temelli analizlerde oldukça karmaşık verilerin anlamlandırılması ve kullanılabilir hale getirilmesi noktasında önem kazanmaktadır. Bu doğrultuda; yazılımlar, algoritmalar ve programlar geliştirilmektedir. Dolayısıyla teknolojinin gelişmesi ile büyümekte ve genişlemekte olan internet teknolojileri, birtakım dönüşümleri de beraberinde getirmektedir. Bu dönüşümler, özelde “bireysel” ve “toplumsal”, genelde ise “fen bilimleri” ve “sosyal bilimler” alanında olabilmektedir. Bu noktada, etki alanındaki artışın temelinde bulunan büyük veri, analiz süreçlerinde karmaşık yapıları anlamlı hale getirmekte ve bu sayede kolay anlaşılabilir bir nitelik kazandırmaktadır. Özellikle sosyal medyanın mobile kadar indirgenmesi ile sosyal medyanın kullanımı arttıkça üretilen içeriklerin “veri” niteliği kazandığı görülmektedir. Dolayısıyla verilerin değerinin, miktarının ve çeşidinin de arttığını söylemek mümkündür. Bu noktada, doğrudan ve / veya dolaylı olarak insanı ve onun yaşam alanını etkiler hale gelen büyük verinin “sosyal bilimler” alanında irdelenmesi zorunlu hale gelmiştir. Bu çalışma, büyük verinin fen bilimlerindeki araştırma alanlarının yanı sıra sosyal bilimlerde ne derece yer aldığını ortaya koymayı amaçlamıştır.

Bu temel sorunsaldan hareketle akademide önemli bir yere sahip olan YÖK Tez’in gelişmiş tarama kısmında “büyük veri” veya “big data” kavramları taratılarak, bir dizin oluşturulmuştur. Çalışmanın periyodu, 2010 – 2020 aralığını kapsamaktadır. Çalışmada yöntem olarak, genel bir tarama modelinin ardından nitel ve nicel içerik analizi kullanılmıştır. Bu bağlamda, toplamda 150 doktora tezine ulaşılmıştır. Tez sayılarının diğer yıllara oranla 2019 yılında yoğunlaştığı görülmüştür. 2010 yılından 2019 yılına doğru tez sayılarında istikrarlı bir artış dikkati çekmektedir. Bu durum, 2020 yılında % 9’luk bir azalma söz konusu olsa da- büyük verinin gün geçtikçe popülerleşen ve kullanımı giderek artan bir kavram olduğunu göstermektedir. Tez sayılarına enstitü bazında bakıldığında ise fen bilimlerinin ilk sırada yer aldığı görülmüştür. Sosyal bilimler enstitülerinin bu alandaki çalışmaları %25’lik bir paya sahiptir. Fen bilimlerinin % 72’lik paya sahip olduğu düşünülürse, bu oranın oldukça düşük olduğunu söylemek mümkündür.

Sosyal bilimler enstitülerinin çalışmalarını kendi içinde değerlendirdiğimizde ise anabilim dallarında, işletme anabilim dalının öne çıktığı anlaşılmaktadır. Bunu sırasıyla ekonometri, gazetecilik, halkla ilişkiler ve tanıtım ve diğer anabilim dalları izlemektedir. Bu noktada, işletmelerin büyük veri analizleri ile kurumsal rekabet sağlama düşüncesinin akademik çalışmalarda yer bulduğunu söylemek mümkündür. İşletme anabilim dalına ait % 46’lık akademik çalışma oranı, bu bilginin kanıtı niteliğindedir.

Sosyal bilimler alanındaki 37 doktora tezinin anahtar kelimelerine ve içindekiler bölümlerine göre yapılan içerik analizlerinde ise büyük verinin genel anlamda bir “araç” olarak kullanıldığı göze çarpmaktadır. Kimi zaman Twitter kullanıcı verileri, kimi zaman bir işletmenin online verileri, kimi zaman da gazetecilikte yeni bir alan olarak ortaya çıkan “veri gazeteciliği”, büyük veriyi bir bilgi havuzu olarak yöntem ve araştırmalarda kullanmıştır. Bunun yanı sıra büyük veriyi kuramsal bazda ve gelecek tahmini çerçevesinde inceleyen çalışmalar da mevcuttur. Ayrıca kimi çalışmalar, anahtar kelimelerinde bu kavrama yer verse de içerik bazında ayrı bir bölüm olarak değinmemiştir.

Dijital çağa girdiğimiz bu dönemde, analiz yöntemleri kolaylaşmış, geçmişe göre çok daha fazla veriye ulaşılması mümkün hale gelmiştir. Bu imkânlarla birlikte; büyük verinin anlaşılması, analizi,

büyük veriden anlamlı ve faydalı çıkarımlar yapılabilmesi için uluslararası nitelikte araştırma merkezlerinin kurulması, özellikle sosyal bilimler alanında bilgisayar tabanlı becerilere ve yetkinliğe sahip veri bilimcilerin yetiştirilmesi yararlı olabilir. Çünkü günümüzde toplumsal alanları anlamaya çalışan sosyal bilimciler için büyük veri, önemli bilgi kaynaklardan biri olmuştur. Büyük veriyi sadece teknoloji ve bilişim kaynaklı bir terim olarak algılayan çalışmaların yanı sıra “birey” ve “toplum” temelli yönlerini de açığa çıkaracak akademik yazınların varlığına ihtiyaç duyulmaktadır. Bu noktada, büyük verinin sosyal bilimler enstitüleri tarafından ana kavram olarak değerlendirilmesinin düşünülmesi yararlı olabilir.

## Kaynakça

- Akcan, B. (2016). Big Data: Üreticiler ve Tüketiciler. *Yeni Medya Araştırmaları - 2*. [https://www.researchgate.net/publication/313402228\\_BIG\\_DATA\\_URETICILER\\_VE\\_TUKETICILER](https://www.researchgate.net/publication/313402228_BIG_DATA_URETICILER_VE_TUKETICILER).
- Aktan, E. (2018). Büyük Veri: Uygulama Alanları, Analitiği ve Güvenlik Boyutu. *Bilgi Yönetimi Dergisi*. 1(1), 1-22. <https://doi.org/10.33721/by.403010>.
- Altunışık, R. (2015). Büyük Veri: Fırsatlar Kaynağı mı Yoksa Yeni Sorunlar Yumağı mı? *Yıldız Social Science Review (YSSR)*. 1(1), 45-76. <https://dergipark.org.tr/tr/pub/yssr/issue/21899/235390>.
- Bayrakçı, S. (2015). Sosyal Bilimlerdeki Akademik Çalışmalarda Büyük Veri Kullanımı (Yayımlanmış Yüksek Lisans Tezi). Marmara Üniversitesi, Sosyal Bilimler Enstitüsü, İstanbul.
- Bayrakçı, S. ve Albayrak, M. A. (2019). Büyük Verinin Akademik Çalışmalarda Kullanımı Üzerine Mukayeseli Bir Veri Tabanı Araştırması. *AJIT-e: Online Academic Journal of Information Technology*. 10(36), 73-94. DOI: 10.5824/1309-1581.2019.1.004.x.
- Boyd, D., Crawford, K. ve Savcı, B. (2015). Büyük Veri Üzerine Eleştirel Sorular: Kültürel, Teknolojik ve Bilimsel Bir Olgu Hakkında Eleştirel Sorgulamalar. *Folklor/Edebiyat*. 21(83), 199-215. <https://dergipark.org.tr/tr/pub/fe/issue/26049/274368>.
- Bozkanat, E. ve Çömlekçi, M. F. (2020). Yeni Medya Çalışmalarının Geleceği ve Büyük Veri Kaynaklı Riskler: Bir Delfi Çalışması. *Connectist: Istanbul University Journal of Communication Sciences*. 58, 37-64. <https://doi.org/10.26650/CONNECTIST2020-0003>.
- Çelik, T. (2019, 4-6 Nisan). Dijital Çağda Gazetecilerin Büyük Veri Kullanımı. *Uluslararası Uygulamalı Sosyal Bilimler Kongresi'nde sunulan bildiri (s. 247-257)*. C-iasoS 2019 – İzmir –Turkey.
- Çevik, M., Öztürkcan, Selcen & Kasap, N. (2015, 10-13 Haziran). Sosyal Medya Analitiği: Twitter için Büyük Veri Yaklaşımı. *20.Ulusal Pazarlama Kongresi'nde sunulan bildiri*. Eskişehir, Turkey.
- Doğan, K. ve Arslantekin, S. (2016). Büyük Veri: Önemi, Yapısı ve Günümüzdeki Durumu. *DTCF Dergisi*. 56.(1), 15-36. <http://dtcdfdergisi.ankara.edu.tr/index.php/dtcf/article/view/4/1357>.
- Doğu, B. (2015). Veri Haberciliği: Demokratik Medya İçin Olanaklar. *Folklor/Edebiyat*. 21(83), 181-197. <https://dergipark.org.tr/tr/pub/fe/issue/26049/274357>.
- Ege, B. (2013). Rastlantının Bittiği Yer Big Data. *Bilim ve Teknik*. 550, 22-26. [http://bortecin.com/rastlantinin\\_bittigi\\_yer\\_big\\_data.pdf](http://bortecin.com/rastlantinin_bittigi_yer_big_data.pdf).
- Erkmen, Ö. (2018). Büyük Veri ve Gazetecilik: Veri Gazeteciliği Demokrasi, Katılım ve Gazeteciliğe Dair Anlayışımızı Nasıl Dönüştürebilir? *Akdeniz Üniversitesi İletişim Fakültesi Dergisi (AKİL)*. (30), s. 322-344. DOI: 10.31123/akil.464511.

- Gülle, M. T. (2013). Büyük Veri ya da İçgörü. *Türk Kütüphaneciliği*. 27(4), 581-582. <https://app.trdizin.gov.tr/makale/TVRjME56azBOQT09/buyuk-veri-ya-da-icgoru>.
- Işıkılı, Ş. (2014). Büyük Veri, Epistemoloji ve Etik Tartışmalar. *AJIT-e: Bilişim Teknolojileri Online Dergisi*. 5(17), 89-122. <https://dergipark.org.tr/tr/pub/ajit-e/issue/54437/740935>.
- Ketizmen, M., Kart, A. (2019). Kişisel Veri ve Rekabet Hukuku Kapsamında 'Big Data.' *Kişisel Verileri Koruma Dergisi*. 1(1), 64-76.
- Kırcı Çevik, N. ve Özdemir, M. (2018). Büyük Veri: Tanımı ve Oluşumu. *ICOMEF'18-Autumn, International Congress of Management, Economy and Policy 2018 Autumn Istanbul/Türkiye*. 1-2. [https://atif.sobiad.com/index.jsp?modul=makale-goruntule&id=AW4Raf\\_4yZgeuuwfeB3b](https://atif.sobiad.com/index.jsp?modul=makale-goruntule&id=AW4Raf_4yZgeuuwfeB3b).
- Koçak, A. ve Arun, Ö. (2006). İçerik Analizi Çalışmalarında Örneklem Sorunu. *Selçuk İletişim*. 4(3), 21-28.
- Milli, M., Şentürk, F., vd. (2016). Büyük Veri Kavramı ve Karakteristik Özellikleri. 183 – 188. [https://www.researchgate.net/publication/338623361\\_Buyuk\\_Veri\\_Kavrami\\_ve\\_Karakteristik\\_Ozellikleri](https://www.researchgate.net/publication/338623361_Buyuk_Veri_Kavrami_ve_Karakteristik_Ozellikleri).
- Mollaoğulları, B. F. ve Özdoğan, B. (2018). İletişim Teknolojilerindeki Gelişmeler, Riskler ve İç Denetimin Rolü. *Yönetim ve Ekonomi Celal Bayar Üniversitesi İktisadi ve İdari Bilimler Fakültesi Dergisi*. 25(3), 625 – 639. DOI: 10.18657/yon-veek.477585.
- Narin, B., Ayaz, B., Fırat, F. ve Fırat, D. (2017). Büyük Veri ve Gazetecilik İlişkisi Bağlamında Veri Gazeteciliği. *AJIT-e: Online Academic Journal of Information Technology*. 8(30), 215-235. DOI: 10.5824/1309-1581.2017.5.010.
- Onay, A. (2020). Büyük Veri Çağında İç Denetimin Dönüşümü. *Muhasebe Bilim Dünyası Dergisi*. 22(1), 127-163. DOI: 10.31460/mbdd.620837.
- Özbilgin, İ, G. (2015, 4-6 Şubat). Kamuda Büyük Veri ve Uygulamaları. *AB2015 Akademik Bilişim Konferansı*. <https://ab.org.tr/ab15/bildiri/483.pdf>.
- Öztürk, A. (2017). Yükseköğretimde Büyük Veri ve Öğrenme Analitikleri: Güncel Teori ve Uygulamalar. *Açıköğretim Uygulamaları ve Araştırmaları Dergisi*. 3(1), 218-228. <https://dergipark.org.tr/tr/pub/auad/issue/34114/377410>.
- Seker, S. E. (2014). Sosyal Ağlarda Akan Veri Madenciliği. *YBS Ansiklopedi*, 1(3), 21 - 27.
- Seker, S. E. (2015). Büyük Veri ve Büyük Veri Yaşam Döngüleri. *YBS Ansiklopedi*. 2(3).
- Uğur, S. (2020). Büyük Veri, Nesnelerin İnterneti, Sosyal Ağlar, Sanal Kimlikler ve Siber Güvenlik. *Açıköğretim Uygulamaları ve Araştırmaları Dergisi*. 6(1), 4-7.
- Yılmazel, Ö. (2019). Yök Ulusal Tez Merkezi'nde Büyük Veri Alanında Kayıtlı Bulunan Lisansüstü Tezlerin Analizi. *Karadeniz Uluslararası Bilimsel Dergi*. 41, 225-240. ISSN: 1308-6200. DOI Number: 10.17498/kdeniz.

### İnternet Kaynakları

- Çifçi, M. A. (2017). [https://www.researchgate.net/publication/329026562\\_BigData-sunusu](https://www.researchgate.net/publication/329026562_BigData-sunusu)DOI: 10.13140 / RG.2.2.12182.80969. adresinden 24.10.2020 tarihinde erişilmiştir.
- Düşünen Beyinler. (2020). Big Data (Büyük Veri) Nedir? <https://medium.com/dusunenbeyinler/big-data-b%C3%BCy%C3%BCk-veri-analizi-d53d8f8ab52b#:~:text=Veri%2C%20ara%C5%9Ft%C4%B1rmalardan%2C%20g%C3%B6zlemlerden%2C%20internetten,genel%20bir%20terimi%20ifade%20etmektedir>. adresinden 24.10.2020 tarihinde erişilmiştir.
- TRT Haber - Bilim& Teknoloji. (2020). İnternette 1 Dakikada Neler Oluyor? <https://www.trthaber.com/haber/bilim-teknoloji/internette-1-dakikada-neler-oluyor-517740.html> adresinden 23.10.2020 tarihinde erişilmiştir.

Türk Dil Kurumu (TDK), (2020). Veri Nedir?. <https://sozluk.gov.tr/> adresinden 24.10.2020 tarihinde erişilmiştir.

Vikipedi, (2020). Büyük Veri. [https://tr.wikipedia.org/wiki/B%C3%BCy%C3%BCk\\_veri](https://tr.wikipedia.org/wiki/B%C3%BCy%C3%BCk_veri) adresinden 23.10.2020 tarihinde erişilmiştir.

Visual Capitalist, (2020). İşte 2020'de İnternette Her Dakikada Ne Olur?. <https://www.visualcapitalist.com/every-minute-internet-2020/> adresinden 23.10.2020 tarihinde erişilmiştir.

## Extended Abstract

### Purpose of Research

In this study, how to interpret the concept of big data, which is more common in the field of science (such as social media, mobilization) within a structure in which technology and Digital are more involved in our lives, is the main problem in terms of social disciplines. Based on this, it is aimed to question the degree to which big data, which interacts with people, is increasingly diverse and increasingly qualified, is involved in the field of Social Sciences. This query is evaluated from the point of view of related disciplines, taking the social and humanities at the forefront.

### Research Questions

The main question of the research: "Are there dissertations in the field of social sciences that subject the concept of 'big data' or include this concept?" it appears in the form. In addition, the proportional distribution of theses by institutes and departments is also among the research questions.

### Literature Review

Systems that become integrated with the development of technology emerge in the form of a directory of digital and ongoing processes. One of these processes is a popular concept that finds a place in the literature as "big data." Big data is expressed as mass or mass information waiting to be analyzed and made sense. This information has been a great store of information in recent times, especially under the influence of digitalization and social media. The big data store is used by institutions and organizations to help businesses generate more profits and to determine their management and IT strategies. Therefore, algorithmic analysis of big data, which we hear more about in information and technological disciplines, is developed and analyzed. Thus complex heaps gain meaning and quality.

### Methodology

For this purpose, the doctoral theses in the dissertation center of the Council of Higher Education (YÖK dissertation) were determined as a sample using the academic summer. In this direction, the concepts of "big data" or "big data" were scanned using the advanced scanning feature of the YÖK thesis. The 2010-2020 year period was the period of the study. The method of the study was determined as qualitative and quantitative content analysis. Accordingly, the results obtained as a result of the analysis were classified around a number of titles (such as Year, Institute, Department). The data obtained as a result of the classification were interpreted according to mathematical and statistical evaluations.

## Results and Conclusion

As a result of the scanning, a total of 150 doctoral dissertations were reached, regardless of field differences. Accordingly, despite a 9% decrease in 2020, it was observed that the number of doctoral dissertations with big data content increased steadily from 2010 to 2019. When the theses are examined on the basis of the institute, it has been seen that disciplines belonging to the science have come to the fore. It was concluded that Social Sciences found an area of 25% compared to science, and this rate was quite low compared to science. Studies belonging to institutes of social sciences, when evaluated in itself, the Department of Business Administration stood out with an academic study rate of 46%. In this context, it has been seen that the strategies of corporate companies to create competition with big data are also evaluated in the academic index. When examining the key words and contents of dissertations in the field of social sciences, it was also noted that big data is considered a “tool”. As a result, it has been seen in the academic literature that big data is mainly considered in the disciplines of science, while social sciences remain in the background.

Big data should be included not only as a “tool” but also as a “main concept” in social science processes. Therefore, the presence of competent data scientists in the field of Social Sciences and qualified research companies can benefit both in order to obtain healthy findings from analysis processes of big data, which is increasing in popularity, diversity and use with digitalization.