

Derleme Makale

**Büyük Veri Çağında Sosyal Medya Verilerinin Sosyal Bilimler İçin Önemi;
Twitter Örneği**

*Importance of Social Media Data for Social Sciences in Big Data Age; The Case of
Twitter*

Hüseyin Tuğrul Oktay

Ankara Hacı Bayram Veli Üniversitesi Doktora Öğrencisi

htoktaya@gmail.com

<https://orcid.org/0000-0002-3656-7733>

Makale Gönderme Tarihi	Revizyon Tarihi	Kabul Tarihi
15.01.2020	12.03.2020	25.05.2020

Öz

Günlük yaşamda sıklıkla kullanılan sosyal medya araçları geleneksel olarak kurulan bağlantıların ötesinde sosyal ilişkiler kurulan ortamlar hâline gelmiştir. Bu ortamlarda ağın neresinde olduğunu bilemese de katılımcılar bir sosyal ağ kurmakta, bazen güçlü bazen de izole bir katılımcı olarak ağın bir parçası olmaktadır. Sosyal medyada çok farklı gruplar hemen hemen her gün ortaya çıkan kamusal tartışmalarda kendi görüşlerini ifade etmektedir. Özellikle Twitter mikroblog sitesi, karşıt grupların hemen her gün yeni bir konu başlığı etrafında karşı karşıya geldiği ve diyaloga girdikleri bir platformdur. Bu yeni durum artan teknik imkân ve becerilerle ele alınınca sosyal bilimciler için de devasa bir veri denizi ortaya çıkarmıştır. Üretilen veri miktarı her geçen gün katlanarak artmaktadır ve eskiden veri peşinde koşup bulmaya çalışılırken artık bireysel imkânlarla hesaplanamayacak kadar çok verinin önümüzden akışına şahit olunmaktadır. Bu çalışmada Twitter'in Dijital Sosyoloji bağlamında araştırmacılara sağladığı imkânlar dijital yöntemlerin ve Büyük Veri'nin imkân ve sınırlılıkları çerçevesinde tartışılmaktadır.

Anahtar Kelimeler: Büyük Veri, Dijital sosyoloji, Sosyal medya, Twitter

Abstract

Social media tools, which are frequently used in daily life, have become socially established environments beyond traditionally established connections. Although they do not know where the network is located in these environments, the participants form a social network, sometimes as a strong and sometimes isolated participant. In social media, very different groups express their views in public debates that occur almost every day. In particular, the Twitter microblogging site is a platform where opposing groups come across a new topic almost every day and engage in dialogue. Given this new situation with increasing technical capabilities and skills, it has created a huge sea of data for social scientists. The amount of data produced is increasing exponentially day by day, and in the past, we are witnessing the flow of data that cannot be calculated by individual means. In this study, the opportunities provided by Twitter in the context of Digital Sociology are discussed within the framework of the possibilities and limitations of digital methods and Big Data.

Keywords: Big Data, Digital sociology, Social media, Twitter

1. Giriş

7,3 milyarlık dünyada 3 milyar aktif sosyal medya kullanıcısı vardır ve bu sayı dünyanın en kalabalık ülkesi Çin'den daha fazla bir nüfustur. 2019 yılı istatistiklerine göre de Twitter mikro

Önerilen Atıf/Suggested Citation

Oktay, H. T. 2020. Büyük Veri Çağında Sosyal Medya Verilerinin Sosyal Bilimler İçin Önemi; Twitter Örneği, *Üçüncü Sektör Sosyal Ekonomi Dergisi*, 55(2), 1090-1110

blog sitesinin kullanıcı sayısı 330 milyondur (Clement, 2019). Yani Twitter kullanıcıları dünyada bir ülke olsa dünyadaki en kalabalık 3. ülke olup Çin ve Hindistan'dan sonra yer alacaktı.

Elbette ki bu sayıların fazlalığı sosyolojik olarak tek başına bir şey ifade etmemektedir. Sonuçta ulus devletlerin nüfuslarıyla parça parça sohbet topluluklarının toplamından oluşan sosyal ağ kullanıcılarını karşılaştırmak sosyolojik bir yaklaşım değildir. Peki ama giderek artan bir kullanıma sahip olan sosyal ağlar sosyolojinin konusu olamaz mı? Elbette olur ancak bu görece yeni ve en geniş tanımlamayla “dijital sosyoloji” olarak da adlandırılan inceleme alanı kendine has araştırma yöntemleri ve sınırlılıklarını da beraberinde getirecektir. Sosyal ağlar üzerinde ele alınan topluluklar canlı insan toplulukları değildir; bilgisayar dolayımı ile iletişime birbiriyle etkileşime geçen, ağın etkisiyle hareket eden, ağa bağlı olarak kendilerine has özellikleri olan, bilim kurgu yazarlarının tabiriyle nitelersek belki de yarı insan yarı robot siborglar şeklinde aktörlerdir.

Sosyal ağlarda benzer ilgi alanlarına sahip kişiler Fruchter'in “5C” olarak tanımladığı (Conversation, Commenting, Community, Collaboration, Contribution) yani sohbet, yorum, topluluk, iş birliği, katkı gibi unsurlarla birbirleriyle bağlantı kurarlar (Fruchter, 2009). Sosyal bilimciler olarak bizi özellikle sosyal ağlarda oluşan topluluklar ve bunların birbirleriyle ilişkileri ilgilendirmektedir. Sosyal medya ağları genellikle yukarıda da belirtilen sohbet ve yorum üzerinden oluştuğu için benzerlikler ve karşıtlıklar üzerinden birçok sanal topluluğun oluştuğu mecralardır. Bu mecraların dijital olmasını sağlayan analog olarak niteleyebileceğimiz gerçek hayattan farklı olarak atılan her adımın elektronik ayak izleri bırakması ve takip edilebilmesidir. Dijital ortamda kurulan bu bağlantı ve verilerin çözümlenmesinin de kendine has yöntemleri vardır.

Belirttiğimiz gibi dijital ortamlar sosyal ortamlardır. Medya yerine “sosyal medya” teriminin kullanılma sebebi de zaten budur. Sosyal medya ortamlarında kullanıcılara sunulan özellikler, bu sosyalliği görünür olma, beğeni alma, yorum alma, paylaşımın yayılmasının sağlanması gibi geri dönüşümlerle sürekli besler. Süreklilik arz eden bu etkileşimler sosyal medyada yer alan bireylerin izi sürülebilir bir ağ kurmalarına neden olur. Toplumsal ağların analizi, öncü birçok çalışmaya karşın sosyolojideki gerçek atılımını ancak 20. yüzyılın ikinci yarısında gerçekleştirmiştir. Bir ağı oluşturan ilişkiler resmî olmayan toplumsal ilişkilerdir ve toplumsal ağ örgütlü ama kurumsallaşmamış bir biçimdir. Toplumsal aktör tek başına yaşayan bir varlık değildir ve sosyoloji de toplumsal ilişkilerin aldıkları biçimleri sınırlandırarak oluşan ağları çözümlenmektedir. Bu ilişki biçimleri, günümüzde iletişim araçlarının gelişmesiyle ulusal çerçevenin dışına taşmıştır ve artık (world wide web, www) dünya çapında ağ tanımına uygun şekilde büyümektedir (Boudon, Borlandi, Cherkau ve Valade, 2011). Ağ formuna kavuşan ilişkiler de artık içinde bulunduğu ağa göre gelişmektedir.

Yalnızca iletişim olgusunu değil, toplumsal hayatımızı da her yönüyle etkileyen yeni medya aletleri ve buna bağlı olarak gelişen sosyal medya, günümüzün en sık vurgu yapılan kavramlarından birisi hâline gelmiştir. Bilgisayar, tablet, akıllı telefonlar, akıllı kol saati, sanal gözlük gibi teknolojik aletler ve bu teknik altyapı üzerinde çalışan çeşitli programlar vasıtasıyla her türlü toplumsal pratiğimizi biçimlendiren bir güce kavuşmuştur. Romantik ilişkiler sosyal medyada başlayıp bitmekte, sevinçlerimiz, üzüntülerimiz ve yüzlerce emojiyle ifade edilen birçok duygumuz paylaştığımız yeni medya ortamlarında tüm dünyada görünür olmaktadır. Dünyanın birçok yerinde karşılaşılan “işgal et” eylemleri, çeşitli doğa felaketleri, toplumsal heyecan oluşturan her türlü olayda sosyal medya araçları başrol oynamaktadır. Resmî kurumların yetkilileri, birçok önemli açıklamayı sosyal medya üzerinden yapmakta, hayranı olduğunuz sanatçının evlenme kararını Twitter adresinden milyonlarca takipçisiyle aynı anda magazin gazetecilerinden değil, bizzat kaynağın kendisinden öğrenebilmekteyiz. Böylece dijital ortamlar, eğlencenin, toplumsallaşmanın, siyasal katılımında bulunmanın olduğu kadar farklı öğrenme biçimlerinin de mecrası hâline gelmiştir (Akbulut, 2013).

Bilgisayarların kişiselleşmesiyle 1980'lerin ortalarında sayısal ortam üzerinden sadece metin değil, çok farklı veri formatlarının iletilmesine imkân veren (ses, fotoğraf, video vs. farklı formatlarda veri) bir aşamaya geçildi. 1990'larla birlikte internetin daha kapsamlı bir nitelik

taşıdığı görüldü ve “yeni medya” kavramı kullanılmaya başladı. 2005 yılından sonra bilgisayar tabanlı diğer araçların da kişisel kullanıma girip yaygınlaşmasıyla ve bu araçlarda internet tabanlı YouTube, Facebook, Twitter vs. uygulamalar yaygınlaştıkça “sosyal medya” kavramı benimsendi. İnternette türlü türlü medya yer alıyor, çevrim içi eğitim alınabiliyor, kredi kartı borcu ödenebiliyor, alışveriş yapılabilir, uydu görüntüleri ve sokak görüntüleri üzerinden dünya gezilebiliyor yani hayatın hemen hemen tüm işlevlerinin gerçekleştirilebildiği bir alan sunuluyor. Sonuç olarak internet ve sosyal medya ortamları düpedüz toplumsal birçok işlemlerle karşı karşıya kaldığımız, gerçek toplumun gerçek mekânlarından birisi hâline gelmiştir (Uğur, 2013) .

Makalenin temel konusu Twitter mikroblog sitesinin sosyal bilimciler için sağladığı yeni ve büyük veri denizi ve sunduğu araştırma imkanlarıdır. Bu konunun ele alınmasındaki temel sebep yakın zamanda yeni veri çeşitlerinin ortaya çıkması, bunları elde etme ve değerlendirme imkanlarının da kolaylaşması ve sosyal bilimciler için yeni imkanlar sunmasıdır. Şimdiye kadar hiç olmadığı kadar çok veriye ulaşılabilmesi ve analiz yöntemlerinin de dijitalleşmeye bağlı olarak kolaylaşması konuyu günümüz sosyal bilimleri için oldukça önemli hale getirmektedir. Bu çalışma ile hem bu yeni veri bolluğu durumu hem de sosyal bilimlerin bu yeni durum karşısındaki sınırlılık ve fırsatları olduğu ön kabulü değerlendirilmeye çalışılmıştır.

Çalışmanın ilk kısmında içinde yaşadığımız dijital toplumun ürettiği yeni veri tipleri ele alınmış, buradan bu veri tiplerinin ürettiği büyük veri kavramına geçilmiş ve dijital sosyolojinin bu yeni durum karşısındaki metodolojik sınırlılıkları değerlendirilmiştir. Son olarak da sosyal bilimciler için diğer sosyal medya platformlarına göre veri sağlamadaki cömertliği nedeniyle ayrıran Twitter platformu ele alınmıştır.

2. Dijital Çağ ve Yeni Veri Tipleri

Bahsi geçen bu gerçek toplumun gerçek mekânlarında çevrim içi hizmetlerin kullanımı ile üretilen veri ve veri depolamadaki yeteneklerin yanı sıra analiz kabiliyetinin artışıyla birlikte araştırmacılar, dijital izleme verilerinin insan davranış kalıpları hakkında neler söyleyebileceğine giderek daha fazla ilgi duymaya başlamıştır. Bu yeni veri kaynakları, insan davranışlarını daha önce görülmemiş detay ve ölçekte belgelemektedir. Büyük Veri için analitik araçların yaygınlığı, iletişim alanındaki araştırmacılar için muazzam fırsatlar ve bununla beraber zorluklar sunmaktadır. Twitter özelinde kullanıcı tarafından oluşturulan metinlerin aksine, dijital izleme verileri (bağlantılar, cevaplama ve retweet’ler gibi çevrim içi kullanıcı etkinliklerinin varlığı) alan için yeni bir metodolojik sınırı temsil eder (Freelon, 2014) .

Çevrim içi bir faaliyetle her kullanıcı, çeşitli veri izleri bırakır ve çevrim içi uygulamadaki etkinliğini belgeler. Her bir Google arama terimi, tweet, Facebook veya Foursquare’de her giriş ve paylaşım dijital bir iz bırakır. Bu etkileşimlerin toplamları mühendisler için olduğu kadar sosyal bilimciler için de ilgi çekicidir. Bu veri kümelerindeki kullanıcı etkileşimlerinin kalıp ve dinamikleri, kullanıcı davranışlarını daha iyi anlamaya yarayabilir (Jungherr, 2015). Dijital izleme verilerine baktığımız zaman bir yandan birincil veri olarak nitelenebilecek kullanıcı tarafından doğrudan söz konusu hizmetle etkileşimde bulunarak üretilen, Google’a yazılan anahtar kelime veya bir tweet metni gibi verilere sahibiz. Diğer yandan meta veri olarak adlandırılacak hizmetin kullanıcılarıyla etkileşimi bağlamsallaştırmak için kullandığı, Google’da bir anahtar kelimenin arandığı zamanı işaretleyen bir zaman damgasına, kullanıcının tweet gönderdiği yerdeki GPS koordinatlarına veya bir mesaj göndermek için kullanılan uygulamanın adına sahibiz. Meta-verilerin analizi, araştırmacıların birincil verilerde bulunan kalıpların yorumlanması için potansiyel olarak anlamlı bağlamlar oluşturmalarına izin verir. Bu veri ve meta-verilere çeşitli API’ler (uygulama ara birimleri) ile araştırmacılar tarafından ulaşılabilir. Dijital izleme verileri, kişilerin zaman içindeki davranışlarını belgeleyen, bireysel ve süreç içinde üretilen tüm verileri kapsayacak şekilde mevcuttur. Bu verilerin bir başka güzel tarafı da tüm aktörlerin davranışlarının tek bir sistemde analiz edilmesine imkân sağlamalarıdır. Dijital izleme verileri bir bilgi sisteminin teknolojik tasarımı bağlamında üretilir ve saklanır. Her hizmetin tasarımı ve programcılarının bilgileri saklamaya karar verdikleri biçim, verilerin araştırmacılar tarafından nasıl kullanılacağı konusunda çok önemli etkilere sahiptir

(Jungheer, 2015).

API'ler hem akademik arařtırmalarda hem de ötesinde giderek daha ilgi çekici bir veri erişim şekli hâline gelmektedir. Pek çok popüler sosyal medya hizmeti, program geliřtiriciler, arřiv ve veri tabanları gibi daha geleneksel veri kaynaklarına sosyal bilimciler ve gazetecilerin erişimi için de kullanılabilen API'ler sunar (Puschmann ve Ausserhofer, 2017) .

Dijital cihazlarda, çevrim içi hizmetlerde ve insan davranışlarını belgeleyen veri setlerindeki artış, sosyal bilimlerin çeşitli dallarında dijital alanlardaki arařtırma çabalarının ortaya çıkmasına neden olmuştur. İnsan davranışlarının dijital veri izlerinin analiziyle en sık ilişkilendirilen üç kavram hesaplamalı sosyal bilimler, dijital yöntemler ve büyük veridir. Bu terimlerin her biri farklı bir arařtırma geleneğinden kaynaklanmaktadır ve arařtırma pratiğı ve arařtırma ilgisi açısından farklı yaklaşımlarla bağlantılıdır (Jungheer, 2015). Burada bu kavramlar arasındaki ayrımlara girmeden sosyal bilimciler için bu yeni dijital verilerin analizinin önemine odaklanacağız.

Dijital çağ hayatın her alanına yayılarak büyümekte ve arařtırmacılar için arařtırma imkânlarını artırmaktadır. Arařtırmacılar dijital ortamlarda davranışları gözlemleyebilir, sorular sorabilir, deneyler yapabilir ve yakın geçmişte mümkün olmayan yöntemlerle iş birliğı yapabilir. Bu yeni fırsatların yanı sıra yeni riskler de gelmektedir; arařtırmacılar insanlara yakın geçmişte olmayan şekillerde zarar da verebilir. Bu fırsat ve risklerin kaynağı analog çağdan dijital çağı geçişte yatmaktadır. Bu geçiři fark etmenin bir yolu, günlük yaşamımızdaki değışiklikleri farketmektir. Hayatımızda analog olan birçok şey şimdi dijital hâlde: Belki filmleri bir kamera kullanıyorduk fakat şimdi muhtemelen akıllı telefonumuzun bir parçası olan dijital kamera kullanıyoruz. Belki fiziksel bir gazete okuyorduk ama şimdi bir çevrim içi gazete okuyoruz. Belki nakit para ile ödeme yapıyorduk ama şimdi bir kredi kartıyla ödeme yapıyoruz. Her durumda analogdan dijital geçiş, hakkımızdaki daha fazla verinin dijital olarak yakalanması ve saklanması anlamına gelir (Salganik, 2017).

Dijital çağ, insanların iletişim kurması için yeni yollar açmıştır. Bu yeni iletişim biçimleri, arařtırmacıların yenilikçi anketler yapmalarına ve meslektaşları ve halk ile kitlesel iş birliğı oluşturmalarına izin vermektedir. Bir atın resmini yakalarsanız o zaman bir fotoğrafınız olur. Bir atın saniyede 24 resmini yakalarsanız o zaman bir filminiz olur. Bir film sadece bir demet fotoğraftan ibarettir ama artık o bir fotoğraf değil, filmidir. Arařtırmacılar, fotoğraftan sinematografiye geçişte bir değışiklik yapma sürecindedir. Ancak bu değışiklik, geçmişte öğrendiğimiz her şeyin göz ardı edilmesi gerektiğı anlamına gelmez. Fotoğrafçılığın ilkeleri sinematografi ile ilgili bilgi verdiğinde son 100 yılda geliřtirilen sosyal arařtırma ilkeleri, önümüzdeki 100 yıl içinde gerçekleşecek sosyal arařtırmalara ışık tutacaktır. Fakat değışim aynı zamanda sadece aynı şeyi yapmaya devam etmememiz gerektiğı anlamına gelir. Aksine geçmişin yaklaşımlarını bugünün ve geleceğın yetenekleriyle birleřtirmeliyiz. Daha genel olarak sosyal arařtırmacıların, dijital çağın olanaklarından yararlanmak için sosyal bilimler ve veri bilimindeki fikirleri bir araya getirmeleri gerekecektir (Salganik, 2017). Burada peki bu "Büyük Veri" nedir, diye sormak gerekir. Şimdi fotoğraftan filme geçişe imkân veren Büyük Veri'ye kısaca bakalım.

3. Büyük Veri

İnsan davranışlarını belgeleyen dijital veri kaynaklarından bilgi edinme potansiyeli ile ilişkili en önemli terimlerden biri Büyük Veri'dir. "Büyük Veri" terimi, iki binli yıllarla beraber ortaya çıkmıştır ve günümüzde sıklıkla kullanılmaktadır. Rob Kitchin "*The Data Revolution*" adlı kitabında kapsamlı bir şekilde Büyük Veri'nin yedi temel özelliğini tanımlar (Kitchin, 2014):

- *Hacim*: Büyük Veri, terabayt veya petabaytlardan oluşan verilerden oluşan büyük bir hacme sahiptir.
- *Hız*: Büyük Veri, genellikle "gerçek zamanlı" olarak büyük bir hızla oluşturulur.
- *Çeşitlilik*: Büyük Veri, yapılandırılmış veya yapılandırılmamış olabilir ve resimler, ses dosyaları, metin dosyaları gibi çeşitli biçimlerden oluşabilir.
- *Yorucu*: Büyük Veri, sadece küçük bir örneğı değil, tüm nüfusu yakalar veya yakalamaya çalışır.

- *Yüksek çözünürlük:* Büyük Veri, olabildiğince ayrıntılı, taneli çözünürlüğe sahip ve benzersiz bir endekse sahip olmayı amaçlamaktadır.
- *İlişkisel:* Büyük Veri, genellikle farklı veri kümelerinin birleştirilmesini mümkün kılan ortak alanları içerir.
- *Esnek:* Büyük Veri, tipik olarak oldukça genişlemecidir (Yeni alanlar veriye eklenebilir.) ve ölçeklenebilir (Boyut olarak hızla genişleyebilir.).

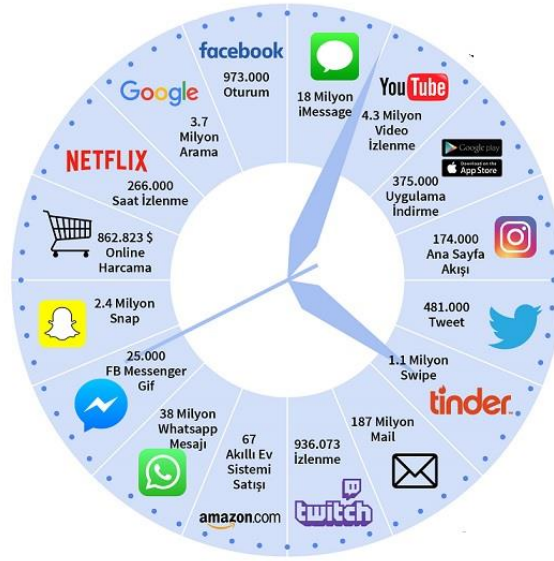
Büyük Veri özünde tahminlerle ilgilidir. Yapay zekâ denilen bilgisayar bilimi dalının bir parçası ve daha spesifik olarak “makine öğrenmesi” şeklinde adlandırılan bir alan ifadesiyle tanımlanmış olsa da bu karakterizasyon yanıltıcıdır. Büyük Veri, bir bilgisayarın “düşünmesini” insanlara benzetmeye çalışmakla ilgili değildir. Bunun yerine olasılıkları çıkarmak için çok miktarda veriye matematik uygulamakla ilgilidir. Bu yaklaşımın uygulandığı sistemler iyi performans gösterir çünkü tahminlerini temel alan çok fazla veri ile beslenir. Dahası sistemler, daha fazla veriyle beslendikçe zaman içinde kendilerini geliştirmek için inşa edilmiştir (Mayer-Schönberger ve Cukier, 2013).

Büyük Veri’nin nerede bulunabileceğini anlamak için “teknolojik etkinleştiricilere” ve “kaynaklara” bakmak gerekir. Teknolojik anlamda Büyük Veri olarak ne toplandığı, nasıl toplandığı ve nerede depolandığı oldukça önemlidir ancak Büyük Veri’yi beslemenin ikinci konusu yani “kaynaklar”, araştırmalarında Büyük Veri kullanmak isteyen sosyal bilimciler için daha çok önemlidir (Bright, 2017). İkinci Dünya Savaşı’ndan bu yana bir dizi teknolojik gelişme ve yenilik Büyük Veri üretimini ve toplamayı kolaylaştırmıştır. Hesaplama gücündeki hızlı ilerleme, ağ teknolojileri, yaygın ve mobil bilgi işlem (ör. akıllı telefonlar, sensörler, cihazlar ve “Nesnelerin İnterneti”) ve özellikle veri depolama alanında yaşanan bir dizi devrim niteliğindeki gelişme bu “teknolojik etkinleştiricileri” oluşturmaktadır. Büyük Veri çok çeşitli kaynaklardan gelebilir. Jonathan Bright, bu kaynakları “veri kurumları” olarak ifade etmektedir. Bunlar, hizmet sağlayıcılardan cep telefonu şirketlerine kadar her şey olabilir; bankalar, devlet daireleri ya da sosyal medya platformları, web sunucuları ve küresel konumlandırma sistemleri (GPS).

Araştırma yapmak için Büyük Veri’ye aşağıdakileri içeren çok çeşitli yollarla erişilebilir:

- Büyük kuruluşların veri sunucularında doğrudan paylaşılan içerik
- Uygulama programlama ara birimi
- Web kazıma (web sayfalarından gelen bilgileri otomatik olarak toplamak ve yapılandırmak için bir bilgisayar programı kullanarak)

Büyük Veri terimi bazen kafa karıştırıcı şekillerde kullanılsa da veri toplama ve işleme konusundaki muazzam ilerlemenin büyük etkisi olduğu yadsınamaz. Büyük Veri’yi analiz etmek için yöntem ve araçların geliştirilmesindeki patlama ile mevcut yeni verilerle sosyal fenomenleri anlamak için fırsatlar artmıştır. Örneğin, 2018 yılına dair sosyal medya verilerinin özetlendiği aşağıdaki grafikte görülen paylaşım içerikleri sosyal bilim araştırmacılarının ele aldığı pek çok sosyal fenomene ilişkin veri içermektedir ve uygulama ara birimleri ile tamamına olmasa da verilere ulaşılabilmektedir.



Resim 1. İnternette bir dakikada neler oluyor (2018) (Lewis, 2019)

Büyük Veri, şehirlerden veya küresel sağlık politikalarından küçük grupların kültürel etkileşimine kadar sosyolojinin ilgi alanına giren hemen hemen tüm sosyal fenomenleri görme ve anlamada yeni yollar sunar. Büyük Veri sağlayıcılar ve kaynaklar büyümeye devam ettikçe bu dinamikler sosyal bilimcilerle giderek daha fazla karşı karşıya gelecek ve bu arada hem büyük fırsatları hem de yeni sorunları önümüze çıkaracaktır. Bu nedenle entelektüel ve pratik olarak Büyük Veri ile ilgilenme becerilerini kazanmak, sosyal bilimcilerin önemli bir konusudur. İş, endüstri ve gündelik yaşamın pek çok alanında Büyük Veri'nin ortaya çıkması, Büyük Veri'ye potansiyel bir sosyal fayda olarak odaklanmanın önemini göstermektedir. Büyük Veri, araştırmayı doğru bir şekilde yapmak için ya da doğru soruları sormak ve Büyük Veri'nin yeni fırsatlarını ve sınırlamalarını anlamak için yeni beceriler gerektirir. Veri bilimi; çok büyük miktardaki bilginin toplanması, hazırlanması, analiz edilmesi, görselleştirilmesi, yönetilmesi ve sunulması ile ilgilenen ve nihai olarak bu veriden anlamlı bilgilerin çıkarılması sürecidir (Çelik, 2018). Veri bilimi geniş bir alandır ve hızla gelişmektedir. Bu durum sosyal bilimcilerin Jonathan Bright'ın "Büyük Veri becerileri boşluğu" dediği şeyle yüzleşmesi anlamına gelmektedir. Ayrıca tüm bunlara ek Büyük Veri kullanımı ile birlikte çok sayıda karmaşık etik sorun ortaya çıkmaktadır.

Büyük Veri ile başa çıkmayı sağlayan pratik yöntemler alanı genellikle veri bilimi olarak adlandırılır. Bu geniş terim büyük miktardaki verilerle uğraşabilmek için gerekli çeşitli süreçleri kapsamaktadır. Veri toplama, veri temizliği ve yapılandırması, veri analizi, veri görselleştirme temel süreçler olarak nitelenebilir. Yukarıda bahsedilen bu becerileri kullanabilmek için de sosyal bilimcilerin bilgisayar tabanlı niteliklerini artırmaları gereklidir. Bundan dolayı da R ve Python gibi programlama dillerinin kullanımı, çeşitli veri görselleştirme programları, Gephi vb. ağ analizi yapmak için kullanılacak programlar tüm dünyada sosyal bilimcilerin gündemine girmiştir. Büyük Veri çağında araştırma yapmak için sosyal bilimciler veri biliminin süreçlerini, potansiyellerini ve sınırlarını anlamalı ve kendi bilim disiplininde bu sürece katkı sunmalıdır.

Dijital çağda sosyal araştırmanın nasıl olması gerektiğini inceleyen Matthew Salganik Büyük Veri'yi sosyal bilim penceresinden değerlendirmiştir. Büyük verinin genellikle araştırma amacıyla oluşturulmadığını vurgulayan Salganik, Büyük Veri kaynaklarının ortak birtakım özelliklere sahip olma eğiliminde olduğunu belirtir. Büyük Veri'nin bazı nitelikleri; büyük, daima açık ve reaktif olmayan şekilde araştırma için yardımcıdır. Diğer tür Büyük Veri ise araştırma için genel olarak sorunlu olan; eksik, erişilemez, temsili olmayan, sürüklenen, algoritmik olarak karmaşık, kirli ve hassas niteliklere sahiptir. Araştırmacılar için bu iki sınıfa tasnif edilebilen Büyük Veri'nin on özelliği aşağıdadır (Salganik, 2017):

- Büyük: Büyük Veri kaynaklarının en çok tartışılan özelliği, büyük olmalarıdır. Örneğin, birçok makale, ne kadar çok veri analiz ettikleri hakkında tartışır ve övünür. Büyük Veri

kümelere kendi içinde bir amaç değildir ancak nadir olayların incelenmesi, heterojenliğin tahmini ve küçük farklılıkların tespiti de dâhil olmak üzere belirli araştırma türlerini mümkün kılabılır.

- Daima açık: Her zaman açık olan Büyük Veri’ler beklenmeyen olayların çalışılmasını ve gerçek zamanlı ölçümün yapılmasını sağlar. Birçok Büyük Veri sistemi her zaman açıktır ve sürekli veri toplar. Bu her zaman açık olan karakteristik, araştırmacılara uzunlamasına veriler sağlar.
- Reaktif olmayan: Büyük Veri kaynaklarında ölçüm yapma davranışını değiştirme olasılığı daha düşüktür. Sosyal araştırmaların bir zorunluluğu, araştırmacılar tarafından gözlemlendiklerini bildiklerinde insanların davranışlarını değiştirebilecekleridir. Sosyal bilimciler genellikle bunu reaktivite olarak adlandırırırılar.
- Eksik: Büyük Veri’leriniz ne kadar büyük olursa olsun, muhtemelen istediğiniz bilgilere sahip değildir. Büyük Veri kaynaklarının çoğu, araştırmanız için isteyeceğiniz bilgilere sahip olmadığı anlamında eksiktir. Bu, araştırma dışındaki amaçlar için oluşturulan verilerin ortak bir özelliğidir.
- Erişilemez: Şirketler ve hükümetler tarafından tutulan verilere araştırmacıların erişebilmesi zor olabilir. Veri erişimini engelleyen ciddi yasal, ticari ve etik engeller vardır ve bu engeller teknoloji geliştikçe de ortadan kalkmaz çünkü teknik engeller değildir.
- Temsili olmayan: Temsil edici olmayan veriler, örnek dışı genellemeler için kötüdür ancak örnek içi karşılaştırmalar için oldukça yararlı olabilir. Bazı sosyal bilimciler, belirli bir ülkedeki tüm yetişkinler gibi iyi tanımlanmış bir popülasyondan gelen olasılıksal rastgele bir örneklemeden gelen verilerle çalışmaya alışkındır. Bu veri türüne temsili veri denir çünkü örneklem daha büyük popülasyonu “temsili eder”, Büyük Veri ise her ne kadar büyük olsa da araştırmacı için temsili olmayabilir.
- Sürüklenen: Nüfus hareketliliği, kullanıcıların değişmesi ve sistem kayması uzun vadeli eğilimleri incelemek için Büyük Veri kaynaklarının kullanılmasını zorlaştırmaktadır. Birçok veri kaynağının en büyük avantajlarından biri, zaman içinde veri toplamasıdır ancak değişimin güvenilir bir şekilde ölçülmesi için ölçüm sisteminin kendisinin kararlı olması gerekir.
- Algoritmik olarak karmaşık: Büyük Veri sistemlerinde davranış doğal değildir, sistemlerin mühendislik hedefleri tarafından yönlendirilir. Her ne kadar birçok Büyük Veri kaynağı reaktif olmasa da (İnsanlar verilerinin kaydedildiğinin farkında değildirler.) araştırmacılar bu çevrim içi sistemlerdeki davranışları “doğal olarak meydana geliyor” düşüncesiyle değerlendirmemelidir.
- Kirlili: Büyük Veri kaynakları önemsiz ve spam olarak yüklenebilir. Yani araştırmacıların gerçek ilgi eylemlerini yansıtmayan verileri içerir. Çoğu sosyal bilimci zaten büyük ölçekli sosyal anket verilerini temizleme sürecine aşındır ancak Büyük Veri kaynaklarını temizlemek daha zor gibi görünmektedir.
- Hassas: Şirketlerin ve hükümetlerin sahip olduğu bilgilerin bir kısmı hassas bilgilerdir. Birçok Büyük Veri kaynağı hassas olan, çoğu zaman erişilememe nedeninin bir parçası olan ekstra bilgilere de sahiptir.

4. Dijital Sosyoloji

4.1 Dijital Verilerin İzlenmesi

Dijital sosyal araştırmanın metodolojik problemlerinin birçoğu, dijital toplumların normatif problemlerine çok benzemektedir. Örneğin, veri merkezilikten kaynaklı olarak sosyal araştırmalar orantısızlık oluşturma riski taşır ve veride yer alan fenomenlere gereğinden fazla dikkat çekebilir. Bununla birlikte dijital toplumlarda bu sorun yalnızca sosyal araştırmayı değil, aynı zamanda sosyal yaşamı da etkiler; bilgi teknolojilerine ne kadar çok güvenirsek o kadar “ışığın en parlak olduğu yer olduğu için anahtarları sokak lambasının altında arayan bir sarhoş” gibi oluruz. Sosyal araştırma ve sosyal yaşam arasındaki bu analogiler bazı açılardan belirsiz olsa

da toplum ve sosyal bilim arasındaki mesafenin kısılmasını sağlayabilir (Marres ve Weltevrede, 2013).

Dijital sosyolojinin vaadi ve değeri sosyal hayatı temsil etme kapasitesinde bulunmakta olup bu metodolojinin günümüz toplumunda reddedilemeyecek bir rolü vardır. Dijital sosyal araştırma, sadece toplumu tanımanın bir yolu değildir, dışarıdan dijital ayarlar, eylemler ve içerik arasındaki karşılıklı uyumun daha geniş dinamiklerine de eşit derecede katılımcıdır. Bu tür etkileşim dinamikleri sosyal toplumlarda sosyal yaşamı tanıma ve bunlara müdahale etme fırsatlarının artmasıyla dijital toplumlarda özel bir öneme sahip hâle gelmiştir (Marres, 2017).

Elde edilebilen verinin büyüklüğü ve artan analiz edebilme yeteneğimiz sayesinde sosyal medya, sosyal bilimciler için tarihin belki de en büyük araştırma alanlarından birini sağlamış bulunmaktadır (Gürsakar, 2009). Dijital ortamda bıraktığımız her türlü iz; çeşitli biçimlerde yayılabilmekte, irademiz dışında piyasada “ürün” hâline gelebilmekte, seçeneklerden hangisini işaretlememiz gerektiği yönlendirilebilmekte ve siyasal gözetimin de nesnesi konumuna düşebilmekteyiz. Dijital ortamda bıraktığımız bütün izler bir aktör hâline gelmekte, farklı sosyal medya uygulamalarında varlığımıza dair her türlü bilgiyi çevrim içi dolaşıma sokabilmekteyiz (Timisi, 2018). Bu çalışmanın da temel savlarından birisi artık sosyal medya olarak nitelenen ortamlarda bir toplumsal yapı oluştuğu, gerçek hayattaki gibi bireylerin kimliklerinin ve buna bağlı olarak grup, topluluk gibi birlikteliklerin oluştuğudur. Bundan dolayı literatürde “dijital sosyoloji” olarak adlandırılan çalışma başlığı altında bu sanal topluluklar incelenmektedir. Bu toplumsal ilişkiler bilgisayar destekli programlarla, büyük veri, yapı analizi / sosyal ağ analizi, grafik teorisi gibi teknik çalışma yöntemlerinin katkılarıyla incelenebilmektedir.

Dijital dünyanın bu denli çok veriyi üretiyor ve Google, Facebook, Amazon, Twitter gibi firmalarca bu Büyük Veri’lerin arşivleniyor olması başta sosyologlar olmak üzere sosyal bilimcilerin rüyasında göremeyeceği bir veriyle karşı karşıya kalmasını sağlamıştır. Bu durum “ampirik sosyolojinin” bir krizle karşı karşıya kalması olarak değerlendirilmesine; bu denli gıpta edilecek büyüklükte verinin sosyolog veya sosyal araştırmacı olmayan çoğunlukla bilgisayar bilimlerine hakim kişilerce de kolaylıkla analiz edilebiliyor olması sosyologların hâlâ araştırma sahasında yerinin olup olmayacağına sorulmasına yol açmıştır. Ancak ilerleyen kısımlarda görüleceği üzere sosyologların bu yeni sahaya deneyimleriyle katkı ve yön verebilecekleri birçok başlık bulunmaktadır. Sosyologlar veri büyük olmadan önce de bu işle meşgul olduklarını esprili bir şekilde belirtmektedirler: “Veri büyük olmadan önce de içindeydim” (I was into data before it was big). Sosyologların hâlihazırda onlarca yıldır üzerine çalıştıkları aile, bilim, sağlık, bilgi, kültür, ekonomi, istihdam, eğitim, iş hayatı, cinsiyet, risk, yaşlanma, ırk, etnisite gibi başlıkların tamamının dijital alanlarla bağı bulunmaktadır. Dijital toplumu çalışmak için tüm bu başlıklarla beraber kişilik, kimlik, güç ilişkileri, sosyal eşitsizlikler, sosyal ağlar, sosyal yapı, sosyal kurumlar ve sosyal teoriye odaklanmak gerekmektedir (Lupton, 2015).

4.2 Sosyal Medya Verilerine İlişkin Metodolojik Sınırlılıklar

Tarih boyunca teknolojideki gelişmeler iletişim şekillerini de değiştirmiştir. Bu değişiklikler de toplumsal hayatın hemen her alanını etkilemiştir. Yeni medya açısından bu değişikliklerden en önemlisi geleneksel medyaya nazaran diyaloga açık yapısıdır. Yeni iletişim ağları tüm toplumsal yapı ile doğrudan etkileşim hâlinindedir. Yeni medya ile birlikte değişen iletişimin yapısı “öteki” ile çok daha sık ve doğrudan diyalog kurabilme imkânı getirmiştir. Pasif kitlelerin yerini hemen tepki üreten kitleler almıştır (Uybadın ve Mumcu, 2013). Bu diyaloga açıklık ve verilerin izlenebilmesi imkânı da bağlantıların oluşmasına ve izlenebilmesine imkân tanımaktadır. Sosyal medyanın benzersizliği bağlantı kurma kapasitesindedir. Tüm medya sosyaldir ancak tüm medyanın yalnızca belirli bir alt grubu temel olarak sosyalikleri tarafından tanımlanır (Bruns, 2015).

Sosyal medya verilerinin sosyal bilimsel analiz için kullanılması, veriler hakkındaki düşüncelerimizin ve bunun sosyal dünyayla ilişkisinin yeniden yapılandırılmasını gerektirir. Veriler gözlemlenip gözlemlenmediğine bakılmaksızın sürekli çoğalmaktadır, araştırma amacıyla oluşturulmamıştır ve bu anlamda akademik çalışmadaki rolü tesadüfi olarak nitelendirilebilir ancak bu durum tutumlar, niyetler, kimlik, ağlar, görüşler, konumlar ve temsiller dâhil olmak

üzere bir dizi sosyal fenomene ışık tutmasının önemini ortadan kaldırmaz. Tabii ki verilerin tesadüfi niteliği sosyal bilimler için tamamen yeni değildir (örnek olarak gözlemsel çalışmalar ve etnografi) fakat sosyal medya verileri doğal olarak üzerinde durmamız gereken özel zorluklar yaratmaktadır. Zorluklar disipline özgü değildir ve hacim, çeşitlilik, hız, gerçeklik, etik ve değer kavramları ile ifade edilir. Bu altı zorluk aşağıdaki gibi açıklanabilir (Williams, Burnap ve Sloan, 2017).

Hacim, sosyal medya platformlarında üretilen veri miktarını ifade eder. Dünyadaki verilerin büyük çoğunluğunun zamanımızda oluşturulduğu tahmin edilmektedir ve bu durum her geçen gün artmaktadır. Örneğin, Twitter'ı ele alırsak tüm dünyada 330 milyon Twitter kullanıcısı, Türkiye'de yaklaşık 8,5 milyon Twitter kullanıcısı bulunmaktadır (Clement, 2019). 2019 itibarıyla tüm dünyada her saniye 6 bin tweet olmak üzere günde 500 milyon tweet atıldığı hesaplanmaktadır (Aslam, 2019). Bu verilerin toplanması ve saklanması önemli zorluklar doğurmaktadır ve faydalı verilerin kirlilik oluşturan veriden ayrılması zaman ve beceri gerektirir.

Çeşitlilik; metin, fotoğraflar, videolar, coğrafi girişler ve ses gibi farklı veri formatlarını ifade etmektedir. Ayrıca sosyal medya platformlarının çoğu zaman birden fazla veri türünü birleştirmesi de söz konusudur. Bu, durum "Büyük Veri" sorununun yalnızca nicel çalışmalar için bir sorun olmadığı ve karışık yöntem çalışmaları da dâhil olmak üzere sosyal medyayı kullanarak birçok veri türünün analizi için devasa bir potansiyeli olduğunu göstermektedir.

Hız, hem sosyal medya verilerinin üretilme hızını hem de kullanıcıların gerçek dünya olaylarına ne kadar çabuk yanıt verdiğini ifade eder. Veri oluşturma hızı, API'ler kullanımı gibi kayıt ve hesaplama çözümlerine ihtiyaç duyan veri toplama için bazı özel çözüm yolları gerektirmektedir. Bu durum sosyal bilimci araştırmacılar arasında bilgi işlem ve kodlama bilgisinin yetersizliğinden dolayı yeni problemlere yol açmaktadır. Bireylerin olaylara cevap verme hızı, bir olay meydana geldiğinde veri toplamaya başlayacak kadar hızlı tepki veren ve "lokomotif" verileriyle zamansallığın rolünü anlayan araştırmacıları gerektirmektedir.

Gerçeklik, öncelikle verilerin doğruluğu, güvenilirliği ve kalitesi ile ilgilidir. Sosyal medya verileri genellikle normalde sosyal araştırmalarda, özellikle de katılımcı ve/veya içerik üreticisinin demografik özelliklerinde standart olarak toplayacağımız önemli bilgilerden yoksundur. Demografik verilerin elde edilmesi, sosyal medyada kimin temsil edildiğini anlamak ve böylece örneklem ve popülasyonlar hakkında daha fazla görüşme sağlamak için anahtardır. Sosyal medya verilerinin gerçek dünya olaylarını nasıl yansıttığı, var olan veri kaynaklarıyla veri bağlantısı yoluyla nasıl ele alınabileceği, şahısların çevrim içi kimliğinin nasıl sunulduğu kaygılarıyla ele alınır.

Sosyal araştırmalara ilişkin mevcut *etik* kurallar, sosyal medya verilerine uygulandığında amaç için uygun değildir ve sosyal bilimler topluluğunun bu tür bir etik çerçevenin nasıl olması gerektiği konusundaki tepkisini koordine etmek için uluslararası çalışmalar yapılmıştır. Örneğin, katılımcıların gizliliği gibi genel etik ilkeler, bazı platformlar için veri kullanımının yasal şart ve koşullarıyla çelişmektedir. Twitter tweet'lerin kullanıcı adı olmadan sunulmasına izin vermeyecektir. Bunun da katılımcıları yanıltıcı olabilecek veriler sunarken (nefret söylemi içeren tweet'ler gibi) zarardan korumaları için etkileri vardır. Sosyal medya için etik bir çerçevenin geliştirilmesi için bir başlangıç noktası, katılımcıların araştırma için kullanılan veriler hakkında ne düşündüklerini anlamaktır.

Değer, sosyal medya verilerinin sosyal dünya hakkındaki anlayışımızı nasıl artırdığına dair bir değerlendirmedir. Şimdiye kadar mevcut olmayan araştırma yollarını açarak ve/veya yeni verilere erişim yoluyla mevcut çalışmaları artırmaktadır. Kuşkusuz yeni "Büyük Veri'lerin" gerçekte ne kadar yeni olduğu ve veri sınırlaması içinde hangi teorilerin rol oynayabileceği hakkında sorular sorulmaktadır. Belki de bu verinin bize getirdiği zorluğun beklenmedik bir sonucu sosyal ve hesaplamalı bilimlerin yanı sıra her iki tarafın da faydalandığı beşeri bilimlerde disiplinlerarası çalışmalarda bir artış olmasıdır.

Sosyal medya çalışmalarını birçok disipline ilişkin yapan, çok büyük miktar ve çeşitte verilerin mevcudiyetidir. Sosyal medya, Büyük Veri olarak adlandırılan veriyi üretir ve yüksek hız, büyük

hacim, çeşitlilik, etraflı bir kapsamlılık, ince taneli çözünürlük ve yaklaşımındaki esneklik ile tanımlanır (Kitchin, 2014). Bu, akademisyenler için yeni zorluklar yaratırken aynı zamanda büyük fırsatlar da sunar. Bu yeni durum birkaç soruya yol açmıştır. İlk olarak sosyal medya araştırması, araştırmacıları doğal zorluklarla başa çıkmak için içe bakmaya zorlayan *metodolojik, etik ve ölçekle* ilgili sorular doğurmuştur (McCay-Peet ve Quan-Haase, 2017):

1- Metodolojik sorular: Sosyal medya verilerinin toplanması, analiz edilmesi ve görselleştirilmesinden dolayı yeni metodolojik sorular ortaya çıkmıştır. Bu sorulardan bazıları platforma özgüdür, diğerleri ise her türlü sosyal medyaya uygulanabilir. Bir dereceye kadar sosyal medya, büyük miktarlardaki verilere kolay ve rahat erişim sağlar, diğer taraftan belirli bir veri setini elde etmek maliyetli hatta imkânsız olabilir.

2- Etik sorular: Veri toplama, bir araya getirme ve sosyal medya verilerinin raporlanması araştırmacıların uğraşmaya yeni başladıkları bu konularla ilgili kişisel gizlilik, doğruluk ve hesap yeteneği gibi konularla ilgili sayısız etik soruyu gündeme getirmiştir. Sosyal medya verileri genellikle kamuya açık olmasına rağmen araştırmacıların üzerine düşünmesi gereken birçok etik husus vardır. Sosyal medya sitelerinin tüm kullanıcıları, analiz için verilerinin kullanımının farkında olmadıklarından onay alınıp alınmaması genellikle tartışmaların merkezindedir (Beninger, Fry, Jago, Lepps, Nass ve Silvester, 2014) .

Araştırmacıların veri yönetimi konusundaki geniş kapsamlı araştırma sorularını ele alırken bir taraftan da kullanıcıların veri gizliliği hakkını koruyarak ne tür etik kuralların oluşturulması gerektiğini belirlemeleri gerekmektedir. Bu tür tartışmalar, çeşitli ülkelerde ortaya konan yeni yasalarla giderek daha fazla tartışılır hâle gelmektedir. Örneğin, Avrupa Birliği, bireylerin web üzerindeki kişisel verileri daha iyi kontrol etmelerine olanak tanıyan ve “unutulma hakkı” adı altında tartışılan yeni bir yasa çıkardı. Yasal zorlukların ortaya çıkmasıyla birlikte araştırmacılar için analiz amacıyla toplanan verilerin ön yargıları ve kullanıcıların silmek isteyebilecekleri verilerin depolanmasının etik ve yasal olması ile ilgili yeni araştırma sorunlarını ortaya çıkarır.

3- Ölçek soruları: Ölçek, sosyal medya araştırmacılarının üstesinden gelmeleri gereken en büyük zorluklardan biridir. Sosyal medya verileri, farklı bir olgunun veya konunun farklı açılardan incelenmesine olanak sağlar. Bir çalışma, terabaytlık bilgileri toplayan Büyük Veri kümelerini ya da küçük ölçekli çalışmalarla birkaç kullanıcının yerel davranışlarını inceleyebilir. Her iki yaklaşım da ilgili ve geçerli olsa da tek bir olguda niteliksel olarak farklı görüşler sağlar. Öyleyse bu kadar farklı bilgi edinme yöntemlerinden elde edilen bulguları nasıl entegre ederiz? Farklı veri setlerini ve bulguları birbirine bağlamak ve entegre etmek için yeni teorik ve metodolojik varsayımlara ihtiyaç vardır.

İkinci olarak sosyal medya verileri, araştırmacılara yeni soru türlerini ele alma ve mevcut araştırma sorunlarına farklı bir açıdan ışık tutabilme imkânı sunmaktadır. Bunlar sosyal medya kullanımı ve sosyal olgulara ilişkin sorulardır (McCay-Peet ve Quan-Haase, 2017):

1 - Sosyal medya kullanımı ile ilgili sorular: Sosyal medya üzerine yapılan çalışmalar, bireylerin, kuruluşların ve hükümetlerin sosyal alanlardaki etkileşimi ve etkileşiminden kaynaklanan yeni sorulara cevaplar sağlayabilir. Sosyal medya etkinliği, bu durumda araştırmanın odak noktasıdır (örneğin, kişilerin kişisel sağlıkla ilgili sorunları Facebook’ta nasıl ele aldıkları, Twitter ağlarının topolojik özellikleri ve sosyal medya kullanım modellerini nasıl tartıştıkları). Örneğin, sosyal medya araştırması iş yerleri, okullar ve üniversiteler açısından önemlidir. Bu kuruluşlar ve kurumlar, desteklenecek uygun politikalar geliştirmek veya bazı durumlarda kullanımını kontrol etmek için sosyal medyayı anlamalıdır. Örneğin, akademik atıf sayımları, akademik ödül sistemini bilgilendiren geleneksel bir araştırma etkisinin ölçüsüdür ancak üniversitelerin araştırmalarının sosyal etkisini ölçmede yeni bir yöntem olarak kullanılıp kullanılmayacağını ve nasıl kullanılabileceğini anlamaları gerekir (Holmberg ve Thelwall, 2014).

2 - Sosyal olguları anlamamızı sağlayan sorular: İkinci bir sosyal medya araştırması yaklaşımı, sosyal medyanın akademik araştırma için bir araç veya yöntem olarak kullanılmasını, araştırma sorunlarını incelemek ve başka, daha geleneksel yöntemlerle yapılan karmaşık sorunları anlamak için kullanmaktır (Otieno ve Matoke, 2014). Bu ikinci tür sosyal medya araştırması, “sosyal

medya araştırması” yapan bilim insanlarının sosyal medyayı açıkça çalışmayan ancak araştırmalarını sosyal medya özelliklerinden faydalanabilecek disiplinlerden alanlara getirerek çeşitli fenomenleri incelemesidir: Bir işe alım platformu olarak potansiyeli, belirli bir demografik yapıya ulaşması ve sosyal medya aracılığıyla kolayca gözlemlenebilen ve çıkarılabilen davranışlar, tutumlar ve algılar gibi temelini genişletmektedir. Toplumsal hareketlere katılım, siyasi katılım ve tüketim gibi sosyal olgular sosyal medya verilerinin analizi ile incelenebilir.

McCay-Peet ve Quan-Haase sosyal medya katılımını analiz edebilmek herhangi bir disiplin, teorik ve metodolojik bakış açısı ve gelenekleri ile incelenebilecek altı sosyal medya unsuru tanımlamışlardır: (1) kendini sunma, (2) eylem ve katılım, (3) kullanımlar ve doyumlar, (4) olumlu deneyimler, (5) kullanım ve faaliyet sayıları ve (6) sosyal bağlam (Quan-Haase ve Sloan, 2017). Bu model artırılarak yedinci unsur olan (7) platform özelliği eklenmiştir. Kendini sunma, eylem ve katılım, kullanımlar ve doyumlar, olumlu deneyimler, kullanım ve faaliyet sayıları, sosyal bağlam ve platform olarak nitelenebilecek yedi özelliğin her biri aşağıda kısaca tarif edilmiştir ve birbirleriyle ilişkili olarak nasıl çalışılabileceklerini göstermektedir (Quan-Haase ve Sloan, 2017).

1. Benliğin sunumu, itibar yönetimi ve mahremiyet: Kimlik, sosyal medyada zaman içinde kişisel bir profil veya sanal benliğin gelişimi yoluyla oluşturulur. İnternette mahremiyet kavramı, kişisel bilgilerin çevrim içi olarak aranabilmesi, etiketlenmesi ve toplanması için geliştirilmiş yetenekler nedeniyle büyük ilgi görmüştür. Özellikle dikkat çeken bir çalışma alanı, sosyal ağ sitelerinde (SAS) kişisel bilgilerin sağlanmasıdır. SAS’lar büyük miktarda kişisel bilginin depolandığı ve toplandığı benzersiz bir sosyal alanı temsil eder ve en önemlisi bu sitelerdeki veriler kolayca kopyalanabilir, iletilebilir, çoğaltılabilir ve bağlamdan çıkarılabilir (Young ve Quan-Haase, 2013).

2. Eylem ve katılım: Sosyal medya, kullanıcıların içeriği, iş birliğini ve tartışmaları görüntüleme, gönderme veya paylaşma gibi çeşitli etkinlikler gerçekleştirmesini sağlar.

3. Kullanımlar ve doyumlar: Sosyal medya kullanıcıları, örneğin, bilgi alışverişi ve kullanımından elde edilen sosyal faydalar için farklı motivasyonlara sahiptir: İnsanların neden sosyal önerileri kullandıklarını anlamak için (örneğin, beğeniler). Üniversite öğrencilerinin Facebook’taki “beğenme” gibi sosyal öneri sistemlerini kullanımını incelemek için araştırma yapan Kim bu sistemlerin kullanıcılara, belirli bir çevrim içi topluluk içindeki görüşlerin iklimini tarama fırsatı sunduğunu belirtir. Araştırma sonucunda sosyal tavsiyelerin ifade, bilgi, sosyalleşme ve eğlence amaçlı kullanıldığı gösterilmiştir (Kim, 2014).

4. Olumlu ve olumsuz deneyimler: Sosyal medyanın, insanları olumlu duygular, huzur ve akış gibi faktörlerle kullanmaya iten yönleri nelerdir. Sosyal medyanın oldukça popüler hâle gelmesi ve sık etkileşimleri desteklemesi nedeniyle kullanıcılar sosyal medya kullanımına bağımlı hâle gelebilmekte, fiziksel dünyadan kopabilmektedir. Bu kullanıcıların medyayı paylaşmak ve diğer kullanıcılarla etkileşim kurmak gibi etkinliklere bu kadar derinden dâhil olma sebebi oyunculuk ve özne katılımının akış durumunu açıklayan akış teorisi ile açıklanır (Chan ve Ma, 2014). Aynı şekilde spam, sahtekârlık ve siber zorbalık gibi olumsuz deneyimler, kullanıcının sosyal medya ile etkileşimini azaltabilir ve bu deneyimler çok sayıda araştırma sorusuna yol açar.

5. Kullanım ve etkinlik sayıları: Kullanım ve etkinlik sayıları, kullanıcıların eylemleriyle ve belirli bir sosyal medya sitesindeki katılımlarıyla ilişkili verileri ifade eder; bu veriler kullanıcılara gerçek zamanlı olarak ham veya toplu biçimde sunulabilir. Bu alandaki araştırmalar, örneğin, Twitter ve Facebook gibi sosyal medya sitelerinin sağladığı istatistiklerin kullanıcılar üzerindeki etkisini inceleyebilir.

6. Sosyal bağlam: Sosyal bağlam, sosyal, politik, ekonomik, iş ve kişisel olayları ya da kullanıcıların bu yerel ve küresel ağların büyüklüğü ve niteliği dâhil olmak üzere sosyal medya sitelerinde bir kullanıcının sosyal ağlarının altında yatan özellikleri ifade eder. (örneğin, küçük, sıkı sıkıya bağlı bir eş grubu; geniş, dağınık bir sosyal aktivist ağı).

7. Platform özellikleri: Belirli platformların özelliklerine ilişkin faktörler, katılımcıları etkileyebilir. Örneğin, kullanıcıların bilgi paylaşmasını veya doğrudan birbirleriyle iletişim

kurmasını sağlayan özellikler. Platform özellikleri, kullanıcıların içeriği nasıl oluşturduklarını, paylaştıklarını, etkileşime girdiklerini ve harekete geçirdiklerini ve aynı zamanda topluluğun farklı platformlarda nasıl oluşturulduğunu ve korunduğunu anlamak için kritik öneme sahiptir. Örneğin, Twitter, kişilerin karşılık vermeden bir kişiyi, kurumu veya hesabı takip etmesine izin verir. Yani Twitter tek yönlü bilgi akışını destekliyor. Buna karşılık Facebook yalnızca iki tarafın da bağlantıyı kabul ettiği durumlarda kullanıcılar arasındaki bağlantılara izin verir: iki kullanıcı birbirine eşit olarak bağlanır. Özelliklerin platformlarda nasıl işlediğine dair bu fark bilgi akışı, eşik bekçilerinin oluşumu ve ağların topolojisi için önemli etkilere sahiptir. Bu nedenle platform özelliklerinin sosyal olgularla nasıl ilişkili olduğunu anlamak bu platformların sosyal davranışlar için uygunluğunun daha iyi kavranmasını sağlar (McCay-Peet ve Quan-Haase, 2017). Sosyal ağ siteleri (SAS'lar) üzerine yapılan araştırmalar, kullanıcı tabanı, kullanıcı uygulamaları ve bu araçların özellik kümelerinin giderek farklı olmasına rağmen SAS kullanımını homojen olarak değerlendiren ifadeler kullanılır. Kullanımlar ve doyumlar perspektifinden Facebook'u değerlendiren araştırmacılar Facebook'un genel kullanımına yönelik motivasyonların, belirli özelliklerin kullanımına yönelik motivasyonlardan farklı olduğunu bulmuştur (Smock, Ellison, Lampe ve Wohn, 2011).

Yukarıda özetlenen araştırma sorularına verilen cevaplar; sosyal medya erişimine, insanların sosyal medya ile nasıl ve niçin etkileşime girdiğine ve bunun nasıl bir etkisi olduğuna ayrıca toplum anlayışımıza ilişkin bilgi vermektedir.

4.3 Twitter Mikroblog Sitesi

Toplumsal araştırmalar için Büyük Veri'nin potansiyeli çok büyük olsa da gizlilik veya ticari duyarlılık gibi sebeplerle bu veri setlerinin çoğu hükümetlerin veya özel şirketlerin elinde kalmaktadır. Bunun önemli bir istisnası, içeriği (neredeyse tamamen) herkese açık olan ve Twitter'ın kendi API'si aracılığıyla erişilebilen mikro blog sitesi Twitter'dır. Twitter veri paylaşımı açısından görece cömert yapısıyla sosyal bilimciler arasında oldukça ilgi çekicidir (Tinati, Halford, Carr ve Pope, 2014).

İnternetin gelişimine bağlı olarak Facebook veya Twitter gibi çeşitli çevrim içi hizmetlerin sürekli artan bir kamusal kullanımına şahit olmaktadır. Bu hizmetler, kullanıcılarına bilgi yayınlamalarını, hayatlarının sıradan yönlerini belgelemelerini, arkadaşlarıyla, aileleriyle ve yabancılarla çevrim içi etkileşimde bulunmalarını ve günlük işleri daha kolay gerçekleştirmelerini sağlar. Bu araçlar kullanıcılarının günlük hayatlarını tamamlamakla beraber politik katılım, kampanya ve aktivizmi de değiştirir. Dünyanın pek çok yerindeki seçim kampanyaları, 2010 ve 2011 yıllarında Arap dünyasındaki ayaklanmalar sırasında ya da çeşitli deprem, sel gibi felaketlerde, aktivistlerce düzenlenen çeşitli kampanyalarda internetin ve özellikle Twitter'ın oynadığı önemli rol aşikârdır. Twitter'da yer alan #konuetiketlerinin hızlıca gözden geçirilmesi dijital araçların kampanya, politik söylem ve aktivizmle bütünleştiğini gösterir. Bugünün sosyal meselelerini ve siyasetini anlamak için çevrim içi ve çevrim dışı etkileşimleri anlamak zorundayız. Birçok online hizmet; kampanya yürütücüler, aktivistler ve halkı bilgilendirmek ya da politika hakkında konuşmak için kullanılıyor olsa da mikroblog servisi Twitter en popüler araçlardan ve tüm bu saydıklarımıza en açık olanlardan biridir (Jungherr, 2015).

Sosyal medya kategorilerinden biri olan mikrobloglar anlık ve kısa içerik paylaşımına izin vermektedir. En çok kullanıcıya sahip olan mikroblog sitesi de Twitter'dır. Twitter kullanıcılarına tweet (cikleme, cıvılda) adı verilen önceleri 140 karakterle sınırlıyken sonra 280 karakter limitli metinler yazma imkânı vermektedir. Kullanıcılar eğer hesabını kısıtlamadıysa atılan tweet'ler tüm kamuoyu tarafından görülebilmekte, kullanıcılar birbirlerini tek yönlü olarak takip edebilmekte, kullanıcıların sadece yazdıkları değil, takip ettikleri kişiler, beğendikleri tweetler görülebilmektedir (Eraslan ve Eser, 2015). Özellikle güncel konulara ilişkin olarak kullanılan #hashtag'ler yani #konu etiketi ile hem atılan tweetlerin görünür olması sağlanmakta hem de belli bir konu hakkında yazılanlar kolaylıkla takip edilebilmektedir.

Twitter 2006'da ortaya çıkmasından bu yana mesaj iletimi temelli kurulmuş bir arkadaş takip ve mesajlaşma uygulamasıdır. Twitter dünyada yaşanan kriz ve seçim gibi yoğun yaşanan süreçlerde

de iyi performans sergileyerek dünya çapında bir olay takip uygulamasına dönüşmüştür. Böylece öğlen-ne-yedim medyası imajından sıyrılarak Amerikan Kongre Kütüphanesinde arşivlenen bir veri seti hâline gelmiştir (Weller, Burns, Burgess, Mahrt ve Puschmann, 2016). Twitter faaliyete geçtiği 2006'dan beri 3 evrede incelenebilir. Aşağıda kısaca özetlendiği üzere birinci evre arkadaş takibi medyası, ikinci evre olay takibi için haber medyası takibi, üçüncü evre arşivlenen bir veri seti ve tahmin medyası olarak adlandırılmaktadır (Weller, Burns, Burgess, Mahrt ve Puschmann, 2016).

Birinci evre olan arkadaş takibi medyasına ilişkin Twitter araştırmalarının çoğunda yazılımın arkadaşlar arasında kullanılan kent ve mobil yaşama biçimine yönelik bir gereç olarak kökleri bir bakıma hizmetin adının etimolojisinin ve genel olarak da tweet'lerin yararsızlığının gölgesinde büyük ölçüde kaybolmuştur. Tweet'lerin ne kadar sıradan olduğuna dair değerlendirmeler bu dönemin özelliğidir.

Twitter'ın ikinci evresi olan olay takibi için haber medyası Kasım 2009'da, Twitter'ın etiket satırının değişmesiyle başlatılabilir. Twitter kullanıcılarının karşılaştığı "Ne yapıyorsun?" sorusu "Neler olup bitiyor?" şeklinde değiştirildi. Dil bilimci David Crystal'a göre Twitter, insanları tweet'leri koruyarak ve izole tweet'leri yayınlamaktan ziyade çok daha fazla etkileşime sokarak daha fazla belirleyici olmuştur. Sonuç olarak odak, bireyden gruba kaymıştır ve bu vurguyu yakalamak için daha açık bir soru sorulmuştur. "Ne yapıyorsun" içe bakar, "Neler olup bitiyor" dışarıya bakar (Tate, 2009).

Twitter, üçüncü döneminden itibaren insanların hayatlarındaki sıradan olaylarla ilgili laf salatası ve dikkat çekici bir olay sırasında arka planda etkileşime girmeyi sağlayan bir yol olarak algılanan hâlimden gittikçe uzaklaşarak, dünyada yaşanan krizlerde acil durum iletişim platformu olarak, seçim dönemlerinde sürece aktif katılım mekanizması olarak, spor karşılaşmalarında taraftarların harp meydanı olarak, sanatçıların hayran kitlelerinin münazara mekânı olarak kullanılmaya başlanmıştır. Çeşitli alanlardan kampanya yürütenler, lobiler, şirketler, sivil toplum kuruluşları, siyasi partiler, aktivistler platformu görüşlerini yaymak ve potansiyel hedef kitleleriyle bağlantı kurmak için kullanılmaktadırlar (Weller, Burns, Burgess, Mahrt ve Puschmann, 2016). Böylece araştırmacıların ulaşabilmesi için Amerikan Kongre Kütüphanesi tarafından arşivlenip ulaşılabilir hâle getirilen bir veri seti hâline ulaşmıştır. Böylelikle Twitter, tweet'leri almak, saklamak ve analiz etmek için veri kaynağı olarak değerlendirilmektedir. Twitter tweet'lerin toplanmasındaki ve koleksiyonların hazırlanmasındaki göreceli kolaylığa ve ayrıca anlamlı tweet'ler için retweet'ler, konu kategorizasyonu için hashtag'ler (konu etiketi), ağ analizi için retweet'ler, reply'lar (cevaplar) ve takipçiler-takip edilenler ve referans analizi için kısaltılmış URL'ler de dâhil olmak üzere analizin dâhilî araçları sayesinde araştırma yapmak için özellikle cezbecidir. Twitter özellikle web kazıma teknikleri ile veri elde etmek için de oldukça elverişli bir platformdur (Marres ve Weltevrede, 2013).

Twitter'ın altı özelliği akademik çalışma konusundaki popülerliğini artırmıştır (Steinert-Threlkeld 2018):

Birincisi, hemen hemen her ülkeden aylık 330 milyon aktif kullanıcısı olan en büyük sosyal ağlardan biridir (Aslam, 2019). Bu kullanıcılar arasında devlet başkanları, şirketler, kâr amacı gütmeyen kuruluşlar, uluslararası sivil toplum kuruluşları, ünlüler, sporcular, gazeteciler, akademisyenler ve temel olarak normal insanlar var. Bu nedenle Twitter, bir araştırmacının ilgileneceği hemen hemen her grubun bir kesitini sunar.

İkincisi, kullanıcılar günde 500 milyon mesaj olan çok miktarda veri üretir. Tüm bu insanlar ve mesajlar, Twitter'ın, başka türlü büyük araştırma ekiplerinin eş zamanlı olarak analiz etmelerini gerektiren nüfusun geniş kesimlerini yansıttığını gösterir. Birlikte ele alındığında bu ilk iki özellik, neredeyse her etkinliğin Twitter'da kaydedildiği ve bunun sonucunda pek çok etkinliğin tahmin edilebileceği anlamına gelir.

Üçüncüsü, bu verileri elde etmek nispeten kolay hâle getirilmiştir. Twitter, kullanıcıların verilerini, Twitter hesabı olan herkesin erişebileceği iki API, akış API'si ve REST API'si aracılığıyla sağlar. API, API'lerin v1.1'ine geçmeden önce üçüncü taraflara, programlama

deneyimi olmayan kişilerin API'lerine erişmesine izin veren arayüzler sunmalarına izin verdi. Şimdi bu yetenek artık mevcut değil ve veri edinme girişine engel oluşturmaktadır. Bununla birlikte Twitter'ın popülerliğinden dolayı Twitter'a Python ve R de dâhil olmak üzere erişilebilecek çok sayıda yazılım kütüphanesi vardır. API'lerle etkileşimde bulunmak için biraz programlama bilgisine ihtiyaç vardır.

Dördüncüsü, API'ler, alınan verilerin belirli bir araştırma sorusuna uyarlanmasını kolaylaştırır. Her gün tüm tweet'lerin %1'i (akış API'sinden) alınabilir veya kullanılan anahtar kelimeler, kullanıcı yerleri, kullanıcı kimlikleri veya dil temel alınarak alınan tweet'ler filtrelenebilir. REST API'si yoluyla belirli tweet'ler, kullanıcının en son tweet'lerinden 3,200'ü, kullanıcının kimleri takip ettiği veya kullanıcıyı kimin takip ettiğini gösteren bir liste ve kullanıcı profili bilgileri indirilebilir. Başka bir deyişle Twitter birçok "Büyük Veri" teknolojisine öncülük etmiş olsa da Twitter'ın verilerine erişmek için bu becerilere sahip olmak gerekmez. Verileri toplamak için yeni beceriler öğrenmek gerekebilir ancak bu verilerin modellenmesi ve görselleştirilmesi işlemi eski araçlarla yapılabilir.

Beşinci olarak Twitter ağ ve ağ dışı analizler için mükemmel bir veri kaynağıdır. Hizmet açıkça bir ağ olarak yapılandırıldığından -Hesaplar arasındaki bağlantılar kullanıcı deneyiminin temel yapı taşlarıdır.- difüzyon süreçleri ve ortaya çıkan davranışlarla ilgilenen araştırmacılar Twitter'ı doğal bir kaynak olarak görmektedirler.

Altıncı olarak Twitter'da Facebook'ta bulunmayan bir halkla konuşma normu vardır. Facebook ayrıca bir API sağlarken çoğu kullanıcı bilgilerini herkese açık tutmamayı tercih eder. Twitter; kullanıcılarının konuşmasını, diğer kullanıcılarla etkileşime girmesini, açıklığı teşvik ettiği kadar araştırmacıları da destekler.

Tablo 1: Twitter kullanım sözleşmeleri (Jungherr, 2015)

Kullanım Tanımı	Tanım
Başka bir kullanıcıya @reply to	Herkese açık olarak diğer bir Twitter kullanıcıya cevap vermek için @reply to özelliği kullanılır. (ör. @kullanıcıadı)
Başka bir kullanıcıya @mention	Yazılan bir tweet metninde başka bir kullanıcıya atıfta bulunulabilir. (ör. Tweet metninde @kullanıcıadı)
RT verbatim	Bir tweet hiç değişiklik yapmadan olduğu gibi retweet (RT) edilebilir. (ör. RT @kullanıcıadı)
RT modified	Bir tweet içeriğinde değişiklik yapılarak da retweet edilebilir.
#anahtar kelimeler	Bir tweet içeriğinde bağlam oluşturmak için anahtar kelimeler kullanılabilir. Bunlara #hashtag yani #konu etiketi denir.
Haricî sitelere link, resim ve video paylaşımı	Tweet içeriğinde haricî bir internet sitesine atıfta bulunup link verilebilir, resim, video gibi içerik paylaşılabilir.

Twitter'da en yaygın kullanım alışkanlığı muhtemelen @reply şeklindedir: Bir kullanıcı doğrudan Twitter'da başka bir kullanıcıyla iletişim kurmak veya doğrudan bir tweet'e cevap vermek isterse Twitter mesajını bir @ işareti ile ve ardından iletişim kurmak istediği kullanıcının kullanıcı adını yazarak iletir. @Reply'lar, yalnızca mesaja hitap ettiği kullanıcıyı ve tweet'i gönderen kullanıcının hesabını takip eden kullanıcıların haber akışlarında görünür. @reply'lar geneldir ve bu yüzden onları gönderen kullanıcının profilinde yayınlanan mesajlar listesinde açıkça görünür ve Twitter arama fonksiyonu kullanılarak bulunabilir. Tweet başlangıcında değil de bir mesajın gövdesinde @ kullanıcı adı kombinasyonunu kullanmak da mümkündür. Buna @mention yani

atf denir. @reply'dan farklı olarak tweet yazısını takip eden her kullanıcının haber beslemelerinde @mention içeren mesajlar görünür. Twitter tarafından @reply veya @mention almışlarsa ve onlara tepki göstermeleri istenirse kullanıcılar bilgilendirilir. Bir kullanıcıyla özel olarak mesajlaşmak da @ username şeklinde mümkündür. Doğal olarak bu doğrudan iletilere Twitter'ın uygulama ara birimi aracılığıyla erişilemez ve bu nedenle Twitter'ın uygulama ara birimini veri kaynağı olarak kullanan araştırmanın temeli olarak kullanılamaz. Twitter kullanıcıları arasında başka popüler bir kullanım türü de retweet etmektir. Bir kullanıcı, haber bülteninde gördüğü bir mesajı beğenir ve iletiyi takipçileriyle paylaşmak isterse iletiyi profiline yeniden gönderebilir. Bunu yapmak için söz konusu mesajı kopyalar, RT @ kullanıcı adıyla (orijinal yazara ait) önceler ve bu yeni mesajı hesabına gönderir. Bir mesajı yeniden gönderen kişinin mesaja yeniden gönderilmesini istediği mesajı değiştirmesi veya eklemesi de mümkündür. Bu değiştirilmiş bir retweet olarak bilinir (Jungheer, 2015).

5. Sonuç ve Tartışma

Dijital çağda dijital yeni veri tiplerinin devasa miktarlarda üretilmeye başlanmasıyla beraber araştırmacıların da dikkati bu alana kaymıştır. Genel olarak “Büyük Veri” kavramı ile ifade edilen bu dijital veri kaynaklarından bilgi edinme potansiyeli ile hacim, hız, çeşitlilik, yorucu, yüksek çözünürlük, ilişkisel, esnek olarak nitelenen devasa veriler ortaya çıkmıştır. Sosyal medya da Büyük Veri olarak tabir edilen veriyi üreten başlıca kaynaklardandır ve bu durum araştırmacılar için yeni zorluk ve fırsatlar ortaya çıkarmıştır. Metodolojik, etik ve ölçekte ilgili sorular ve sosyal medya kullanımı ve sosyal olgulara ilişkin sorular olarak sıralanabilecek soruların cevaplandırılması gerekmektedir. Sosyal bilimciler de dijital çağ ile yeni bir döneme girmiştir ve önlerine çıkan bu sorulara cevap vermek durumundadırlar.

İnsan davranışlarını belgeleyen dijital cihazlarda, çevrimiçi hizmetlerde ve veri setlerindeki artış, sosyal bilimlerin çeşitli dallarındaki dijital alanlarda araştırma çabalarına yol açmıştır. İnsan davranışının dijital veri izlerinin analizi ile ilişkili en yaygın üç kavram, hesaplamalı sosyal bilimler, dijital yöntemler ve büyük veridir. Bu terimlerin her biri farklı bir araştırma geleneğinden kaynaklanmaktadır ve araştırma pratiği ve araştırma ilgisine farklı yaklaşımlarla bağlantılıdır.

Yaptığımız çalışma ile sosyal ortamlar olan sosyal medya platformlarında “Büyük Veri” olarak nitelenen çeşit ve çoklukta veri üretildiği gösterilmiştir. Bu ortamlarda sohbet, yorum, iş birliği, katkı gibi unsurlarla bireyler bağlantı kurmakta ve ağda bir topluluk inşa etmektedirler. Bu toplulukların incelenmesinin sosyal bilimciler için hem yeni fırsatlar hem de sınırlılıklar getirdiği de çalışmada gösterilmiştir.

Çevrimiçi hizmetlerin artan kullanımı ile, veri ve veri depolama yeteneklerinin yanı sıra analiz yeteneği de artmıştır. Böylece, araştırmacılar dijital izleme verilerinin insan davranış kalıpları hakkında neler söyleyebileceğiyle giderek daha fazla ilgilenmektedirler. Bu yeni veri kaynakları, insan davranışını benzeri görülmemiş ayrıntı ve ölçekte belgelemektedir. Büyük Veri için analitik araçların yaygınlığı, iletişim alanındaki araştırmacılar için muazzam fırsatlar ve zorluklar sunmaktadır.

Dijital çağın yeni veri kaynaklarından başlıcası Twitter' dır ve veri sağlamadaki cömertliği ile de sosyal bilimciler için cazip alanlardan biridir. Twitter'da kullanıcı tarafından oluşturulan metnin aksine, dijital izleme verileri (bağlantılar, yanıtlar ve retweetler gibi çevrimiçi kullanıcı olaylarının varlığı) alan için yeni bir metodolojik sınır oluşturur. Genel olarak sosyal medya platformları özel olarak da Twitter'ın her biri farklı mimari ve içerikte olmaları sebebiyle sosyal bilimciler için gözden kaçırılmaması gereken mecralar olduğu ancak bunun için araştırmacılar tarafından yeni beceriler kazanılması gerektiği çalışma ile gösterilmiştir. Nihai olarak Büyük Veri üreticisi en önemli kaynaklardan biri olan sosyal medyanın ve özel olarak Twitter'ın bağlantıların kaydedilip sosyal ağ analizi yöntemleriyle incelenmesi ve topluluk seviyesinde değerlendirilebilmesi nedeniyle sosyal bilimciler için hayatın her alanından bir veri kaynağı olduğu değerlendirilmektedir.

Kaynakça

- Akbulut, H., (2013). "Kongre Açılış Konuşması". Özçetin, B., Göker, G., Bayraktutan, G., Sayımer, İ. ve Çomu, T. (Der.), Yeni Medya Çalışmaları. Kuram, Yöntem, Uygulama ve Siyasa I. Ulusal Kongresi Kongre Kitabı s. 8. Alternatif Bilişim Derneği Genel Merkezi, Ankara.
- Aslam, S., (2019, Eylül). "Twitter by the numbers: Stats, demographics & fun facts", <https://www.omnicoreagency.com/twitter-statistics/> adresinden edinilmiştir.
- Beninger, K., Fry, A., Jago, N., Lepps, H., Nass, L., and Silvester, H., (2014). "Research using social media; users' views", NatCen Social Research, ss. 1-40.
- Boudon, R., Borlandi, M., Cherkaui, M., and Valade, B., (2011). Sosyolojik Düşünce Sözlüğü, Bülent Arıbaş (Çev.), İletişim Yayınları, İstanbul.
- Bright, J. (2017). Big Social Science: Doing Big Data İn The Social Sciences. In Fielding, N., Lee, R., and Blank, G. (Eds.). The SAGE Handbook of Online Research Methods, 125-139. SAGE Publications Ltd., London. doi: 10.4135/9781473957992
- Bruns, A., (2015). "Making Sense of Society Through Social Media", Social Media + Society, Sayı. 1, No. 1, s. 1-2.
- Chan, W. W. L., Ma, W. W. K., (2014). "The influence of playfulness and subject involvement on focused attention when using social media", Journal of Communication and Education, Sayı. 1, No. 1, 16-27.
- Clement, J. (2019, Kasım). "Leading countries based on number of Twitter users as of October 2019", <https://www.statista.com/statistics/242606/number-of-active-twitter-users-in-selected-countries/> adresinden edinilmiştir.
- Clement, J., (2019). "Number of monthly active Twitter users worldwide from 1st quarter 2010 to 1st quarter 2019", <https://www.statista.com/statistics/282087/number-of-monthly-active-twitter-users/> adresinden edinilmiştir.
- Çelik, S., (2018). Büyük Veri, Gece Akademi
- Eraslan, L., Eser, D. Ç., (2015). Sosyal Medya Toplum Araştırma Sosyal Medya Sosyolojisine Giriş. Beta Yayıncılık, İstanbul.
- Freelon, D., (2014). "On the interpretation of digital trace data in communication and social computing research", J. Broadcast Electron Media, Sayı. 58, No. 1, ss. 59-75. Doi:10.1080/08838151.2013.875018
- Fruchter, M., (2009). "What Social Media Is and What Social Media Is Not Article title", <https://blog.louisgray.com/2009/01/what-social-media-is-and-what-social.html> adresinden edinilmiştir.
- Gürsakal, N., (2009). Sosyal Ağ Analizi, Dora Yayınları, Bursa.
- Holmberg, K., Thelwall, M., (2014). "Disciplinary differences in Twitter scholarly communication", Scientometrics, Sayı. 101, No. 2. Doi: 10.1007/s11192-014-1229-3.
- Jungherr, A., (2015). "Analyzing Political Communication with Digital Trace Data", The Role of Twitter Messages in Social Science Research. Doi:10.1007/978-3-319-20319-5. 27
- Kim, J. W. (2014). "Scan and click: The uses and gratifications of social recommendation systems", Computers in Human Behavior, Sayı. 33, ss. 184-191.
- Kitchin, R. (2014). The Data Revolution: Big Data, Open Data, Data İnfrastructures & Their Consequences, SAGE Publications Ltd., London.

- Lewis, L., (2019). “1 dakikada internette neler oluyor?”
<https://businessht.bloomberght.com/grafik/haber/2382145-1-dakikada-internette-neler-oluyor> adresinden edinilmiştir.
- Lupton, D. (2015). *Digital sociology*. Abingdon, Oxon: Routledge, 5.
- Marres, N., (2017). *Digital Sociology: The Reinvention of Social Research*, John Wiley & Sons.
- Marres, N., Weltevrede, E., (2013). “Scraping The Social? Issues in Live Social Research”, *Journal of Cultural Economy*, Sayı. 6, No. 3, ss. 313–335.
- Marres, N., Weltevrede, E., (2013). “Scraping the Social? Issues in live social research”, *Journal of Cultural Economy*, Sayı. 6, No. 3, ss. 313-335. Doi: 10.1080/17530350.2013.772070
- Mayer-Schönberger, V., Cukier, K., (2013). *Big data: A Revolution That Will Transform How We Live, Work, and Think*, Houghton Mifflin, New York.
- McCay-Peet, L., Quan-Haase, A., (2017). “What is social media and what questions can social media research help us answer”, *The SAGE Handbook of Social Media Research Methods*, s.19.
- McLuhan, M., (1962). *The Gutenberg Galaxy: The Making of Typographic Man*, University of Toronto Press. Toronto:
- Morris, T., (2009). *All a Twitter: A Personal and Professional Guide to Social Networking with Twitter*. Que Publishing.
- Otieno, D. O., Matoke, V. B., (2014). “Social media as tool for conducting academic research”, *International Journal of Advanced Research in Computer Science and Software Engineering*, Sayı. 4, No. 1, ss. 962–967.
- Puschmann, C., Ausserhofer, J., (2017). *Social Data APIs: Origin, Types, Issues*.
- Salganik, M. J., (2017). *Bit by Bit: Social Research in the Digital Age*. Princeton University Press, Princeton, NJ.
- Schäfer, M. T., Van Es, K. F., (2017). *The Datafied Society: Studying Culture Through Data*. Amsterdam University Press, Amsterdam.
- Smock, A. D., Ellison, N. B., Lampe, C., Wohn, D. Y., (2011). “Facebook as a toolkit: A uses and gratification approach to unbundling feature use”, *Computers in Human Behavior*, Sayı. 27, No. 6, ss. 2322–2329.
- Steinert-Threlkeld, Z., (2018). *Twitter as Data (Elements in Quantitative and Computational Methods for the Social Sciences)*. Cambridge University Press. Cambridge.
Doi:10.1017/9781108529327. 4-5-6
- Tate, R., (2009). *Twitter’s New Prompt: A Linguist Weighs*,
<https://gawker.com/5408768/twitters-new-prompt-a-linguist-weighs-in>, adresinden edinilmiştir.
- Timisi, N., (2018). *Teknoloji ve Toplum, Ağlar ve Aktörler, Dijital Kavramlar, Olanaklar, Deneyimler*. Kalkedon Yayıncılık, İstanbul.
- Tinati, R., Halford, S., Carr, L., Pope, C., (2014). “Big data: methodological challenges and approaches for sociological analysis”, *Sociology*, Sayı. 48, No. 4, ss. 663-681.
Doi:10.1177/0038038513511561.
- Uybadın, A., Mumcu, N., (2013). “Diyalog 2.0: Yeni Medyanın “Öteki ile Ben Diyaloguna Katkıları Üzerine Hermeneutik Bir Yaklaşım”. Özçetin, B., Göker, G., Bayraktutan, G., Sayımer, İ. ve Çomu, T. (Der.), *Yeni Medya Çalışmaları. Kuram, Yöntem, Uygulama ve Siyasa I. Ulusal Kongresi Kongre Kitabı* ss. 82-92, Alternatif Bilişim Derneği Genel Merkezi, Ankara.

- Uğur, A. (2013). “Cehennem Vaadi: Yeni Medya” Özçetin, B., Göker, G., Bayraktutan, G., Sayımer, İ. ve Çomu, T. (Der.), Yeni Medya Çalışmaları. Kuram, Yöntem, Uygulama ve Siyasa I. Ulusal Kongresi Kongre Kitabı s. 10. Alternatif Bilişim Derneği Genel Merkezi, Ankara.
- Weller, K., Burns, A., Burgess, J., Mahrt, M., Puschmann, C. (2016). Anders Olof Larsson, Hallvard Moe, Siyasette ve Seçimlerde Twitter: İskandinavya'nın Düşündürdükleri, Twitter ve Toplum, Epsilon Yayıncılık Ltd. Sti. İstanbul.
- Weller, K., Burns, A., Burgess, J., Mahrt, M., Puschmann, C., (2016). Twitter ve Toplum. Epsilon Yayıncılık Ltd. Sti. İstanbul.
- Williams, M. L., Burnap, P., Sloan, L., (2017). “Crime sensing with big data: The affordances and limitations of using open-source communications to estimate crime patterns”, The British Journal of Criminology, Sayı. 57, No. 2, ss. 320-340.
- Young, A. L., Quan-Haase, A., (2013). “Privacy protection strategies on Facebook”, Information, Communication & Society, Sayı. 16, No. 4, ss. 479–500.

Review Article

**Büyük Veri Çağında Sosyal Medya Verilerinin Sosyal Bilimler İçin Önemi;
Twitter Örneği**

*Importance of Social Media Data for Social Sciences in Big Data Age; The Case Of
Twitter*

Hüseyin Tuğrul OKTAY

Ankara Hacı Bayram Veli Üniversitesi Doktora Öğrencisi

htoktaya@gmail.com

<https://orcid.org/0000-0002-3656-7733>

Extensive Summary

In social networks, people with similar interests connect with each other such as chat, comment, community, collaboration, contribution. As social scientists, we are particularly interested in the communities formed in social networks and their interrelationships. Since social media networks are usually formed through chat and commentary, they are mediums where many virtual communities are formed through similarities and contrasts. It is the fact that every step you take is different from the real life, which can be described as analog which makes these channels digital. There are unique ways of analyzing these connections and data which are established in digital environment.

Digital environments are social environments and that is why the term “social media” is used instead of media. The features offered to users in social media environments feed this sociality through recycling such as being visible, receiving appreciation, receiving comments, and spreading sharing. These continuous interactions cause individuals in social media to establish a traceable network. The relationships that make up a network are informal social relationships, and the social network is an organized but non-institutionalized form. The social actor is not a living entity alone and sociology analyzes the networks formed by limiting the forms of social relations. These forms of relationship have now taken off the national framework with the development of communication tools and are now growing in accordance with the worldwide network definition (world wide web, www). The relations that have reached the network form are now developing according to the network in which they are located.

With the increasing use of online services, the capabilities of data and data storage, as well as the analysis capability, have increased. Thus, researchers have become increasingly interested in what digital monitoring data can say about human behavior patterns. These new data sources document human behavior in unprecedented detail and scale. The prevalence of analytical tools for Big Data presents enormous opportunities and challenges for researchers in the field of communication. Unlike user-generated text on Twitter, digital tracking data (the presence of online user events such as links, replies, and retweets) represents a new methodological limit for the field.

The increase in digital devices, online services and data sets documenting human behavior has led to research efforts in digital fields in various branches of social sciences. The three most common concepts associated with the analysis of digital data traces of human behavior are computational social sciences, digital methods, and big data. Each of these terms derives from a different research tradition and is linked to different approaches to research practice and research interest.

In social networks, people with similar interests connect with each other such as chat, comment, community, collaboration, contribution. As social scientists, we are particularly interested in the communities formed in social networks and their interrelationships. Since social media networks are usually formed through chat and commentary, they are mediums where many virtual communities are formed through similarities and contrasts. It is the fact that every step you take is different from the real life, which can be described as analog which makes these channels digital. There are unique ways of analyzing these connections and data which are established in digital environment.

Digital environments are social environments and that is why the term “social media yerine is used instead of media. The features offered to users in social media environments feed this sociality through recycling such as being visible, receiving appreciation, receiving comments, and spreading sharing. These continuous interactions cause individuals in social media to establish a traceable network. The relationships that make up a network are informal social relationships, and the social network is an organized but non-institutionalized form. The social actor is not a living entity alone and sociology analyzes the networks formed by limiting the forms of social relations. These forms of relationship have now taken off the national framework with the development of communication tools and are now growing in accordance with the worldwide network definition (world wide web, www). The relations that have reached the network form are now developing according to the network in which they are located.

With the increasing use of online services, the capabilities of data and data storage, as well as the analysis capability, have increased. Thus, researchers have become increasingly interested in what digital monitoring data can say about human behavior patterns. These new data sources document human behavior in unprecedented detail and scale. The prevalence of analytical tools for Big Data presents enormous opportunities and challenges for researchers in the field of communication. Unlike user-generated text on Twitter, digital tracking data (the presence of online user events such as links, replies, and retweets) represents a new methodological limit for the field.

The increase in digital devices, online services and data sets documenting human behavior has led to research efforts in digital fields in various branches of social sciences. The three most common concepts associated with the analysis of digital data traces of human behavior are computational social sciences, digital methods, and big data. Each of these terms derives from a different research tradition and is linked to different approaches to research practice and research interest.

The digital age has opened up new ways for people to communicate. These new forms of communication allow researchers to conduct innovative surveys and build mass collaboration with colleagues and the public. If you capture a picture of a horse, then you have a photo. If you capture 24 pictures of a horse per second, then you have a movie. A film is just a bunch of photos, but it is no longer a photograph. Researchers are in the process of making a change in the transition from photography to cinematography. However, this change does not mean that everything we have learned in the past should be ignored. When the principles of photography give information about cinematography, the principles of social research developed in the last 100 years will shed light on social research in the next 100 years. But change also means that we should not just continue to do the same thing. Rather, we must combine the approaches of the past with the capabilities of today and the future.

One of the most important terms associated with the potential to obtain information from digital data sources documenting human behavior is Big Data. The term "Big Data" is defined by seven key features: volume, Speed, Diversity, Exhausting, High resolution, Relational, Flexible.

Although the potential of Big Data for social research is enormous, most of these data sets remain in the hands of governments or private companies for reasons such as confidentiality or commercial sensitivity. An important exception is Twitter, a micro-blog site whose content is (almost entirely) public and accessible through Twitter's own API. Twitter is very interesting among social scientists with its relatively generous structure for data sharing.

The global aspect of Twitter's reach is important and easily understood by Marshall McLuhan's

concept of “global village”. McLuhan speaks of a interdependence that will enable the communication infrastructure in a globalized electronic environment to turn every part of the world into a global village. Twitter has likewise been associated with the strengthening of the global village. Twitter is thought to be one of the closest platforms to the Global Village where the Internet is foreseen. If anything happens anywhere in the world, someone will tweet it. Ultimately, social media, one of the most important sources of Big Data producers, and Twitter in particular, are considered to be a very important data source for social scientists, including reflection from all areas of life, since connections can be recorded, analyzed and evaluated at the community level.