



# Bir Arařtırma Kaynađı Olarak Arřivlenen Sosyal Medya Verilerinin Kullanımı

## Archived Social Media Data as a Research Resource

Aykut Kaya<sup>1</sup>



<sup>1</sup>Kütüphaneci, Bursa Uludađ Üniversitesi, Kütüphane ve Dokümantasyon Daire Başkanlığı, Bursa, Türkiye; Doktora Öğrencisi, İstanbul Üniversitesi, Sosyal Bilimler Enstitüsü, İstanbul, Türkiye

ORCID: A.K. 0000-0001-6477-5655

### Sorumlu yazar/Corresponding author:

Aykut Kaya,  
Kütüphaneci, Bursa Uludađ Üniversitesi, Kütüphane ve Dokümantasyon Daire Başkanlığı, Bursa, Türkiye; Doktora Öğrencisi, İstanbul Üniversitesi, Sosyal Bilimler Enstitüsü, İstanbul, Türkiye  
E-posta: aykutkaya@uludag.edu.tr

**Başvuru/Submitted:** 11.10.2021

**Revizyon Talebi/Revision Requested:** 01.11.2021

**Son Revizyon/Last Revision Received:** 10.11.2021

**Kabul/Accepted:** 12.11.2021

**Online Yayın/Published Online:** 09.12.2021

**Atf/Citation:** Kaya, A. (2021). Bir arařtırma kaynađı olarak arřivlenen sosyal medya verilerinin kullanımı. *Bilgi ve Belge Arařtırmaları Dergisi*, 16, 49-79.

<http://doi.org/10.26650/bba.2021.16.1008002>

### ÖZ

Günümüzde sosyal medya, bireysel olarak kullanımının yanında devlet kurumlarından medya kuruluşlarına, özel sektörden akademik arařtırmalara kadar bir dizi alanda kullanılmaktadır. Sosyal medyanın bu çok sayıda alanda kullanımı ile birlikte, arařtırmacıların çalışmalarında kullanabilecekleri değerli yeni veriler ortaya çıkmaktadır. "Sosyal medya verileri" olarak adlandırılan bu değerli veri türü daha önce görülmemiş bir nitelikte bir arařtırma kaynađı yaratmıştır. *Uygulama Programlama Arayüzlerinin (Application Programming Interface, API)* ortaya çıkması ve sosyal medya platformlarının API'ler etrafında şekillenmesi, arařtırmacıları sosyal medya platformlarından veri toplamanın ve verileri arřivlemenin en etkili yollarından biri olarak platform API'lerini kullanmaya yöneltmiştir. Bilgi ve belge yönetimi alanında değeri yeni fark edilen sosyal medya verileri üzerine pek fazla çalışma bulunmamaktadır. Bu çalışmada; sosyal medya verileri kavramı ile sosyal medya platformlarının API'leri aracılığıyla toplanan arřivlenen sosyal medya verilerinin arařtırmalarda yeni bir kaynak türü olarak önemine vurgu yapılmış, bu verilerin hâlihazırda kullanıldığı akademik arařtırma alanlarına işaret edilmiştir. Çalışma kapsamında ayrıca sosyal medya büyük verisi ve sosyal medya veri madenciliđi kavramları da açıklanmıştır. Literatüre dayalı nitel arařtırma yönteminin kullanıldığı çalışmada elde edilen verilerin analizi için de doküman analizi tekniğinden yararlanılmıştır. **Anahtar kelimeler:** Sosyal medya, sosyal medya verisi, API, veri kaynađı, sosyal medya arařtırması

### ABSTRACT

Social media is used today in numerous contexts, from government institutions to media organizations and from the private sector to academic research, as well as by individuals. The use of social media across a wide range of fields generates valuable new data that researchers can use in their studies. This so-called "social media data" represents an unprecedented research resource. The emergence of *Application Programming Interfaces (APIs)* and the shaping of social media platforms around them have led researchers to use platform APIs as a highly effective way to collect and archive social media data. Few studies exist on social media data, the value of which has only recently been recognized in the field of information and records management. This paper emphasizes the importance of social media data collected and archived through social media platform APIs as a new source for research. It also highlights areas of academic research where these data are currently used. Furthermore, the paper explains the concepts of big data and data mining in the context of social media. A literature-based qualitative research method was used to obtain data for the study, and these data were analyzed through the document analysis technique.

**Keywords:** Social media, social media data, API, data source, social media research



## EXTENDED ABSTRACT

Although social media is generally regarded as the name given to platforms such as Facebook, Instagram, Twitter, YouTube, Reddit, and Pinterest, various definitions of the term have been offered by different studies. Broadly, social media can be defined as web/mobile based applications or media forms that enable users to come together with other users, share information, communicate, and form a community, and in which content is user-created.

In addition to its individual use, social media is employed in a range of contexts such as public institutions, private companies, non-profit organizations, and educational activities. Organizations may also use social media for data mining to better understand attitudes and trends and to predict future behavior. With the increasing number of social media users and its use in the many areas described above, social media data has started to be produced very quickly. Social media contains a significant amount of natural data on many issues related to today's society, including various aspects of human behavior and corporate activity, from social movements to political views and preferences and modes of communication. Consequently, a new type of data has emerged that researchers can use in their studies. This valuable "social media data" has become a research resource of unprecedented scale. Furthermore, as social media data has been accepted as a type of big data because of the many features they share, it may also called "social big data" or "social media big data."

In previous scholarly research, traditional methods such as library, archive, and internet searches, literature reviews, surveys, interviews, observations, and scales were used to obtain data. Over time, however, traditional data acquisition methods have become increasingly difficult and expensive, forcing researchers to seek new data sources. Once social media appeared and its potential as a data source was recognized, new methods of research to exploit it quickly emerged. Although not initially created and thought of as a data source, the development of social media over time has made it a source and object of study in recent years. This has generated new data types that researchers now commonly use in academic research to address new problems and find appropriate solutions.

Social media data can be used for a wide variety of purposes, such as measuring and analyzing information produced through social media. It also has the feature of being both a common social memory and a documentation tool for transmitting this memory to the future. What is important at this point is to identify the type of tool that researchers will require to use the memory-based information in question. APIs have been employed as the most important tool to enable researchers to collect and evaluate social media data in a meaningful and reliable way. The emergence of APIs and the shaping of social media platforms around them have led researchers to use platform APIs as a highly effective way to collect and archive data from social media platforms. Today, academic studies based on data obtained through social media platform APIs have become widespread.

The fundamental starting point of this study is that there has been little work in the field of information and records management on social media data, especially given the growing recognition of its importance as a resource. Accordingly, this paper considers the concept of social media data and the use of social media data collected and archived through social media platform APIs as a new source for research. Moreover, it examines the academic research areas (e.g., crisis moments such as epidemics and natural disasters, health, psychology, linguistics, communication studies, marketing, political activities, and geography) in which these data are being used. The paper also explains the concepts of big data and data mining in the context of social media. A literature-based qualitative research method was used to obtain data for the study, and these data were then subjected to the document analysis technique.

## GİRİŞ

Önceden bilimsel araştırmalardaki veri elde etme yöntemleri, kütüphane, arşiv, internet taramaları, literatür incelemeleri, anketler, görüşmeler, gözlemler ve ölçekler gibi çabalara dayanmaktaydı. Ancak, zaman geçtikçe, geleneksel veri elde etme yöntemlerinin daha zor ve daha pahalı hale gelmesi ile araştırmacılar yeni veri kaynakları arayışına girişmiştir. Bu arayışlara bağlı olarak 2000’li yılların başlarında sosyal medyanın (bloglar, sosyal ağlar, mikrobloglar gibi) ortaya çıkışı ve bir iletişim aracı olarak yaygın kullanılışıyla birlikte araştırmacılar, veri elde etme ve bu verilerin analizi için önemli bir fırsat görmüştür (Fondren ve McCune, 2018, s. 38).

Sosyal medya ve sosyal medya platformları, insanların ve kurumların kendilerini ifade edebildikleri, yapılanları/faaliyetleri duyurdukları, her türden güncel sorunu tartıştıkları, özgün değerlendirmeler yapabildikleri daha da önemlisi daha önce hiçbir şekilde kullanılmamış olan yeni bir veri kaynağının ortaya çıkmasını sağlamıştır. Artık sosyal medya ve sosyal medya platformları, sadece bilginin yayıldığı, toplumsallaştığı, sosyalleşme ve günlük yaşamın tüm yönlerini yansıtan alanlar değil, aynı zamanda veri elde etmek için sıklıkla kullanılan araçlara dönüşmüştür (Zeller, 2017, s. 388).

Başlangıçta bir veri kaynağı olarak oluşturulmaması ve düşünülmemesine rağmen, sosyal medyanın zaman içinde geçirdiği gelişim sosyal medyayı son yıllarda akademik araştırmalarda, araştırmacıların yeni soru(n)ları ele almak ve bu soru(n)lara cevap/çözüm bulabilmek için yaygın olarak kullandığı yeni veri türleri üreten bir kaynak ve çalışma nesnesi haline getirmiştir. Günümüzde geldiğimiz noktada, internet araştırmaları, -özellikle sosyal medya araştırmaları- çok sayıda kullanıcı verisine çevrimiçi erişim sağlayan teknolojik etmenlerden faydalanmıştır/faydalanmaktadır.

Araştırmacılar, *Twitter*, *Facebook* ve *Instagram* gibi sosyal medya platformlarının *Uygulama Programlama Arayüzü (Application Programming Interface, API)*<sup>1</sup> aracılığıyla sosyal medya verilerinin toplanmasını ve arşivlenmesini sağlamıştır. Araştırmacıların sosyal medya verilerinden oluşan veri kümelerini toplamaları çevrimiçi faaliyetler ile ilgili “Büyük Veri” çalışmalarının yaygınlaşmasına zemin hazırlamıştır (Croeser ve Highfield, 2019, s. 670). Günümüzde sosyal medya kullanımı ile ortaya çıkan ve “sosyal büyük veri” veya “sosyal medya büyük verisi” olarak adlandırılan bu kavramlar bilim insanlarının ve uygulayıcıların erişebileceği ölçek ve veri türlerinde önemli bir artış sağlamıştır. Sosyal medya, toplumsal hareketlerden, politik görüşler ve tercihler ile iletişim biçimlerine kadar insan davranışının ve kurum faaliyetlerinin çeşitli yönlerini de kapsayan, günümüz toplumuyla ilgili pek çok konuda önemli miktarda doğal veri içermektedir. Sosyal medya, elde edilmesi zor ve zaman alan bu türden verilere ulaşmak isteyen araştırmacılar için hem benzeri görülmemiş fırsatlar hem de zorluklar sunmaktadır.

1 API bir yazılım ürünü veya uygulama ile bir web sunucusu arasındaki veri aktarımını sağlayan programlama kümesi, arayüzdür.

Sosyal medya verileri, bu ortam üzerinden üretilen bilginin ölçülmesi ve analiz edilmesi gibi çeşitli amaçlar için kullanılmaktadır. Yine sosyal medya verileri ortak toplumsal bellek niteliğindedir ve bu bellek günümüz dünyasının geleceğe aktarılması için bir belgeleme (toplumsal dokümantasyon) aracı olma özelliğine sahiptir. Bu noktada önemli olan, söz konusu belleğe dayalı bilginin kullanılabilmesi için araştırmacıların ne tür bir araca ihtiyaç duyacaklarıdır. Araştırmacıların sosyal medya verilerini anlamlı ve güvenilir bir şekilde toplayıp değerlendirilmelerini sağlayacak en önemli araç olarak API'ler kullanıma sunulmuştur. API'lerin ortaya çıkması ve sosyal medya platformlarının API'ler etrafında şekillenmesi, araştırmacıları sosyal medya platformlarından veri toplamanın ve verileri arşivlemenin en etkili yollarından biri olarak platform API'lerini kullanmaya yöneltmiştir.

Çalışmada, sosyal medya verileri kavramı ile sosyal medya platformlarının API'leri aracılığıyla toplanıp arşivlenen sosyal medya verilerinin araştırmalarda kullanımına değinilmiş, bu verilerin kullanıldığı akademik araştırma alanları (salgın ve doğal afet gibi kriz anları, sağlık, psikoloji, dilbilim, iletişim çalışmaları, pazarlama, siyasal faaliyetler ve coğrafya) incelenmiştir. Çalışma kapsamında ayrıca sosyal medya büyük verisi ve sosyal medyada veri madenciliği kavramları da açıklanmıştır. Değeri son zamanlarda daha çok fark edilen bir kaynak olarak sosyal medya verileri üzerine bilgi ve belge yönetimi alanında şimdiye kadar pek fazla çalışılmamış olması bu araştırmanın temel çıkış noktasını ve özgün yanını oluşturmaktadır. Bu bağlamda çalışma, arşivlenen sosyal medya verilerinin araştırma kaynağı olarak nasıl ve hangi araştırma alanlarında kullanılabileceğinin incelenmesini hedeflemektedir. Çalışmanın temel kaynakları, araştırma makaleleri, bilimsel kitaplar ve tezlerdir. Çalışmanın verileri literatür araştırması, kütüphane ve internet taraması ile kurumsal veri tabanları yoluyla elde edilmiştir. Literatüre dayalı nitel araştırma yönteminin kullanıldığı çalışmada elde edilen verilerin analizi için doküman analizi tekniği kullanılmıştır.

## 1. Kavramsal Çerçeve

Çalışma kapsamında öncelikle bazı temel kavramların açıklanmasında yarar vardır. Bu kavramlar veri, sosyal medya, sosyal medya verisi, yapılandırılmış veri ve yapılandırılmamış veri şeklinde belirtilebilir. Veri kavramı hakkında literatürde ve teknoloji ile ilişkili terimleri açıklayan/inceleyen terminoloji sözlüklerinde pek çok tanım yer almaktadır. Türkçe Bilim Terimleri Sözlüğü'nde veri, "sonuç çıkarmak, çözümlenme yoluyla çıkarsamalar yapmak ya da bir incelemeyi sürdürmek için sormaca, gözlem, deney ya da ölçme gibi tekniklerle toplanmış, ancak işlenmemiş nicel ya da nitel değerler kümesi" olarak tanımlanmaktadır ("TÜBA: Türkçe Bilim Terimleri Sözlüğü", y.y.). Bir başka değerlendirmeye göre veriler, temel bir araç ve araştırma ürünü; bilgiyi elde etmek, saklamak ve yaymak için kullanılabilen kaynaklardır (Bishop ve ark., 2019, s. 4).

Sosyal medya, Kaplan ve Haenlein (2010, s. 61)'in yirmi dört binden fazla atıf alan makalesinde, "her yerde erişilebilir bir şekilde kullanıcı tarafından oluşturulan içeriğin

üretilmesine, erişilmesine ve paylaşılmasına olanak tanıyan web tabanlı ve mobil tabanlı internet uygulamaları” olarak tanımlanmıştır. Günümüzde ise sosyal medya; *Facebook, Twitter, Instagram, Youtube, Reddit, Pinterest* vb. platformlara verilen genel ad olarak bilinmektedir.

Sosyal medya verileri, sosyal medya kullanıcılarının oluşturduğu veya paylaştığı içerik (fotoğraflar, görseller, metinler, yorumlar vb.), bağlantılı içerik, etkileşim verileri, zaman damgaları, hesap bilgileri ve tıklama akışları gibi sistem tarafından oluşturulan üstverilerin bir birleşimi olup genellikle büyük boyutlu, hızlı oluşturulabilen ve çabuk kaybolabilen yani dinamik ve geçici nesnelere (Hemphill, Hedstrom ve Leonard, 2020, s. 106). Sosyal medya verileri, sosyal medya kullanıcılarının sosyal medyadaki her türlü etkinliğine ilişkin veriler olup kullanıcıların, sosyal medyayı kullandıkları sırada oluşturduğu veya paylaştığı bilgileri ifade etmekte ve genellikle kişisel olarak nitelendirilebilen veriler olarak bilinmektedir (ESOMAR [European Society for Opinion and Marketing Research], 2011, s. 3). Sosyal medya içeriği, sayısız kanaldaki metinler, hikâyeler, yorumlar, hashtagler<sup>2</sup>, videolar, gönderiler, fotoğraflar ve tweetler gibi çeşitli verilerden oluşmaktadır. Bu sebeple, sosyal medya verileri çoğunlukla kişisel veri olarak değerlendirilmektedir (“Social media archiving in regulated industries: why it matters in 2021”, 2020). Birçok kullanıcı nerede eğitim aldığı, yaşadığı veya çalıştığı, hobileri gibi bilgilerini açıkça paylaştığından sosyal medya platformlarında bunlar gibi büyük miktarda kişisel veri bulunmakta, bu veriler çok bilgilendirici olabilmektedir (Qi ve Edgar-Nevill, 2011, s. 75).

Sosyal medya verilerinin önemi kendine özgü doğasından kaynaklanmaktadır. Bu veriler, kişisel bilgilerden ya da duygulardan oluştuğundan potansiyel olarak insanın duygusal durumunun sonucunu önemli ölçüde etkileyebilecek olaylar için önem taşımaktadır (Huang ve Xu, 2014, s. 267). Her ne kadar kişisel veriler içerse de, sosyal medya verileri birer araştırma kaynağı olarak kullanılabilir. Bahsedildiği üzere, sosyal medya platformları; metinler, videolar, fotoğraflar, sesler ve coğrafi konumlar gibi birçok veri türüne sahiptir.

Genel olarak sosyal medya verileri, yapılandırılmış (structured) veya yapılandırılmamış (unstructured) veriler şeklinde gruplanabilmektedir. Yapılandırılmış veriler, “erişimi ve incelemesi geleneksel yöntemlerle en kolay ve hızlı biçimde yapılabilen”, yapılandırılmamış veriler ise “herhangi bir biçimsel kurala bağlı olmayan, büyük ölçüde metinsel içerik taşıyan, işlenip incelenebilmesi için metin işleme yöntemleri gereken” veriler olarak tanımlanmaktadır. *Twitter*’da gönderilen tweetler yapılandırılmamış metinsel verilere; arkadaş/takipçi bilgileri ise yapılandırılmış verilere örnek oluşturmaktadır (Stieglitz, Mirbabaie, Ross ve Neuberger, 2018, s. 156; Şeker Elmas, 2019, s. 24).

2 Çoğunlukla sosyal medya platformlarında kullanılan, “#” sembolü ile işaretlenen ve beraberinde yazıldığı metni sınıflandıran kelime ya da kelime öbeklerine verilen ad.

Sosyal medya uygulamaları, esas olarak dış bağlantılardan ve çeşitli medya türlerinden oluşmaktadır. Bu uygulamalar, genellikle oldukça karmaşık profillerle ilişkili olan kullanıcılar arasında bilgi alışverişini kolaylaştırmak için işlev gördüğünden, aynı zamanda önemli miktarda üstveri üretmektedir. Söz konusu üstverilerin üretilmesi, veri kümelerinin coğrafi konum, tarih, saat, cinsiyet, meslek, yaş, milliyet ve bir dizi diğer türev kullanıcı ögesiyle potansiyel olarak zenginleştirilebileceği anlamına gelmektedir (Sloan, Morgan, Burnap ve Williams, 2015).

Sosyal medya kaynaklarından elde edilen veriler değişiklik gösterebilir, ancak genellikle dört sınıfa ayrılır. Bunlar; *kullanıcı bilgileri*, *etkinlik*, *ağ* ve *içerik*'tir. *Kullanıcı bilgileri*, profil verileri olarak da bilinmektedir. Kullanıcı profil verileri; kullanıcının adı, doğum yeri ve tarihi, e-posta adresi, iş ve eğitim bilgisi, yaşadığı şehir gibi sosyal medya hesabına sahip kişi hakkındaki bilgileri göstermektedir. *Etkinlik*, sosyal medya kullanıcılarının oluşturduğu içeriklerin IP adresi<sup>3</sup>, zaman, tarayıcı bilgisi, oturumların uzunluğu gibi her faaliyet için sosyal medya platformlarının kayıt altına alındığı ve otomatik olarak tutulan verilerdir. *Ağ* kısmı, takipçi ya da arkadaş adı ve sayısı gibi bilgileri içermektedir. Benzer ilgi alanlarına sahip kişilerin belirlenmesi, kuvvetli bağlantıların bulunması, özel bilgilerin elde edilmesi gibi konular da bu çerçevede içindedir. *İçerik*, kullanıcılar tarafından oluşturulan tweetler, beğeniler, yorumlar, videolar ve fotoğraflar gibi materyalleri tanımlamaktadır (Arshad, Jantan ve Omolara, 2019, ss. 127-128; Gürsakal, 2009, s. 233; Social Media Research Group, 2016, s. 7).

Sosyal medya verilerinin temel ayırt edici özellikleri; anlık, doğal olarak daha önce toplanması zor olan yeni tipte öznel veriler sağlaması ve nüfus ölçeğinde oluşturulmalarıdır. Bu özellikler, -dijital ortamda üretilen hatta sonradan dijital ortama aktarılan veriler gibi- sosyal medyayı verilerini geleneksel araştırma yöntemleri ile elde edilen materyallerden ayırmaktadır. Bunlar, özellikle araştırmacıların, geleneksel araştırma verilerinde karşılına çıkan bazı kısıtlamaları aşmalarına da imkân sağlamaktadır (Procter, Voss ve Lvov, 2015, s. 484). Çünkü sosyal medyada kullanıcı tarafından üretilen veriler, makine tarafından okunabilir biçimde yakalandığından bunların işlenmesi hızlı ve kolaydır (Fält, 2016).

Sosyal medya verileri büyük ölçüde piyasa güçlerine, giderek artan şekilde sahte kullanıcı hesaplarına ve bot hesaplara bağlı olan ticarileştirilmiş platformlarda bulunmaktadır. Ayrıca, veriler platformlardan elde edildiğinde, standartlaştırılmış değildir ve çeşitli biçimlerde (çoğunlukla JSON<sup>4</sup> veya XML<sup>5</sup> gibi) olabilmektedir (Acker ve Kriesberg, 2017, s. 2). Doğası gereği dinamik olan sosyal medya verileri çoğunlukla uzaktan barındırılmaktadır ve iş birliğine dayanan bir yapıya sahiptir. Bu verilere genellikle, sosyal medya platformlarının API'leriyle,

3 IP (Internet Protokol) adresi, Internet'te bağlanan her cihaza verilen eşsiz numara serisidir.

4 JSON (JavaScript Object Notation), JavaScript'ten türetilen insan tarafından okunabilir, veri değişimi için tasarlanmış metin tabanlı bir bilgi saklama formatıdır.

5 XML (Extensible Markup Language), internetteki veri alışverişini hızlı ve pratik bir şekilde gerçekleştirmeyi sağlayan ve platformlar arasındaki iletişimi ayarlayan bir işaretleme dilidir.

platform sağlayıcılarla müzakere edilerek belirlenen erişim yöntemleriyle ve üçüncü taraf veri satıcıları aracılığıyla erişilmektedir (Hemphill ve ark., 2020, s. 105; “Strategies for managing social media information”, y.y.).

## 2. Sosyal Medya Verilerinin Kullanım Alanları

Günümüzde mevcut olan en özgün sosyal veri kaynağı, kullanıcılar tarafından sosyal medyada oluşturulan içeriktir. İçeriğin kullanıcı tarafından oluşturulması, bireyler arasındaki deneyim ve iş birliğinin paylaşılması araştırmacılara çok çeşitli konularda araştırabilecekleri büyük miktarda veri sunmakta ve bu sayede söz konusu verilerden pek çok alanda yararlanılmaktadır (Young, 2013, s. 429). Örneğin; sosyal medya platformları son yıllarda iletişim faaliyetleri için potansiyel bir veri kaynağı haline gelmiştir. Sosyal medya verileri, iletişimdeki kötü tanımla sonuçlanabilecek yeni eğilimleri ve tehlikeleri belirlemek için yararlı olabilmektedir.

İşletmeler kitle iletişimi hakkında bilgi edinmek için kullanıcıların ürettiği sosyal medya verilerini kullanmaktadır (Çelik, 2018, s. 461). Siyasi iletişim alanında da sosyal medya verilerinden sıkça faydalanılmaktadır. Devlet kurumları ve siyasi parti yetkilileri sosyal medyayı hem seçmenleri ile iletişim kurmak hem de büyük kitlelere hızlı ve kolay yoldan ulaşmak için bir araç olarak kullanmaktadır. Hükümetler ve siyasi partiler, sosyal medyadan veri analizleri yaparak, sonraki seçim dönemi için yararlı stratejiler oluşturmaktadır. Kriz iletişimi, sosyal medya verilerinin etkin olarak kullanıldığı iletişim alanlarından bir diğeridir. Sosyal medya; doğal afetler, terör olayları vb. kriz anlarında afet ve acil durum yönetimi kurumları tarafından, krizden etkilenen insanlara ilgili krizin durumu veya insanların kriz anında nasıl hareket etmeleri gerektiği konusunda bilgilendirmek için de kullanılmaktadır. Örneğin, sosyal medya verilerinden krizin meydana geldiği yerin konumunu belirlemek için de yararlanılabilmektedir (Stieglitz ve ark., 2018, ss. 157-158).

Sosyal medya, devlet kurumlarının oluşturduğu değerli kamu verilerinin üretilmesinin yanında vatandaşlardan geri bildirim alınmasını da sağlamaktadır. Kurumlar, sosyal medya platformlarını kullanan vatandaşlarından toplanan bilgileri ve verileri analiz etmektedir. Devlet kurumları ihtiyaç duydukları sosyal medya verilerini bir *Twitter* anketi ya da *Facebook* gönderisi vasıtasıyla gerçek zamanlı olarak elde edebilmektedir (van Rooyen, y.y.).

Hem ticari hem de kâr amacı gütmeyen kuruluşlar da çeşitli amaçlar için sosyal medya verilerini kullanmaktadır. Örneğin; ticari kuruluşlar, sosyal medyayı müşterileriyle bir iletişim kanalı olarak kullanmasının yanında sosyal medya verilerinden müşterilerine sunacağı hizmet ve karar verme süreçlerini desteklemek amacıyla da yararlanmaktadır (Stieglitz ve ark., 2018, s. 157). Sosyal medya verilerini ilk kez kullanan şirketler finans ve perakende alanından olmuştur. Perakende şirketleri, müşteri hizmetlerinin yanında reklam ve pazarlama stratejileri geliştirmek, marka tanınırlığını ölçmek ve sahtekârlığı tespit etmek gibi amaçlar için sosyal



medya verilerini kullanırken, ticarete sosyal medyadaki haber verilerinden, finans alanında ise sosyal medya verilerinden piyasa duyarlılığını ölçmek için yararlanılmaktadır (Batrinca ve Treleaven, 2015, s. 91). Ticari kuruluşlar, nüfus grupları ve potansiyel müşterileri hakkında veri toplamak ve bu verileri analiz etmek için sosyal medya verilerini giderek daha fazla kullanmaktadır (Thomson, 2017, s. 2). Örneğin, turizm ve pazarlama alanlarında sosyal medya verileri coğrafi konum belirlemek için kullanılmaktadır (Çelik, 2018, s. 463).

Sosyal medya verileri potansiyel olarak dünyanın her yerindeki kullanıcıların davranışlarını ifade ettiklerinden çevrimiçi davranışın çoğunlukla yeni bir biçimi olarak kabul edilmektedir. Araştırmacılar, sosyal medya verilerini hem çevrimiçi tartışmalardan veri kazıma veya çıkarma<sup>6</sup>, çevrimiçi görüşme gibi çalışmalarında hem de birincil veya ikincil kaynak olarak bilimsel yayınlarda kullanmaktadır (Acker ve Kriesberg, 2017, s. 1; Beninger ve ark., 2014, s. 1). Araştırmacılar, önceden verileri literatür incelemesi, kütüphane, arşiv, internet taramaları, anketler, görüşmeler ve gözlem gibi çeşitli yöntemlerle elde ederken, araştırma verilerinin artık belli bir kısma, yoğunlukla sosyal medya platformları üzerinden erişilebilmektedir (Townsend ve Wallace, 2016, s. 4). Anketler ve görüşmeler gibi geleneksel veri kaynaklarının aksine, sosyal medya verileri web tabanlı sosyal medya platformları aracılığıyla makine tarafından okunabilen bir formatta oluşmaktadır. Bu veriler, insan eliyle hatta çoğunlukla bizzat verinin sahibi tarafından (birinci elden) doğrudan üretildiğinden insan davranışı ve sosyal etkileşim hakkında yeni keşiflere yol açabilecek karmaşık veri analizlerini kolaylaştırmıştır (Thomson ve Kilbride, 2015, s. 262). Zaman alıcı anketler veya röportaj araçları olmadan, sosyal medya verileri, araştırmacıların siyasi olaylar üzerine tartışmalar gibi önceden gözlemlenmesi zor olan olayları yerinde incelemesine de olanak tanımaktadır (Kinder-Kurlanda, 2019, ss. 105-106). Geçmişte veri azlığından veya yokluğundan şikâyet eden araştırmacılar, bugün büyük miktardaki sosyal medya verilerini ne şekilde elde edebileceklerini ve bunları nasıl kullanabileceklerini tartışmaktadır (Gürsakal, 2009, s. 36). Bu tartışmalardan biri de veri madenciliği üzerinedir.

### 3. Sosyal Medya Veri Madenciliği

Sosyal medya verilerinin sıklıkla kullanıldığı alanlardan biri olan veri madenciliği, yani eldeki büyük miktardaki verilerden ilk bakışta çok net anlaşılabilen ancak potansiyel olarak kullanmaya elverişli olan bilginin çıkarılmasıdır (Şimşek, 2012, s. 4). Veri madenciliği tekniği; yeni bilgi üretmek için önceden var olan büyük veri kümelerini ölçülebilir yöntemlerle inceleyen, belirli durum veya eğilimleri keşfetmek, bunlardan bir sonuç elde etmek için çok sık kullanılan bir tekniktir (Beninger ve ark., 2014, s. 5; “Data Mining VS Data Extraction: What’s the Difference?”, 2021). Veri madenciliği, müşterileri sınıflandırmaya, ürün satışlarını artırmaya, dolandırıcılığı tespit etmeye ve daha pek çok şeye yardımcı olabildiğinden 1990’lı yıllardan itibaren yaygınlaşmış ve daha sonra ulaşım, finans, sağlık, telekomünikasyon gibi

6 Çıkarılması istenen web sitesi ya da sosyal medya platformundaki değerli bilgileri bilgisayardaki kayıtlı bir yerel dosyaya veya elektronik tabloya aktarma tekniği veya işlemi olarak bilinmektedir.

çeşitli alanlardan birçok şirket, verilerden yararlanmak için veri madenciliği tekniklerini kullanmaya başlamıştır (“Data Mining VS Data Extraction: What’s the Difference? ”, 2021).

Organizasyonlar sürekli olarak veri üretmekte ve bunları yine sürekli olarak incelemektedir. Geçen zamanla beraber, ürettikleri bu verilere sosyal medya verileri de dâhil olmuş sosyal medya platformlarında veri madenciliği önemli hâle gelmiştir (Çelik, 2018, s. 462).

Sosyal medya veri madenciliği, sosyal medya platformlarında yer alan çeşitli bilgilerdeki fikirlerin anlamsal açıdan ortaya çıkarılması olarak tanımlanmaktadır (Çelik, 2018, s. 459). Sosyal medya veri madenciliği, kullanıcılar, şirketler ve müşteriler için yararlı olabilecek eyleme geçirilebilir kalıpları sosyal medyadan çıkarma potansiyeline sahiptir. Sosyal medya veri madenciliği geleceğe yönelik tahminler yapılmasının yanında çevrimiçi davranışlar, satın alma, bireyler arasındaki bağlantılar gibi ilk anda anlaşılabilen ham sosyal medya verilerinin incelenerek analiz edilmesine yardımcı olabilmektedir (Pushpam ve Jayanthi, 2017, s. 149). Sosyal medya veri madenciliği, sosyal medya, sosyal medya ağ analizi ve veri madenciliği alanlarını birleştirmektedir (Ertam, 2017, ss. 67-68).

Tüm bunlarla birlikte veri madenciliğinin bir alt dalı olarak oluşturulan sosyal medya veri madenciliği, sosyal medya araştırmaları verilen yeni bir araştırma alanının ortaya çıkmasını sağlamıştır. Sosyal medya veri madenciliği, çok farklı disiplinlerden araştırmacıların sosyal medya araştırmalarına önemli katkılarda bulunabileceği, akademik çalışmalarda önemi yeni yeni fark edilen bir alandır (Gundecha ve Liu, 2012, s. 1). Bu teknik, sosyal medya araştırmacıları için kısa bir zamanda oldukça güvenilir veri sağlayabilecek önemdedir. Araştırmacılar son yıllarda sosyal medya platformlarından veri madenciliği yöntemleri ile büyük miktarlarda veri elde etmişler ve bunları çalışmalarında kullanmışlardır (Beninger ve ark., 2014, s. 5). Sosyal medya veri madenciliği, araştırmacıların sosyal medya kullanımı sonucu ortaya çıkan yeni durumları anlamalarına yardımcı olabileceği gibi, daha iyi hizmetler sunmak ve yaratıcı beklentiler oluşturmak için iş zekâsı geliştirmelerine de yardımcı olabilir (Gundecha ve Liu, 2012, s. 1). Sosyal medya platformlarından duygu veya algı analizi, metin madenciliği ve kümeleme algoritması gibi çeşitli şekillerde veri madenciliği yapılabilmektedir (Çelik, 2018, ss. 462-463). Hem sosyal hem de akademik açıdan, sosyal medyadaki modern bağlantıları ortaya çıkarmanın veya gelecekteki çeşitli durumları tahmin etmenin bir yolu olarak sosyal medya veri madenciliği kullanımı zaman geçtikçe yaygınlaşmaktadır (Taylor ve Pagliari, 2018, s. 3).

#### **4. Sosyal Büyük Veri/Sosyal Medya Büyük Verileri**

Geleceğe yönelik tahminler üretmek ve insan davranışlarını anlama ihtiyaçları için büyük miktarlarda ve sayısız şekillerde bulunan veri kümelerini tanımlayan bir terim olan “Büyük Veri (Big Data)” kavramı; sistem kayıtları, kullanım istatistikleri ve iklim veya trafik sinyalizasyon verileri gibi farklı alanlardaki yığın verilerden meydana gelmektedir. Bu veriler, ilk bakıldığında

değersiz gibi görünse de zaman geçtikçe önemi anlaşılmış ve günümüzde oldukça kullanılır hale gelmiştir (Caplan ve Purser, 2019, s. 418; Çelik, 2018, s. 454). Tüfekçi (2014, s. 505), büyük verilerin, özellikle insan davranışını içeren birçok alan araştırmasında kullanıldığını söylemektedir.

Birbiriyle ilişkisiz ve yönetilmesi zor olan büyük miktardaki veriyi ifade eden büyük verilerin bilinen dört özelliği bulunmaktadır. Bunlar; hacim, hız, çeşitlilik ve doğruluk ile karakterize edilmektedir (Lynn ve ark., 2015, s. 1):

1. Hacim (Volume): Verinin miktarını belirtmektedir.
2. Hız (Velocity): Verinin hızlı üretildiği/alındığı sıradaki veya akışındaki orandır.
3. Çeşitlilik (Variety): Farklı biçimlerdeki veri kümelerinin bir koleksiyonudur. Bir diğer ifade ile çeşitlilik; video, fotoğraf ve metin gibi yapılandırılmamış ya da yarı yapılandırılmış veri türlerini açıklamaktadır.
4. Doğruluk (Veracity): Verilerin kesinlik ve güvenilirlik derecesi olarak tanımlanabilir. Veriler doğru ve güvenilir olmazsa kuruluşlar tarafından alınan kararlar anlamlı olmayacaktır (Al-Hajjar, Jaafar, Al-Jadaan ve Alnutaifi, 2015, s. 301-302; Çelik, 2018, s. 455).

Sosyal medya verileri, büyük verinin bir türü ve büyük verilere atfedilen birçok kategoriye uygun olarak kabul edilmektedir (Caplan ve Purser, 2019, s. 418; Derczynski, Yang ve Jensen, 2013, s. 137). Bu verilere “sosyal büyük veri” veya “sosyal medya büyük verisi” denmektedir. Sosyal medya verilerinin, çok farklı kaynaklardan elde edilenleri içeren “büyük veri”nin birçok özelliğini paylaştığı söylenebilir. Bu nedenle sosyal medya verilerini de içine alan çalışmalar büyük verilerle ilgili yayınlarda yaygın olarak bulunmakta ve sosyal medya verilerini kullanan araştırmacılar da sıklıkla büyük veri literatürüne atıf yapmaktadırlar (Stieglitz ve ark., 2018, s. 158).

Büyük veri, başlangıçta süper bilgisayarlar gerektirecek kadar büyük veri kümelerine atıfta bulunmaktaydı. Günümüzde ise “gelişmiş öngörü ve karar verme için düşük maliyetli, yenilikçi bilgi işlem biçimleri talep eden yüksek hacimli, hızlı ve çeşitli bilgi varlıkları” olarak değerlendirilmektedir (Boyd ve Crawford, 2012, s. 663; Vis, 2013). Manovich (2012, s. 460) ise büyük veri terimini, “boyutu yaygın olarak kullanılan yazılım araçlarının verileri kısa bir sürede toplama, yönetme ve işleme yeteneğinin ötesinde olan veri setleri” olarak tanımlamaktadır. 21. yüzyılın başında, özel şirketlerin üretmeye başladığı veri setlerinden büyük veriler ortaya çıkmış ve *Google*, *Facebook* ve *LinkedIn* gibi bazı sosyal veri üreten şirketler en başından beri büyük veriler etrafında oluşturulmuştur (Gökkurt Demirtel ve Demirtel, 2016, s. 226; Social Media Research Group, 2016, s. 4).

Günümüzde sosyal medya kullanımındaki ve kullanıcı sayısındaki yükselişle beraber sosyal medya platformlarında üretilen veriler çok hızlı bir şekilde artmaya başlamıştır. Sosyal medya

kullanımındaki bu büyük artış, hem “sosyal medya büyük verisi” olarak adlandırılan kavramın hem de büyük miktarda verinin ortaya çıkmasını sağlamıştır (Stieglitz ve ark., 2018, s. 156). Bir başka ifade ile sosyal medya platformlarında yapılan faaliyetler arttıkça, üretilen veriler de artarak büyük veri denilen veri kümelerini meydana getirmiştir (Çelik, 2018, s. 453). Tüfekçi (2014, s. 505)’ye göre, sosyal medya büyük verilerinin ortaya çıkması ve bu sayede insan davranışlarının araştırılması, biyolojide mikroskopun, astronomide ise teleskopun icadına benzer bir devrim yapmıştır.

Sosyal medya büyük verileri, önemli kavrayışların anahtarı olarak görülerek; akademisyenler, kurumlar, şirketler, gazeteciler, politikacılar ve hükümetler tarafından kapsamlı bir şekilde analiz edilmiş, bu verilerin sürekli olarak kullanılmasıyla yapılan sosyal medya büyük veri analizleri, kurum ve şirketlerin politikalarına yön vermeye başlamıştır (Ghani, Hamid, Targio Hashem ve Ahmed, 2019, s. 426; Gökurt Demirtel ve Demirtel, 2016, s. 226; Tufekci, 2014, s. 505). Ancak sosyal medya verileri doğası gereği büyük miktarda ilk bakışta anlaşılamayan bilgi içermekte, bu da doğal olarak verilerin kullanılmasında ve analiz aşamasında bir zorluk oluşturabilmektedir. Sosyal medya büyük verileri anlamlı bir şekilde getirilip doğru şekilde kullanılmazsa onlardan faydalanmak pek mümkün olamamaktadır (Ghani ve ark., 2019, s. 418). Bu nedenle *Google*, *Twitter* ve *Facebook* gibi sosyal medya platformlarına öncülük eden şirketler, kendi büyük verilerinin doğru analizleri için kendi yazılımlarını, sistemlerini ve programlama dillerini geliştirmiştir. Bunun neticesinde sosyal medya platformları büyük veri analizi için kullanılabilir hâle gelmiştir (Çelik, 2018, s. 466).

Büyük veri kullanımının ortaya çıkaracağı analitik güç potansiyeli onun çok çeşitli sektörler ve disiplinler arasında yaygınlaşmasını sağlamıştır (Larson, 2020, s. 7). Zelenkauskaitė ve Bucy (2016), sosyal medya büyük verilerinin yararlarını, bölünmüş ve yapılandırılmamış veri kaynaklarını birleştirme ile sosyal olaylar konusundaki anlayışları geliştirme potansiyeli üzerinden değerlendirmektedir. Sosyal medya büyük verilerinin yapısının, araştırmacılar için bu verilere önemli zorluklar olmadan erişebilmeyi sağlaması sebebiyle çeşitli akademik alanlardaki araştırmacılar sosyal medyadaki büyük veriler üzerinde daha fazla araştırma yapmak istemektedir (Shneiderman, Preece ve Pirolli, 2011, s. 34). Büyük veri kaynağı olarak sosyal medya, diğer veri kaynaklarıyla bütünleşebilmekte, mevcut eğilimler hakkında bilgiye sahip olmak, araştırmalardaki fırsatları belirlemek ve araştırma sürecini desteklemek için kullanılabilir. Sosyal medya, araştırma sürecini desteklemenin yanında araştırma konularını geliştirmek için bir araç olabilmektedir (Lafferty ve Manca, 2015, s. 87). Sosyal medya verileri, kısa bir kod ya da program yazılması vasıtasıyla milyarlarca kullanıcıdan otomatik olarak elde edilmesine izin vermektedir. Bu sebeple araştırmacıların, önemli olayları inceleyebilmek için büyük veri kümelerini kullanma ihtiyacı, “büyük” ve “sosyal” verilerin miktarında ve türünde önemli bir artış sağlamıştır (Walker, 2017, s. 1).

Yapılandırılmış ve yapılandırılmamış şekilde toplanan çok sayıda veri kümesinden elde edilen sosyal medya büyük verileri sağlık, iletişim, siyaset gibi farklı bilimsel alanlarda

kullanılmıştır. Sosyal medya büyük verilerinin en yaygın kullanıldığı alanlar; sosyal medya analizi, duyarlılık analizi ve veri madenciliği olmuştur (Ghani ve ark., 2019, s. 417-418). Tüm bunların sonucu olarak, çeşitli sosyal medya araştırmacılarının, çok çeşitli alanlardaki araştırmalarında doğrudan sosyal medyadan elde ettikleri büyük verileri/büyük veri kümelerini kullandığı görülmüştür (Weller ve Kinder-Kurlanda, 2016, s. 166).

## **5. Arşivlenen Sosyal Medya Verilerinin Bir Araştırma Kaynağı Olarak Kullanımı ve Sosyal Medyadaki Araştırma Alanları**

### **5.1. Sosyal Medya Araştırması**

Sosyal medyanın ortaya çıkışı ile birlikte zamanla, dünya çapında toplumların düşündüklerini, ilgilerini, beğendiklerini, beğenmediklerini, isteklerini vb. açıkça ifade etmelerini sağlayan bir kamusal iletişim şekli hâline gelmiştir.

Son yıllarda sosyal medya adeta bir iletişim devrimi yapmış ve bu nedenle birçok alandan büyük ilgi görmüştür. Sosyal medyanın geniş kabulü, çeşitli topluluklardaki toplumsal ve ticari yaşamlarla ilişkileri büyük ölçüde değiştirmiş, bu da direkt olarak sosyal medya ile ilgili çalışmaların artışı sağlamıştır (Bishop ve ark., 2019, s. 4). Sosyal medyanın gösterdiği bu hızlı devrim, birçok alana uygunluğu ve kullanım özellikleri ile ilgili sonuçları dikkate alındığında bilgi teknolojileri de dâhil olmak üzere çeşitli disiplinlerden akademik araştırmacıların ilgisi sosyal medya üzerinde toplanmıştır (Kapoor ve ark., 2018, s. 532). Sosyal medyayı birçok disiplinle ilgili kılan şey, çok çeşitli alanlarda çok çeşitli kullanıcıların ürettiği verilerin sistemde mevcut olmasıdır. Sosyal medya, başlangıçta bir veri kaynağı olarak oluşturulmamıştır ancak sosyal medyanın zamanla kullanımının artması ve onun akademik çalışmalar için veri üreten bir araç olarak fark edilmesi ile birlikte sosyal medya, araştırmacıların çalışmalarında yaygın olarak kullandığı ana veri kaynaklarından biri hâline gelmiştir.

Hem sosyal medyanın araştırma amacıyla nasıl kullanılabilceği hem de sosyal medya kullanımı üzerine yapılan araştırmalar, son yıllarda dikkat çekici bir artış göstermiştir. Sosyal medya, araştırmacılara özellikle kullanıcıların davranış, inanç ve tutumları gibi konularda potansiyel olarak zengin bir bilgi kaynağı sunmaktadır (Sloan ve ark., 2015). İnternet ortamında devasa miktarda bulunan tweetler, gönderiler, etkileşimler, yorumlar, fotoğraflar gibi veriler, sosyal medya alanında istatistiksel ve bilimsel çalışma ve araştırmalara doğru yeni bir hareket kazandırmıştır. Veri odaklı bu hareket, çevrimiçi sosyal medya verilerinin analizine yönelik benzeri olmayan ciddi sayıda araştırma yapılmasını sağlamıştır (Aggarwal, 2011, s. 2-3).

Araştırma için sosyal medyayı ve sosyal medya verilerini kullanmanın olası avantajlarından bazıları, büyük nüfus alanlarında araştırma yürütme maliyetlerini azaltma, daha fazla katılımcıya ulaşarak uzun bir zaman süresince veri elde etme imkânıdır (Bishop ve ark., 2019, s. 4). Sosyal medyada üretilen verilerin, onu özellikle sosyal bilim araştırmaları için güçlü bir kaynak haline

getiren özellikler taşıması, bunun yanında *Facebook*, *Twitter*, *Instagram* gibi çok kullanılan sosyal medya platformlarında (fotoğraflar, videolar gibi) metin dışı verilerin giderek artan varlığı, yine özellikle sosyal bilim araştırmacılarının bu verileri birçok şekilde inceleyerek, toplumun bunları anlamasına da imkân vermektedir (Pennington, 2016, s. 232; Thomson ve Kilbride, 2015, s. 261).

Sosyal medya araştırması, sosyal medyayı kullanan bir araştırma şekli; çeşitli araçlar ve veri elde etme yöntemleri kullanarak insan topluluklarının çeşitli temalarla nasıl ilişki kurduğunu anlamasına yardımcı olmak ve çeşitli araştırmalar yapmak için sosyal medya verilerini analiz etme sürecidir (“Social Media Research: tools, techniques and outputs”, y.y.). Sosyal medya araştırması terimi, sosyal medyadan sağlanan verileri kullanan her türlü araştırmayı ifade etmektedir (Bishop ve ark., 2019, s. 4). Sosyal medyanın akademik araştırmalarda kullanımı çoğunlukla iki şekildedir. İlki araştırmanın yürütülmesini sağlamak için sosyal medya platformlarını araştırma aracı olarak kullanmak, ikincisi ise sosyal medyayı araştırma için veri kaynağı olarak kullanmaktır. Daha önce değinildiği gibi, sosyal medya araştırması, sosyal medya platformlarının ve verilerinin doğası nedeniyle, web tabanlı anketler ya da diğer çevrimiçi araştırma türlerinden farklıdır. Sosyal medya araştırması hemen hemen yeni bir çalışma alanıdır. Veriler gerçek zamanlı olarak veya buna yakın olarak elde edildiğinden, genellikle diğer analiz türlerinden çok daha hızlı ve ucuzdur (Social Media Research Group, 2016, s. 6-7). Sosyal medya araştırması, çevrimiçi olarak oluşturulan verileri incelemekte ve bu verilerden elde edilen eğilimleri ortaya çıkararak sosyal uyum sağlama ve insanların davranışlarını inceleme şeklinde sınırları çizilebilecek bir çerçeveye sahiptir. Sosyal medya araştırmaları, kullanıcıların sosyal medya platformlarını nasıl kullandıklarını, bu platformlardaki hareketlerini anlamak için yeni teknik ve yöntemler arayan uzmanlarla beraber sürekli olarak bir gelişim çabası içindedir (“Social Media Research: tools, techniques and outputs”, y.y.).

Sosyal medya verileri çok çeşitli alanlardan araştırmalarda kullanılmıştır/kullanılmaktadır. Ancak veriler, bunlarla sınırlı değildir. Akademik bağlam dışında, sanatçılar, yazarlar ve gazeteciler gibi diğer profesyoneller de sosyal medya verilerinin kullanılabilirliğinden yararlanmaktadır. Bu tür farklı alanların zaten sosyal medyadaki verileri kullanması, onların içeriğinin araştırma değerini göstermekte ve giderek daha fazla araştırmacı, elde ettikleri verilerin sonuçlarının tekrarlanabilirliği ve karşılaştırılabilirliği dolayısıyla sosyal medyadan toplanan araştırma verilerini kullanmaktadır (Weller ve Kinder-Kurlanda, 2016, s. 166).

API’lerin ortaya çıkışı ve sosyal medya platformlarının API’ler etrafında oluşturulması araştırmacıları, sosyal medya platformlarından en doğru ve hızlı şekilde veri toplayabilmek için platform API’lerini kullanmaya yöneltmiştir. API’lerin kullanımı, veri toplamanın ve verileri arşivlemenin en etkili yollarından biri olarak görülmektedir. Kullanıcılar tarafından paylaşılan gönderiler, çeşitli *hashtagler veya anahtar kelimeleri* içeren veriler, API’ler kullanılarak toplanabilmektedir (Bruns ve Burgess, 2016, s. 19). Sosyal medya verilerine erişim açısından,

API'lerin istenilen verilere gerçek zamanlı bir bağlantı oluşturabilmesi, genellikle ücretsiz olarak sunulması ve bunları daha sonra analiz etmeye, görselleştirebilmeye olanak sağlaması (Social Media Research Group, 2016, s. 9) gibi nedenlerle birçok alandaki araştırmacılar çalışmalarında sıklıkla sosyal medya API'leri aracılığıyla arşivlenen sosyal medya verilerini kullanmaktadır.

*API tabanlı sosyal medya araştırması*, sosyal medya platformları tarafından belirtilen standartlaştırılmış bilgileri içeren ve yine kontrolü platformlar tarafından sağlanan teknik arayüzler aracılığıyla elde edilen “önceden biçimlendirilmiş” verilere dayanmaktadır (Rieder, Abdulla, Poell, Woltering ve Zack, 2015, s. 2). *API tabanlı* bir araştırmanın araştırmacılar tarafından sıkça kullanılmasının bir nedeni de *API*lerin yüksek verimli araştırma araçları haline getirilerek veri toplama, organizasyon, temizleme, koruma ve analiz işlemlerinin otomatikleştirilmesidir. Araştırmacılar veri toplama süreçlerine platformları dâhil etmek zorunda kalmak yerine *API*ler sayesinde belli programlama becerileri dâhilinde ilgili verileri kendileri toplayabilmektedir (Lomborg ve Bechmann, 2014, s. 256-257). *API tabanlı sosyal medya araştırmalarında* en sık kullanılan sosyal medya platformu *Twitter* olarak görülmektedir. *Twitter*, *API*leri sayesinde araştırmacıların kolayca veri toplamasına olanak sağladığı için popüler bir kaynak hâline gelmiştir (Torres, Anu ve Varde, 2021). *Facebook* ve *Instagram* gibi diğer popüler sosyal medya platformlarının *Twitter*'a göre araştırmacılara daha kapalı olması (çoğunlukla *API* erişiminin kısıtlanması) araştırmacıların genel olarak çalışmalarını *Twitter*'a yöneltmiştir. Bu sebeple birçok akademik çalışma *Twitter* üzerine yoğunlaşmıştır. *Twitter*'da yapılan araştırmaların çokluğu, bu platformdaki verilerin hem toplanmasına hem de analiz edilmesine yönelik yaklaşımların standartlaştırılması yolundaki çabaların doğmasına da olanak sağlamıştır. *Twitter*'da belirli bir konu ile ilgili tweetleri bulmanın en sık kullanılan yöntemi *anahtar kelimeleri* veya özellikle *hashtagleri* kullanmaktır. *Hashtaglerin* kullanımı belirli sosyal medya avantajları sağladığından araştırmaların içeriğini tanımlamak için *hashtagler* yararlı filtreler olarak görülmektedir. Çeşitli çalışmalarda, aranan bilgilerin bulunması için *Twitter API*'sine erişilerek arama sorguları olarak tanımlanmış terimler kullanılmaktadır. Örneğin, belirli bir *hashtag* aranarak, o *hashtagi* içeren herkese açık tweetler bulunabilmektedir (Highfield ve Leaver, 2014).

## 5.2. Sosyal Medya Verilerinin Kullanıldığı Bazı Araştırma Alanları

Aşağıda dünyada sosyal medya platformlarının *API*leri aracılığıyla arşivlenen sosyal medya verileri kullanılarak yapılan çeşitli alanlardaki çalışmalardan örnekler verilmeye çalışılmıştır. Bu tür veriler kullanılarak yapılan akademik çalışmalar, doğal olarak bu çalışmada verilen örneklerden çok daha fazladır. Buradaki amaç, genel olarak sosyal medya platform *API*leri vasıtasıyla toplanarak arşivlenen veriler kullanılarak yapılan çeşitli alan (salgın ve doğal afet gibi kriz anları, sağlık, psikoloji, dilbilim, iletişim çalışmaları, pazarlama, siyasal faaliyetler ve coğrafya) araştırmaları ile ilgili bir farkındalık oluşturmaktır.



### 5.2.1. Terör

Sosyal medyada terör saldırıları gibi kriz anları ile ilgili çok sayıda akademik çalışma bulunmaktadır. Steensen (2018), 22 Temmuz 2011’de Norveç’te meydana gelen terör saldırısı sırasında ve sonrasında atılan tweetleri ve bu süreçte hashtaglerin oynadığı rolü analiz etmiştir. Veriler Twitter API’lerinden toplanmıştır. Çalışmada terör saldırısı gibi büyük bir acil durum olayı gerçekleştiğinde hashtaglerin Twitter iletişimi için hemen alakalı olmayabileceği sonucuna ulaşılmıştır. Merrill, Sumartojo, Stephens ve Coward (2020), 22 Mayıs 2017’de *Manchester Arena* bombalamasının birinci yıldönümü anmasında sosyal medyada belirgin olan birliktelik biçimlerini ve duygularını incelemiştir. Ayrıca çalışmada, Twitter içeriğinin hesaplamalı olarak toplanmasını ve analizini birleştiren metodolojik bir yaklaşım da sunulmuştur. Tweetler API’ler vasıtasıyla #ManchesterTogether hashtagi temel alınarak toplanmıştır. Sonuç olarak dijital olarak oluşturulan kamusal atmosferler için anma etkinlikleri bağlamında bir çerçeve sağlanmış ve özellikle teröre verilen tepkiler anlamında sosyal medyada birlikteliğin nasıl yeniden işlendiğine dair katkıda bulunulmuştur. Simon, Goldberg, Aharonson-Daniel, Leykin ve Adini (2014), 21 Eylül 2013’te 24 Eylül 2013’e kadar devam eden Kenya’daki Westgate alışveriş merkezine yapılan saldırı ile ilgili 21-25 Eylül 2013 arasındaki coğrafi konumlar, terör saldırısı, sosyal destek ve organizasyonlar gibi ana hashtag kategorilerine dayalı tweetleri Twitter API’leri vasıtasıyla toplamış, depolamış ve analiz etmiştir. Buna göre, Twitter’ın Kenya’da, acil durumlar için çok değerli bir iletişim ve müdahale koordinasyon aracı olarak hizmet ettiği sonucuna ulaşılmıştır.

### 5.2.2. Doğal Afetler

Sosyal medya çalışmalarında arşivlenmiş sosyal medya verilerinin kullanıldığı bir diğer kriz anı alanı da doğal afetlerdir. Örneğin, Bruns Burgess, Crawford ve Shaw (2012), 2011’de Yeni Zelanda’nın Güneydoğu Queensland eyaletindeki sel felaketinde Twitter’da yöneticilerin yanında, sıradan vatandaşların afet bilgilerini ve paylaşımları yaymadaki rolünü incelemiştir. 10-16 Ocak 2011 tarih aralığında Twitter API’lerinden sel felaketi ile ilgili toplanan hashtagler ve tweetler sonucunda, Twitter’ın Yeni Zelanda’daki bu doğal afette bilgi paylaşmada ve yaymada aktif bir rol oynadığı görülmüştür. Morstatter, Lubold, Pon-Barry, Pfeffer ve Liu (2014), çalışmalarında hem 2012’deki *Sandy Kasırgası* gibi bir doğal afete hem de 2013’teki *Boston Maratonu Bombalı Saldırısı* gibi bir terör olayına odaklanmıştır. Bu iki kriz anı ile ilgili Twitter API’si kullanılarak Amerika kıtasından coğrafi etiketli tweetler toplanmıştır. Buna göre bu iki kriz yeri üzerinden bir kriz bölgesinden atılan tweetlerin dili arasında fark olup olmadığı incelenmiştir. Sonuç olarak bir kriz bölgesinden atılan tweetlerin, başka zaman dilimlerinde başka yerlerden atılan tweetlerden farklı olduğu görülmüştür. Cresci, Tesconi, Cimino ve Dell’Orletta (2015), 2009’daki *L’Aquila Depremi*, 2012’deki *Emilia Depremi*, 2013’teki *Sardunya Seli* ve 2014’te İtalya’da meydana gelen 4 doğal afete yönelik tweetleri, API’ler aracılığıyla her doğal afet için belirlenen farklı hashtaglere göre toplamış ve bunları



arşivleyerek bir veri seti oluşturmuştur. Çalışmanın sonucunda afet hasar değerlendirmesi için açıklanmalı ilk İtalyan sosyal medya mesaj dizisi meydana getirilmiştir.

### 5.2.3. İletişim

Sosyal medyada akademik olarak incelenen bir diğer alan da iletişimdir. Bruns ve Burgess (2016, s. 18)'e göre, iletişim alanındaki sosyal medya araştırması, genellikle büyük ölçekli ve veriye dayalı olup geniş ve çeşitli kullanıcı gruplarında görülebilen sistemik iletişim süreçleri hakkındaki sorularla ve bunların belirli kısa zamanlı olaylara verdiği tepkileri araştırmakla ilgilenmektedir. Sosyal medyadaki iletişim araştırmaları, işletmelerin pazarlama birimleri tarafından, müşterilerden çeşitli bilgileri toplamak için kullanılabilir. Bu tür bilgileri toplamak, işletmelerin ve kurumların, strateji oluşturma süreçlerinde daha müşteri temelli olmalarına olanak tanımaktadır (Kavoura, 2014, s. 491). Bunun yanında siyasi iletişim de sosyal medyadaki iletişim araştırması alanlarından biridir. Bu çalışmalara bir örnek Kavanaugh, Song, Li ve Fox (2019), tarafından yapılan çalışmadır. Çalışmada seçimlerde siyasi hoşgörü ve iletişim özgürlüğünü araştırma amacıyla 2014 Tunus Parlamento ve Cumhurbaşkanlığı seçimlerinin üç aylık döneminde gönderilen Twitter içerikleri API'ler aracılığıyla toplanmıştır. Sonuçlar 2011 *Tunus Devrimi*'nden sonra yapılan bu ilk düzenli seçimlerde siyasi görüş ve iletişimlerin daha demokratik ve özgür olarak yapıldığına işaret etmektedir. Sosyal medyadaki iletişim araştırmalarına bir başka örnek de ElAlfy, Darwish ve Weber (2020) tarafından yapılan çalışmadır. Standard & Poor's<sup>7</sup> 500 adıyla bilinen 500 büyük Amerikan şirketini içeren ve bu şirketlerin kurumsal sürdürülebilir kalkınma hedefleri ile ilgili Twitter'da iletişim kurup kurmadıklarını inceleyen çalışma, Standard & Poor's 500'ün Twitter'da hesabı bulunan 433 şirketinin Twitter zaman çizelgesindeki tweetlerinin Twitter API'si aracılığıyla toplanması ile yapılmıştır. Sonuçlar şirketlerin temel faaliyetleri içeren sürdürülebilir kalkınma hedefleri hakkında Twitter'da iletişim kurduklarını göstermiştir.

### 5.2.4. Sağlık

Arşivlenmiş sosyal medya verileri sağlık alanındaki çalışmalarda da sıklıkla kullanılmaktadır. Örneğin, Ashok, Guruprasad, Prakash ve Shylaja (2019), yaptıkları çalışmada Twitter'dan elde edilen verilerin hastalıkları tespit etmek ve bunları verimli bir şekilde izlemek için kullanılabilirliğini araştırmış. Twitter API'si ile toplanıp arşivlenen veriler sayesinde salgınların erken tahmini için kullanılabilirliğini ve dolayısıyla salgın durumunun daha hızlı ve kolay bir şekilde ele alınabileceği sonucuna ulaşmıştır. Young, Rivers ve Lewis (2014), HIV önleme için coğrafi konumlu konuşmaların sosyal medya verilerinden çıkarılıp çıkarılmayacağını değerlendirerek, "HIV önleme" için gerçek zamanlı Twitter verilerini kullanmıştır. Bunun için 6 Mayıs ile 9 Aralık 2012 tarihleri arasında HIV riskiyle ilişkili hashtagler (örneğin cinsel davranışlar ve ilaç kullanımı) Twitter API'sinden toplanmıştır. Çalışma sonucunda,

7 Uluslararası kredi değerlendirme kuruluşu.

HIV önleme ve halk sağlığını korumak için gerçek zamanlı sosyal medya verilerinin, HIV risk davranışlarını tespit etmek ve sonuçlarını değerlendirmek için kullanılabileceği saptanmıştır. Joshi ve arkadaşları (2020), 2014 yılında Batı Afrika’da ortaya çıkan *Ebola* salgınının erken belirtilerinin sosyal medya paylaşımlarından elde edilip edilemeyeceğini incelemiştir. Batı Afrika’da salgından etkilenen bölgenin önemli şehirlerinde Aralık 2011 ve 2014 yılları arasında Twitter API’sinden *Ebola* ile ilgili toplanan tweetler araştırmanın veri setini oluşturmuştur. Bu araştırma sonucunda Ebola gibi salgın hastalıkların erken tespiti için sosyal medyadan faydalanılabileceği ortaya çıkmıştır. Sağlık alanındaki bir başka çalışmada Zhang, Hall ve Bastola (2018), hastaların ve sağlık hizmeti sağlayıcılarının kemoterapi hakkındaki algılarını değerlendirmek için Twitter API’sinden toplanan kanser ile ilgili tweetleri analiz etmiştir. Sonuç olarak çalışma, sosyal medyanın, kanser hastalarının nasıl iletişim kurduklarını anlamak ve ihtiyaçlarını öğrenebilmek için önemli bir sağlık veri kaynağı olduğunu göstermiştir. Edwards, Markovic, Petrunova, Lin ve Corsar (2018), Twitter’den 10 Kasım 2017-18 Aralık 2017 tarihleri arasındaki gıda kaynaklı hastalıkların nedenlerini araştırmak için kullanılan tweetleri Twitter API’si aracılığıyla toplayarak arşivlemiştir. Gelecekte bu alanda çalışacak araştırmacıların kullanabilmesi için arşivlenen veriler, tweet kimlikleri ve hashtaglerden oluşmaktadır. Benzer şekilde Beguerisse Diaz (2016) da Twitter API’si aracılığıyla Mart 2013 ile Ocak 2014 tarihleri arasında Twitter’da yayınlanan ve “diyabet” terimini içeren İngilizce tweetleri arşivleyerek bir veri seti oluşturmuştur.

### 5.2.5. Salgınlar

Bu çalışmanın hazırlandığı sırada 2019 yılında Çin’de ortaya çıkan milyonlarca insana bulaşan ve yine milyonlarcasının ölümüne sebep olan korona virüs salgını -literatürdeki adıyla COVID-19- devam etmekteydi. 21. yüzyıldaki en büyük salgınlardan biri olan bu salgın doğal olarak akademik çalışmaları da etkilemiş ve sosyal medyada COVID-19 salgını konu alan çok sayıda çalışma yapılmış ve veri arşivlenmiştir. Bunlardan birinde Liu, Xia ve Lang (2021), COVID-19 salgını sırasında insanların giyim tüketiminin nasıl değiştiğini ve bunun stresle başa çıkmayı sağlayıp sağlamadığını anlamak için API’ler aracılığıyla Ocak ve Eylül 2020 arasındaki tweetleri toplamış ve bunlarla duygu analizi gerçekleştirmiştir. Bunun sonucunda tüketim değişikliğinin stresle başa çıkmanın temel mekanizması olduğuna dair bulgular tespit edilmiştir. Ye, Özbay, Zuo ve Chen (2021), yine korona virüs salgını sırasında insanların seyahat ile ilgili tweetlerinin seyahat davranışları üzerindeki etkilerini New York şehri örneğinden API’ler aracılığıyla 13 Ocak ile 31 Ağustos 2020 tarihleri arasında toplanan verilerle incelemişler, sonuç olarak New York şehrinde salgın esnasında Twitter gönderilerinin zaman serisi ile seyahat oranı arasında benzer bir eğilim olduğunu tespit etmişlerdir. Bir başka çalışmada Sharma ve Sharma (2021), Twitter’da duygu analizi için COVID-19’un getirdiği kısıtlama ve karantina süreçlerinden etkilenen insanların depresyon ve intihar eğilimlerini araştırmıştır. 15 Şubat ile 19 Mayıs 2020 tarih aralığındaki tweetler Twitter API’si ile toplanıp

saklanarak analiz edilmiştir. Sonuç olarak bir yerde kısıtlı kalan insanların depresyon ve intihar benzeri duygular geliştirme olasılığının yüksek olduğu görülmüştür.

### 5.2.6. Psikoloji

Psikoloji alanında da sosyal medyadan araştırma faaliyetleri için faydalanılmaktadır. KVTKN ve Ramakrishnudu (2021), sosyal medyada psikolojik stres tespiti için Twitter API'sinden veri toplamıştır. Colombo, Burnap, Hodorog ve Scourfield (2016) intihar düşüncesi olan ve intihar düşüncelerini içeren tweetler paylaşan Twitter kullanıcılarının özelliklerini ve sosyal medya etkileşimlerini analiz etmiştir. Twitter API'sinden toplanan veriler ve intihar içerikli tweet paylaşımı yapan kullanıcı grubundan elde edilen sonuçlar göstermiştir ki Twitter'da intihara meyilli içerik paylaşımı yapan kullanıcılar yüksek oranda karşılıklı bağlantı kurmaktadır.

### 5.2.7. Siyasi Faaliyetler

Sosyal medya ile ilgili diğer büyük araştırma alanlarından biri de siyasi faaliyetlere katılım, siyasi faaliyetlerin analizi ve siyasi seçimlerdir. Meclis veya başkanlık seçimleri (özellikle Twitter'da) üzerine çok sayıda arşivlenmiş sosyal medya verisi ve bu veriler kullanılarak yapılan araştırma bulunmaktadır. Örneğin; Kaczmarek ve arkadaşları (2013), sosyal medyadaki siyasal iletişim için, 2013'teki *Alman Federal Meclisi* (Deutscher Bundestag) seçimlerinden faydalanarak Twitter'dan Alman Federal Meclisi adaylarının ve gazeteciler gibi haber yapımcılarının tweetlerini Twitter API'leri aracılığıyla hashtag ve tarih aralıklarına göre toplayarak arşivlemiştir. Sonuç olarak, Twitter verilerinin elde edilmesi ile sosyal medyadan veri toplamaya yönelik genel bir görünüş ortaya konmuş, seçim adayları tarafından gönderilen tüm tweetlerden bir dizi kimlik ve adaylar hakkındaki bilgiler arşivlenmiş ve yeniden kullanıma sunulmuştur. Bu sayede gelecekteki araştırmacıların veri kaynağı olarak kullanılabilecekleri bir veri seti de oluşturulmuştur. Aynı şekilde Sältzer ve arkadaşları (2021) da hem araştırmacılar Twitter'da seçim kampanyasını inceleme fırsatı vermek, hem de araştırmacıların ağ, metin ve görüntü analizi için ilgili verileri toplamasına olanak sağlamak için 2021 Almanya Federal Meclis seçiminde (Bundestagswahl) yarışan tüm adayların Twitter hesabını içeren verileri Twitter API'si aracılığıyla toplamış ve bunları arşivlemiştir. Chukwumeka ve Abdul (2017), Birleşik Krallık 2015 Genel Seçimleri için standart Twitter API'sini kullanarak "Sosyal Medya- Anlayış, Altyapı ve Katılımın Geliştirilmesi" adlı bir proje kapsamında tweetlerin kimliklerini arşivlemiştir. Sosyal medya verilerinin analizini çevreleyen etik hususlar hakkında rehberlik sağlamayı amaçlamıştır. Birch, Ounis, Macdonald, Craig ve Yang (2019), seçmenlere, adaylara, seçim görevlilerine ve gazetecilere yönelik seçim şiddetini açıklama hususunda 2015 Venezuela Parlamento Seçimlerine, 2016 Filipinler Genel Seçimlerine ve 2016 Gana Genel Seçimlerine odaklanmış ve bu amaçla Twitter gönderilerini arşivlemiştir. Çalışmanın, gelecekteki araştırmacıların sosyal medya verilerine dayalı seçim şiddeti araştırmaları yapmalarına kolaylık sağlayacağı düşünülmektedir.

Veri kaynağı olarak Instagram'ın kullanıldığı az sayıdaki araştırmadan biri, Mahoney, Feltwell, Ajuruchi ve Lawson (2016) tarafından 2014 İskoç Bağımsızlık Referandumu ve 2015 İngiltere Genel Seçimlerinde İskoç seçmenin sosyal medyada demokratik sürece katılım ve kendini siyasi bakımdan ifade etme aracı olarak fotoğraf paylaşımını nasıl kullandığına dair yapılan araştırmadır. Çalışmada veriler, Instagram API'sı kullanılarak #IndyRef (İskoç bağımsızlık referandumu için) ve #GE2015 (2015 Birleşik Krallık genel seçimleri için) hashtaglerinden her birinin oylamadan önceki 7 gün içinde bu hashtagler kullanılarak paylaşılan gönderilerden toplanması ile elde edilmiştir. Sonuç olarak İskoç seçmenlerinin siyasi olarak kendini ifade etmek için görsel paylaşımlar yaptığı anlaşılmıştır. Çalışma ayrıca, vatandaşların demokratik siyasi katılımı ile ilgilenirken ortaya çıkardığı görsel medya kullanımının anlaşılmasına da katkıda bulunmuştur. Görüldüğü üzere sosyal medyayı içeren birçok çalışmada seçimlerden ve bu seçimlerden elde edilen arşivlenmiş sosyal medya verilerinden faydalanılmıştır.

### 5.2.8. Dilbilim

Sosyal medyadan dilbilim alanındaki araştırmalar için de faydalanılmaktadır. Örneğin, Wang ve Inbar (2021), ABD siyasi elitleri tarafından kullanılan ahlaki dili incelemiştir. Çalışmada, Amerika Birleşik Devletleri Kongresi'nin Cumhuriyetçi ve Demokrat üyelerinin Twitter hesaplarından 1 Ocak 2016'dan 31 Ocak 2018'e kadar gönderdiği tweetler analiz edilmiş ve bunlar, üyelerin Kongre'de yaptığı konuşmalarla karşılaştırılmıştır. Tweetler, Twitter API'sinden toplanmıştır. Sonuç olarak ahlaki dil kullanımındaki partisel farklılıkların, partiler siyasi güç kaybettikçe veya kazandıkça zaman içinde değiştiği gözlemlenmiştir. Hegazi, Al-Dossari, Al-Yahy, Al-Sumari ve Hilal (2021), Twitter'dan Arapça metinleri işlemek, bunlardan bilgi çıkarımı yapabilmek ve gelecekteki çalışmalarda kullanabilmek amacıyla Twitter API'sinden toplanan Arapça tweetlerden veri seti oluşturulmuştur. Sonuç olarak gelecekte Arapça bilgi çıkarıcı uygulamalar geliştirmek için bir yaklaşım sunulmuştur.

### 5.2.9. Pazarlama

Sosyal medya günümüzde, pazarlama faaliyetlerinin de çok önemli bir parçası olmuştur. Milyarlarca kullanıcının davranışları, düşünceleri ve ihtiyaçları hakkında bilgi alışverişinde bulunduğu bir platform olan sosyal medya aynı zamanda ticari firmaların ürünlerini tanıttığı ve çeşitli kampanyaların yürütülmesini sağlamak ve öngörü kazanmak için de önemli bir kaynağı haline gelmiştir ("Social Media Research: tools, techniques and outputs", y.y.). Bu sebeple özellikle sosyal medya pazarlamasındaki tüketici/müşteri davranışları ve satın almayı da içine alan sosyal medyada hizmet sektörü en çok araştırma yapılan alanlardan biri haline gelmiştir. Örneğin, Liu, Shin ve Bruns (2021), lüks bir markanın sosyal medya pazarlama faaliyetlerinin müşteri katılımı üzerindeki etkisini araştırmak için en fazla Twitter takipçisine sahip ilk 15 lüks markanın Twitter API'si ile Temmuz 2012 ile Haziran 2017 tarih aralığında toplanan tweetlerini incelemiş sonuç olarak lüks bir markanın sosyal medya pazarlama çabalarının

etkileşim, eğlence ve popülerlik gibi özelliklerine odaklanmasının pazarlama faaliyetleri ve müşteri katılımını önemli ölçüde artırdığı görülmüştür.

### 5.2.10. Coğrafi Konum

Sosyal medya verileri coğrafi konumla ilgili araştırmalarda da sıklıkla kullanılmaktadır. Konum bilgileri, özellikle Twitter'daki, örneğin nüfus sayımlarından veya anketlerden elde edilen sosyo-demografik verilerin, sağlık verileri, çevresel veriler, coğrafi bilgiler gibi diğer çevrimdışı verilerle karşılaştırılmasına da imkân sağlamaktadır. Coğrafi konumla ilgili çalışmalara bir örnek, Kinder-Kurlanda, Weller, Zenk-Möltgen, Pfeffer ve Morstatter (2017)'in, yaptığı çalışmadır. Çalışmada Amerika Birleşik Devletleri'nden 2014 ve 2015 yıllarındaki coğrafi etiketli tweetler Twitter API'si ile toplanarak arşivlenmiştir. Böylelikle bu çalışma hem konum bilgisi taşıyan tweetlerin benzer özelliklerini ve bu verileri mevcut çalışmalarla karşılaştırmak isteyen araştırmacılar için bir kaynak oluşturmuş hem de coğrafi etiketli sosyal medya veri kümelerinin başarıyla arşivlenebileceğini göstermiştir. Benzer şekilde Sloan ve Morgan (2015) da Twitter kullanıcılarının demografik özelliklerini belirlemek için coğrafi hizmetler ve konum etiketlemeden faydalanmıştır. Twitter API'sinden elde edilen verilere göre Twitter'da coğrafi hizmetlerin kullanımının tweetlerin yazıldığı dil, sınıf yaş ve cinsiyet gibi demografik özelliklere bağlı olduğu sonucuna ulaşılmıştır.

Arşivlenmiş sosyal medya verilerininin coğrafi konum araştırmalarına konu olabileceğini gösteren diğer bir araştırma söz konusudur. Bu araştırmada Littman, 2017 yılında Amerika Birleşik Devletleri'ni etkileyen *Irma* ve *Harvey* adlı iki kasırga ile ilgili Twitter API'si kullanarak tweet kimliklerini toplayarak bir koleksiyon oluşturmuştur (Littman, 2017). Bu koleksiyondaki tweetler kullanılarak Brangbour ve arkadaşları (2019) tarafından bir çalışma yapılmıştır. Çalışmada, bir bölgedeki sel olasılıklarını hesaplamak için tweetlerdeki coğrafi bilgileri analiz etmiştir. Analiz sonucunda tweetlerin yaklaşık %17'sinin bir sel olayıyla ilgili coğrafi bilgilere sahip olduğu görülmüştür. Aggarwal (2019), sosyal medya gönderilerini, özellikle tweetleri, acil durumlarda haber veya haber dışı olarak sınıflandırmanın zorluğunu ele aldığı çalışmasında 5 veri koleksiyonundan faydalanmıştır. Bunlardan biri de Littman'ın tweet kimlik koleksiyonudur. Çalışma sonucunda Aggarwal, acil durumlarda bilgi edinmek için faydalı olabilecek, 5 farklı acil durum kategorisindeki Twitter gönderilerinden oluşan ilk veri setini oluşturmuştur.

### 5.2.11. Duygu Analizi

Yapılan literatür taramasında bir konu hakkında arşivlenmiş sosyal medya verilerininin başka araştırmalara da kaynak olabileceği görülmektedir. Örneğin Chen, Deb ve Ferrara (2021), Amerika Birleşik Devletleri'nin 2020'deki Başkanlık Seçimleri ile ilgili, çok dilli 1,2 milyarlık fazla tweeti toplamış ve bunları arşivlemiştir. Arşivlenmiş bu veriler hem Masood, Peterson

ve Voleti (y.y.) hem de Chaudhry ve arkadaşları (2021) tarafından Twitter'daki duygu analizi çalışmalarında kullanılmıştır. Yine aynı arşivlenmiş verileri kullanan bir başka çalışmada Chen ve arkadaşları (2021) COVID-19 salgını sırasında yapılan 2020 ABD Başkanlık Ön Seçimlerindeki halk sağlığı uygulamaları ile sağlıkla ilgili yanlış bilgilerin yayılımını incelemiş ve sonuç olarak bazı muhafazakâr kullanıcı gruplarının salgını önlemek için maske takma ve oy verme konularında Twitter'da yanlış bilgiler yaydığını bulmuştur.

### 5.2.12. Davranış Bilimi

Başka bir örnek Wrubel ve Kerchner (2020) tarafından 3 Mart 2020 ile 3 Aralık 2020 tarihleri arasındaki Twitter API'si kullanılarak oluşturulan "Coronavirus" veya "COVID-19" ile ilgili tweet arşividir. Bu arşiv kullanılarak yapılan çalışmalardan biri Sanders, Tosi, Obradovic, Miligi ve Delaney (2021) tarafından hazırlanmıştır. Çalışmada Birleşik Krallık'taki COVID-19 salgını sırasındaki karantina ve kısıtlamaların insanlar üzerindeki etkisi, davranış bilimi üzerinden sosyal medya da analiz edilerek incelenmiştir. Bu veri arşivi kullanılarak yapılan bir diğer çalışmada, Torres, Anu ve Varde (2021), Twitter kullanıcılarının COVID-19 ile ilgili bilgileri nereden edindiklerini hangi web kaynaklarını (URL'ler) kullandıklarını anlamayı amaçlamıştır. Sonuç olarak, COVID-19 salgını hakkında halkın sosyal medya ruhuna ilişkin görüş sağlanmıştır. Örnek çalışmalarda da görüldüğü üzere COVID-19 salgını birçok akademik çalışmaya konu olmuş/olmaya devam etmekte ve gelecekte de COVID-19 salgınının ve salgının etkilerinin akademik çalışmalarda sıkça kullanılacağı öngörülmektedir.

### 5.2.13. Yolsuzluk

Ye ve arkadaşları (2017), Çin'in yolsuzlukla mücadele politikası hakkında halkın sosyal medya üzerinden verdiği tepkilerin geleceğe aktarılmasını sağlamak için Çin'in en büyük sosyal medya platformlarından biri olan Sina Weibo adlı sosyal medya platformunun API'si aracılığıyla yolsuzlukla ilgili milyonlarca Weibo gönderisi ve binden fazla Sina blog sayfasını arşivlemiştir. Bunun sonucunda hem dijital bir koleksiyon oluşturulmuş hem de Çin'in dijital tarihinin önemli bir bölümü korunmuştur.

### 5.2.14. Toplumsal Bellek Oluşturma

Schafer, Truc, Badouard, Castex ve Musiani (2019), yaptıkları çalışmada Fransa'da Ocak ve Kasım 2015'te Paris, Temmuz 2016'da Nice kentlerinde düzenlenen terör saldırıları temelinde sosyal medyada saldırılarla ilgili yapılan paylaşımların *Fransa Milli Kütüphanesi* ve *Fransa Ulusal Görsel İşitsel Enstitüsü* tarafından API'ler aracılığıyla nasıl arşivlendiğini incelemiştir. Bu sayede sosyal medyadaki gerçek zamanlı çevrimiçi tepkilerin tüm anları kayıt altına alınarak toplumsal hafızanın korunmasına yardımcı olunmuştur.

## SONUÇ

Günümüzde birçok insan için *Twitter*'da düşüncelerini yazmak, *Instagram*'da bir fotoğraf paylaşmak veya *Facebook*'ta bir yorumda bulunmak artık günlük hayatın bir parçası haline gelmiştir. Sosyal medya bireysel kullanımının yanında çok farklı alanlardaki artan kullanımı ile ortaya çıkan kullanıcı verilerinin de artması adeta bir veri patlaması yaratmıştır. Sosyal medya, insan iletişiminin ortaya çıkardığı, günümüz toplumuyla ilgili pek çok konuda önemli miktarda doğal veri içerdiğinden bu da yeni bir veri türünün ortaya çıkmasını sağlamıştır.

Araştırmacıların sosyal medyanın bir veri kaynağı potansiyelini fark etmesi, geleneksel veri elde etme yöntemlerinin zor ve pahalı bir hale gelmesi sebebiyle araştırmacıların yeni veri kaynakları arayışına yönelmesi bu arada sosyal medyanın ortaya çıkışı ve kullanımının artması ile birlikte olmuştur. Ortaya çıkan bu yeni veri türü zamanla, akademik araştırmacıların çalışmalarında kullanabileceği alternatif bir veri kaynağı haline gelmiştir. Geleceğe yönelik tahminler yapmak, insan davranışlarını anlama ihtiyaçları için büyük miktarlarda, sayısız şekillerde bulunan veri kümelerini tanımlayan ve hacim, hız, çeşitlilik ve doğruluk gibi özellikler taşıyan bir terim olan “Büyük Veri”nin bir yönü olarak da kabul edilen bu veriler, “sosyal medya verisi” ya da “sosyal veri” olarak da adlandırılmıştır.

Sosyal medya verileri, akademik çalışmaların yanında bahsedildiği gibi çok çeşitli amaçlar için kullanılmaktadır. Bu kullanım özelliklerinin yanında sosyal medya verileri, ortak toplumsal hafıza özelliği de taşımaktadır. Sosyal medya verilerinin taşıdığı ortak toplumsal hafıza niteliği, ona, günümüz toplumlarının yaşamlarını geleceğe aktarılmasını sağlayacak bir belgeleme (toplumsal dokümantasyon) aracı olma görevi de yüklemektedir. Bununla birlikte araştırmacıların, sosyal medya verilerini anlamlı ve güvenilir bir şekilde toplaması, koruması ve geleceğe taşınması için ihtiyaç duyduğu araç olarak *API*'ler kullanıma sunulmuştur. Kısaca *API* (*Application Programming Interface*) olarak bilinen uygulama programlama arayüzlerinin ve bunu takip eden süreç içinde sosyal medya platformlarının da *API*'ler etrafındaki gelişimi, araştırmacıların verilere erişmek, verileri toplamak, arşivlemek ve veri koleksiyonları oluşturmak için *API*'lere yönelmesini sağlamıştır. Günümüzde araştırmacılar sosyal medya *API*'leri aracılığıyla şimdiye kadar görülmemiş bir ölçekte ve hızda sosyal medya verilerini toplayıp arşivleyebilmektedir.

Araştırmacılara; siyasi faaliyetler, sağlık, psikoloji ve iletişim gibi geniş alanlarda faydalanmaları için yeni fırsatlar sunan, *API*'ler yoluyla toplanıp arşivlenen sosyal medya verileri kullanılarak yapılan araştırmalar, bahsi geçen çok sayıda disiplinde gerçekleştirilmiştir. Ayrıca farklı disiplinler arasındaki çeşitli kullanımlar, araştırmada sosyal medya içeriğinin değerini ve sosyal medyanın birden fazla amaç ve bağlam için kullanıldığını da göstermiştir.

Sosyal medya platformlarından elde edilen verilerin araştırmalar için sunduğu fırsatlar açıktır. Sosyal medya verileri açıkça insan davranışının en büyük, en zengin ve en dinamik kanıt temeli



olduğundan araştırmacılar, bu zengin veri kaynağının nasıl daha kolay elde edip farklı alanlarda kullanma ve analiz etmenin yeni yöntemlerini her geçen gün daha fazla araştırmaktadır.

Sosyal medya platform verilerine araştırmalar için daha kolay erişim, sosyal medya verileri için dokümantasyon standartlarının geliştirilmesi için çalışmaların sürdürülmesi sosyal medya araştırmalarının gelişiminin bir yolu olarak görülmektedir. Bunun yanında veri kaynakları araştırmacılar için geçen zamanla birlikte sürekli olarak değişmeye devam etse de sosyal medya içeriğinin hâlihazırda kullanım miktarı, araştırmacıların bu verileri çok değerli bulduklarını göstermektedir.

Sosyal medya verileri kullanılarak yapılan sosyal medya araştırmaları için henüz yeni sayılabilecek bir aşamadır denilebilir. Bilgi ve belge yönetimi alanında da değeri yeni fark edilen bu veriler üzerine şimdiye kadar çok fazla çalışmanın yapılmamış olması, sosyal medya verilerinin güncel ve tarihsel bir kaynak olarak araştırmacılar tarafından kullanılabilceğini ortaya çıkmaktadır. Sosyal medya verilerinin geçici doğası sebebiyle bu veriler, hızla kaybolabilmektedir. Bu nedenle, gelecekteki akademik çalışmalar için sosyal medya verilerinin nasıl korunacağı ve yine gelecekte bu verilere nasıl erişilebileceği konusunda özellikle bilgi ve belge yöneticilerine önemli bir sorumluluk düşmektedir. Sosyal medya verileri ve bu verilerin arşivlenmesi üzerine bilgi ve belge yönetimi alanındaki çalışmaların çoğaltılması, bu verilerin kullanıldığı gelecekteki çalışmalar için yol gösterici olabilir.

**Teşekkür:** Bu çalışmanın hazırlanmasında değerli yardımlarını esirgemeyen Hocam Prof. Dr. İshak KESKİN'e ve meslektaşım aynı zamanda çalışma arkadaşım Dr. Özhan SAĞLIK'a teşekkürlerimi sunarım.

---

**Hakem Değerlendirmesi:** Dış bağımsız.

**Çıkar Çatışması:** Yazar çıkar çatışması bildirmemiştir.

**Finansal Destek:** Yazar bu çalışma için finansal destek almadığını beyan etmiştir.

**Peer-review:** Externally peer-reviewed.

**Conflict of Interest:** The author has no conflict of interest to declare.

**Grant Support:** The author declared that this study has received no financial support.

---

## Kaynakça/References

- Acker, A. ve Kriesberg, A. (2017). Tweets may be archived: Civic engagement, digital preservation and Obama White House social media data. *Proceedings of the Association for Information Science and Technology*, 54(1), 1-9. <https://doi.org/10.1002/pr2.2017.14505401001>
- Aggarwal, C. C. (2011). An introduction to social network data analytics. In C. Aggarwal (Ed.), *Social Network Data Analytics* (pp. 1-15). Boston, USA: Springer. [https://doi.org/10.1007/978-1-4419-8462-3\\_1](https://doi.org/10.1007/978-1-4419-8462-3_1)
- Aggarwal, P. (2019). Classification approaches to identify informative tweets. In *Proceedings Of The Student Research Workshop Associated with RANLP 2019* (pp. 7-15). Bulgaria: Varna. doi:10.26615/issn.2603-2821.2019\_002



- Al-Hajjar, D., Jaafar, N., Al-Jadaan, M. ve Alnutaifi, R. (2015). Framework for social media big data quality analysis. In N. Bassiliades et al. (Eds.), *New Trends in Database and Information Systems II* (pp. 301-314). Cham, Switzerland: Springer International Publishing. [https://doi.org/10.1007/978-3-319-10518-5\\_23](https://doi.org/10.1007/978-3-319-10518-5_23)
- Arshad, H., Jantan, A. ve Omolara, E. (2019). Evidence collection and forensics on social networks: research challenges and directions. *Digital Investigation*, 28, 126-138. <https://doi.org/10.1016/j.diin.2019.02.001>
- Ashok, A., Guruprasad, M., Prakash, C. O. ve Shylaja, S. S. (2019). A Machine learning approach for disease surveillance and visualization using Twitter data. In *2019 International Conference on Computational Intelligence in Data Science (ICCIDS)* (pp. 1-6). doi:10.1109/ICCIDS.2019.8862087
- Batrinca, B. ve Treleaven, P. C. (2015). Social media analytics: a survey of techniques, tools and platforms. *AI & SOCIETY*, 30(1), 89-116. <https://doi.org/10.1007/s00146-014-0549-4>
- Beguiris Diaz, M. (2016). Study of dynamic communities on networks, diabetes tweets 2013-2014 [Data collection]. *UK Data Service*. <http://doi.org/10.5255/UKDA-SN-852474>
- Beninger, K., Fry, A., Jago, N., Lepps, H., Nass, L., Silvester, H., ... Silvester, H. (2014). *Research using social media; users' views*. *NatCen Social Research*. Retrieved from <http://www.natcen.ac.uk/media/282288/p0639-research-using-social-media-report-final-190214.pdf%5Cnhttp://www.natcen.ac.uk>
- Birch, S., Ounis, I., Macdonald, C. ve Yang, X. (2019). Electoral violence incident dataset 2015-2016 [Data collection]. *UK Data Service*. <http://doi.org/10.5255/UKDA-SN-853262>
- Bishop, E. L., Hagen, S., Koščik, M., Vavra, M., Štebe, J., Ryan, L., ... Rød, L.-M. (2019). *Report on legal and ethical framework and strategies related to access, use, re-use, dissemination and preservation of social media data*. *SERISS (Synergies for Europe's Research Infrastructures in the Social Sciences)* (C. 6.3). Retrieved from [https://seriss.eu/wp-content/uploads/2019/11/D6.3-Report-on-legal-and-ethical-framework-and-strategies...\\_FINAL.pdf](https://seriss.eu/wp-content/uploads/2019/11/D6.3-Report-on-legal-and-ethical-framework-and-strategies..._FINAL.pdf)
- Boyd, D. ve Crawford, K. (2012). Critical questions for big data: Provocations for a cultural, technological, and scholarly phenomenon. *Information Communication and Society*, 15(5), 662-679. <https://doi.org/10.1080/1369118X.2012.678878>
- Brangbour, E., Bruneau, P., Marchand-Maillet, S., Hostache, R., Chini, M., Matgen, P. ve Tamisier, T. (2019). Computing flood probabilities using Twitter: Application to the Houston urban area during Harvey. In J. Brajard, A. Charantonis, C. Chen ve J. Runge (Eds.), *9th International Workshop on Climate Informatic* (pp. 22-26). Boulder, USA: National Center for Atmospheric Research. Retrieved from <https://opensky.ucar.edu/islandora/object/technotes%3A581/datastream/PDF/view>
- Bruns, A. ve Burgess, J. (2016). Methodological innovation in precarious spaces: the case of Twitter. In H. Snee, C. Hine, Y. Morey, S. Roberts ve H. Watson (Eds.), *Digital methods for social science: an interdisciplinary guide to research innovation* (pp. 17-33). London, UK: Palgrave Macmillan. [https://doi.org/10.1057/9781137453662\\_2](https://doi.org/10.1057/9781137453662_2)
- Bruns, A., Burgess, J., Crawford, K. ve Shaw, F. (2012). *#qldfloods and @QPSMedia: crisis communication on Twitter in the 2011 South East Queensland floods - research report*. Brisbane, Australia: ARC Centre of Excellence for Creative Industries and Innovation. Retrieved from <https://eprints.qut.edu.au/48241/>
- Caplan, M. A. ve Purser, G. (2019). Qualitative inquiry using social media: a field-tested example. *Qualitative Social Work*, 18(3), 417-435. <https://doi.org/10.1177/1473325017725802>
- Croeser, S. ve Highfield, T. (2019). Blended data: Critiquing and complementing social media datasets, big and small. In J. Hunsinger, M. M. Allen ve L. Klastrup (Eds.), *Second international handbook of internet research* (pp. 669-690). Dordrecht, Netherlands: Springer. [https://doi.org/10.1007/978-94-024-1555-1\\_15](https://doi.org/10.1007/978-94-024-1555-1_15)

- Çelik, U. (2018). Sosyal medya platformlarında veri madenciliği. H. Yıldız (Ed.), *Sosyal Medyanın İş Yaşamındaki Yeri* (ss. 451–471) içinde. İstanbul: Beta Yayıncılık.
- Chen, E., Chang, H., Rao, A., Lerman, K., Cowan, G. ve Ferrara, E. (2021). COVID-19 misinformation and the 2020 U.S. presidential election. *Harvard Kennedy School Misinformation Review* 1, 1-17. <https://doi.org/10.37016/mr-2020-57>
- Chen, E., Deb, A. ve Ferrara, E. (2021). #Election2020: The first public Twitter dataset on the 2020 US Presidential election. *Journal of Computational Social Science*. <https://doi.org/10.1007/s42001-021-00117-9>
- Chukwuemeka, D. ve Abdul, A. (2017). The UK 2015 General Election, Twitter data [Data collection]. *UK Data Service*. <http://doi.org/10.5255/UKDA-SN-852772>
- Colombo, G. B., Burnap, P., Hodorog, A. ve Scourfield, J. (2016). Analysing the connectivity and communication of suicidal users on Twitter. *Computer Communications*, 73, 291-300. <https://doi.org/10.1016/j.comcom.2015.07.018>
- Cresci, S., Tesconi, M., Cimino, A. ve Dell’Orletta, F. (2015). A Linguistically-driven approach to cross-event damage assessment of natural disasters from social media messages. In A. Gangemi, S. Leonardi ve A. Panconesi (Eds.), *Proceedings of the 24th International Conference on World Wide Web* (pp. 1195-1200). New York, USA: Association for Computing Machinery. <https://doi.org/10.1145/2740908.2741722>
- Data Mining VS Data Extraction: What’s the Difference? . (2021). Retrieved from <https://www.octoparse.com/blog/data-mining-vs-data-extraction-whats-the-difference>
- Derczynski, L., Yang, B. ve Jensen, C. (2013). Towards context-aware search and analysis on social media data. In G. Guerrini (Eds.), *ACM International Conference Proceeding Series* (pp. 137-142). <https://doi.org/10.1145/2452376.2452393>
- Edwards, P., Markovic, M., Petrunova, N., Lin, C. ve Corsar, D. (2018). Tweets used to explore causes of self-reported foodborne illnesses on social media 2017 [Data collection]. *UK Data Service*. <http://doi.org/10.5255/UKDA-SN-853375>
- EIAlfy, A., Darwish, K. M. ve Weber, O. (2020). Corporations and sustainable development goals communication on social media: Corporate social responsibility or just another buzzword? *Sustainable Development*, 28(5), 1418-1430. <https://doi.org/10.1002/sd.2095>
- Ertam, F. (2017). Sosyal medya verileri için etkili bir sınıflandırma yaklaşımı. *Fırat Üniversitesi Mühendislik Bilimleri Dergisi*, 29(2), 67-73. <https://dergipark.org.tr/pub/fumbd/339547> adresinden erişildi.
- ESOMAR. (2011). *Guideline on Social Media Research*. Retrieved from [https://www.esomar.org/uploads/public/knowledge-and-standards/codes-and-guidelines/ESOMAR\\_Guideline-for-online-research.pdf](https://www.esomar.org/uploads/public/knowledge-and-standards/codes-and-guidelines/ESOMAR_Guideline-for-online-research.pdf)
- Fondren, E. ve McCune, M. M. (2018). Archiving and preserving social media at the Library of Congress: Institutional and cultural challenges to build a Twitter archive. *Preservation, Digital Technology and Culture*, 47(2), 33-44. <https://doi.org/10.1515/pdte-2018-0011>
- Ghani, N. A., Hamid, S., Targio Hashem, I. A. ve Ahmed, E. (2019). Social media big data analytics: A survey. *Computers in Human Behavior*, 101, 417-428. <https://doi.org/10.1016/j.chb.2018.08.039>
- Gökkurt Demirtel, Ö. ve Demirtel, H. (2016). Sosyal medyada kurumsal arşivleme sorunu. T. Çakmak ve N. Özel (Ed.), *Uluslararası Kültürel Mirasın ve Kültürel Bellek Kurumlarının Yönetimi Kongresi* (ss. 226-234) içinde. Ankara: Üniversite ve Araştırma Kütüphanecileri Derneği.
- Gundecha, P. ve Liu, H. (2012). Mining social media: A brief introduction. *TutORials in Operations Research* (pp. 1-17). <https://doi.org/10.1287/educ.1120.0105>

- Gürsakan, N. (2009). *Sosyal ağ analizi: Pajek Ucinet ve Gmine uygulamalı*. Bursa: Dora Yayıncılık.
- Hegazi, M. O., Al-Dossari, Y., Al-Yahy, A., Al-Sumari, A. ve Hilal, A. (2021). Preprocessing Arabic text on social media. *Heliyon*, 7(2), 1-15. <https://doi.org/10.1016/j.heliyon.2021.e06191>
- Hemphill, L., Hedstrom, M. L. ve Leonard, S. H. (2020). Saving social media data: Understanding data management practices among social media researchers and their implications for archives. *Journal of the Association for Information Science and Technology*, 72(1), 97-109. <https://doi.org/10.1002/asi.24368>
- Highfield, T. ve Leaver, T. (2014). A methodology for mapping Instagram hashtags. *First Monday*, 20(1). <https://doi.org/10.5210/fm.v20i1.5563>
- Huang, Q. ve Xu, C. (2014). A data-driven framework for archiving and exploring social media data. *Annals of GIS*, 20(4), 265-277. <https://doi.org/10.1080/19475683.2014.942697>
- Joshi, A., Sparks, R., Karimi, S., Yan, S.-L. J., Chughtai, A. A., Paris, C. ve MacIntyre, C. R. (2020). Automated monitoring of tweets for early detection of the 2014 Ebola epidemic. *PLOS ONE*, 15(3), e0230322. Erişim adresi: <https://doi.org/10.1371/journal.pone.0230322>
- Kaczmarek, L., Mayr, P., Vatrapu, R., Bleier, A., Blumenberg, M., Gummer, T., ... Wolf, C. (2013). *Social media monitoring of the campaigns for the 2013 German Bundestag elections on Facebook and Twitter*. Köln. Retrieved from [https://www.gesis.org/fileadmin/upload/forschung/publikationen/gesis\\_reihen/gesis\\_arbeitsberichte/WorkingPapers\\_2014-31.pdf](https://www.gesis.org/fileadmin/upload/forschung/publikationen/gesis_reihen/gesis_arbeitsberichte/WorkingPapers_2014-31.pdf)
- Kaplan, A. M. ve Haenlein, M. (2010). Users of the world, unite! The challenges and opportunities of Social Media. *Business Horizons*, 53(1), 59-68. <https://doi.org/10.1016/j.bushor.2009.09.003>
- Kapoor, K. K., Tamilmani, K., Rana, N. P., Patil, P., Dwivedi, Y. K. ve Nerur, S. (2018). Advances in social media research: Past, present and future. *Information Systems Frontiers*, 20(3), 531-558. <https://doi.org/10.1007/s10796-017-9810-y>
- Kavanaugh, A., Song, Z., Li, L. ve Fox, E. (2019). Communication behavior in an emerging democracy. In Y. Chen, F. Salem ve A. Zuiderwijk (Eds.), *Proceedings of the 20th Annual International Conference on Digital Government Research* (pp. 445-455). New York, USA: Association for Computing Machinery. <https://doi.org/10.1145/3325112.3325263>
- Kavoura, A. (2014). Social media, online imagined communities and communication research. *Library Review*, 63(6/7), 490-504. <https://doi.org/10.1108/LR-06-2014-0076>
- Kerchner, D. ve Wrubel, L. (2020). Coronavirus tweet ids. Harvard Dataverse. <https://doi.org/10.7910/DVN/LW0BTB>
- Kinder-Kurlanda, K. (2019). Privatheitsschutz durch open data und trusted third parties: plädoyer für die öffentliche kontrolle sozialer daten. In C. Ochs, M. Friedewald, T. Hess ve J. Lamla (Eds.), *Die Zukunft der datenökonomie: zwischen geschäftsmodell, kollektivgut und verbraucherschutz* (pp. 103-116). Wiesbaden, Germany: Springer Fachmedien Wiesbaden. [https://doi.org/10.1007/978-3-658-27511-2\\_6](https://doi.org/10.1007/978-3-658-27511-2_6)
- Kinder-Kurlanda, K., Weller, K., Zenk-Möltgen, W., Pfeffer, J. ve Morstatter, F. (2017). Archiving information from geotagged tweets to promote reproducibility and comparability in social media research. *Big Data & Society*, 4(2), 1-14. <https://doi.org/10.1177/2053951717736336>
- KVTKN, P. ve Ramakrishnu, T. (2021). A novel method for detecting psychological stress at tweet level using neighborhood tweets. *Journal of King Saud University - Computer and Information Sciences*. <https://doi.org/10.1016/j.jksuci.2021.08.015>

- Lafferty, N. T. ve Manca, A. (2015). Perspectives on social media in and as research: A synthetic review. *International Review of Psychiatry*, 27(2), 85-96. <https://doi.org/10.3109/09540261.2015.1009419>
- Larson, E. (2020). Big questions: Digital preservation of big data in government. *The American Archivist*, 83(1), 5-20. <https://doi.org/10.17723/0360-9081-83.1.5>
- Littman, J. (2017). Hurricanes Harvey and Irma tweet ids. Harvard Dataverse. <https://doi.org/10.7910/DVN/QRKIBW>
- Liu, C., Xia, S. ve Lang, C. (2021). Clothing consumption during the COVID-19 Pandemic: Evidence from mining tweets. *Clothing and Textiles Research Journal*, 39(4), 314-330. <https://doi.org/10.1177/0887302X211014973>
- Liu, X., Shin, H. ve Burns, A. C. (2021). Examining the impact of luxury brand's social media marketing on customer engagement: using big data analytics and natural language processing. *Journal of Business Research*, 125, 815-826. <https://doi.org/10.1016/j.jbusres.2019.04.042>
- Lomborg, S. ve Bechmann, A. (2014). Using APIs for data collection on social media. *The Information Society*, 30(4), 256-265. <https://doi.org/10.1080/01972243.2014.915276>
- Lynn, T., Healy, P., Kilroy, S., Hunt, G., van der Werff, L., Venkatagiri, S. ve Morrison, J. (2015). Towards a general research framework for social media research using big data. In J. Kaye ve A. Druin (Eds.), *2015 IEEE International Professional Communication Conference, IPCC 2015, Limerick, Ireland, July 12-15, 2015* (pp. 1-8). <https://doi.org/10.1109/ipcc.2015.7235843>
- Mahoney, J., Feltwell, T., Ajuruchi, O. ve Lawson, S. (2016). Constructing the visual online political self: An analysis of Instagram use by the Scottish electorate. In *CHI '16: Proceedings of the 2016 CHI Conference on Human Factors in Computing Systems* (pp. 3339-3351). New York, USA: Association for Computing Machinery. <https://doi.org/10.1145/2858036.2858160>
- Manovich, L. (2012). Trending: The promises and the challenges of big social data. In M. K. Gold (Ed.), *Debates in the digital humanities* (pp. 460-475). Minnesota, USA: University of Minnesota Press. Erişim adresi: <http://www.jstor.org/stable/10.5749/j.ctttv8hq.30>
- Masood, Z., Peterson, H. ve Voleti, S. (y.y.) The sentiment of U.S. Presidential Elections on Twitter. Retrieved from [https://dsc-capstone.github.io/projects-2020-2021/reports/project\\_1.pdf](https://dsc-capstone.github.io/projects-2020-2021/reports/project_1.pdf)
- Merrill, S., Sumartojo, S., Stephens, A. C. ve Coward, M. (2020). Togetherness after terror: The more or less digital commemorative public atmospheres of the Manchester Arena bombing's first anniversary. *Environment and Planning D: Society and Space*, 38(3), 546-566. <https://doi.org/10.1177/0263775819901146>
- Morstatter, F., Lubold, N., Pon-Barry, H., Pfeffer, J. ve Liu, H. (2014). Finding eyewitness tweets during crises. In *Proceedings of the ACL 2014 Workshop on Language Technologies and Computational Social Science* (pp. 23-27). Baltimore, USA: Association for Computational Linguistics. doi:10.3115/v1/W14-2509
- Nazeer Chaudhry, H., Javed, Y., Kulsoom, F., Mehmood, Z., Iqbal Khan, Z., Shoaib, U., ... Abualigah, L. (2021). Sentiment analysis of before and after elections: Twitter data of U.S. election 2020. *Electronics*, 10(17), 1-26. <https://doi.org/10.3390/electronics10172082>
- Pennington, D. R. (2016). Coding of non-text data. In A. Q.-H. Luke Sloan (Eds.), *The SAGE handbook of social media research methods* (pp. 513-529). London: SAGE Publications. <https://dx.doi.org/10.4135/9781473983847.n15>
- Procter, R., Voss, A. ve Lvov, I. (2015). Audience research and social media data : Opportunities and challenges. *Participations: Journal of Audience Reception Studies*, 12(1), 470-493. Retrieved from <https://core.ac.uk/download/pdf/30319349.pdf>

- Pushpam, C. A. ve Jayanthi, J. G. (2017). Over view on data mining in social media. *International Journal of Computer Sciences and Engineering*, 5(11), 147-157. <https://doi.org/10.26438/ijcse/v5i11.147157>
- Qi, M. ve Edgar-Nevill, D. (2011). Social networking searching and privacy issues. *Information Security Technical Report*, 16(2), 74-78. <https://doi.org/10.1016/j.istr.2011.09.005>
- Rieder, B., Abdulla, R., Poell, T., Woltering, R. ve Zack, L. (2015). Data critique and analytical opportunities for very large Facebook pages: Lessons learned from exploring “We are all Khaled Said”. *Big Data & Society*, 2(2), 1-22. <https://doi.org/10.1177/2053951715614980>
- Sältzer, M., Stier, S., Bäuerle, J., Blumenberg, M., Mechkova, V., Pemstein, D., ... Wilson, S. (2021). Twitter accounts of candidates in the German federal election 2021. GESIS Data Archive. <https://doi.org/10.4232/1.13790>
- Sanders, J. G., Tosi, A., Obradovic, S., Miligi, I. ve Delaney, L. (2021). Lessons from the UK’s lockdown: Discourse on behavioural science in times of COVID-19. *Frontiers in Psychology*, 12, 1-24. <https://doi.org/10.3389/fpsyg.2021.647348>
- Schafer, V., Truc, G., Badouard, R., Castex, L. ve Musiani, F. (2019). Paris and Nice terrorist attacks: Exploring Twitter and web archives. *Media, War & Conflict*, 12(2), 153-170. <https://doi.org/10.1177/1750635219839382>
- Şeker Elmas, Ş. (2019). *Sosyal medya mesajlarının veri madenciliği yöntemi ile duygu analizi (Sivas ili örneği)* (Yüksek lisans tezi). Yükseköğretim Kurulu Başkanlığı Ulusal Tez Merkezi. (Tez No. 538097)
- Sharma, S. ve Sharma, S. (2021). Analyzing the depression and suicidal tendencies of people affected by COVID-19’s lockdown using sentiment analysis on social networking websites. *Journal of Statistics and Management Systems*, 24(1), 115-133. <https://doi.org/10.1080/09720510.2020.1833453>
- Shneiderman, B., Preece, J. ve Pirolli, P. (2011). Realizing the value of social media requires innovative computing research. *Communications of the ACM*, 54(9), 34-37. <https://doi.org/10.1145/1995376.1995389>
- Simon, T., Goldberg, A., Aharonson-Daniel, L., Leykin, D. ve Adini, B. (2014). Twitter in the cross fire-the use of social media in the Westgate Mall terror attack in Kenya. *PLOS ONE*, 9(8). <https://doi.org/10.1371/journal.pone.0104136>
- Şimşek, M. U. (2012). *Sosyal ağlarda veri madenciliği üzerine bir uygulama* (Yüksek lisans tezi). Yükseköğretim Kurulu Başkanlığı Ulusal Tez Merkezi. (Tez No. 321573)
- Sloan, L. ve Morgan, J. (2015). Who tweets with their location? Understanding the relationship between demographic characteristics and the use of geoservices and geotagging on Twitter. *PLOS ONE*, 10(11), 1-15. <https://doi.org/10.1371/journal.pone.0142209>
- Sloan, L., Morgan, J., Burnap, P. ve Williams, M. (2015). Who tweets? Deriving the demographic characteristics of age, occupation and social class from Twitter user meta-data. *PLOS ONE*, 10(3), e0115545. <https://doi.org/10.1371/journal.pone.0115545>
- Social media archiving in regulated industries: Why it matters in 2021. (2020). Retrieved from <https://jatheon.com/blog/social-media-archiving-in-regulated-industries/>
- Social Media Research: tools, techniques and outputs. (y.y.). Retrieved from <https://www.pulsarplatform.com/what-is-social-media-research-top-tools-and-techniques/>
- Social Media Research Group. (2016). Using social media for social research : An introduction, (May), 1-27. Retrieved from [https://assets.publishing.service.gov.uk/government/uploads/system/uploads/attachment\\_data/file/524750/GSR\\_Social\\_Media\\_Research\\_Guidance\\_-\\_Using\\_social\\_media\\_for\\_social\\_research.pdf%0Ahttp://dera.ioe.ac.uk/26600/](https://assets.publishing.service.gov.uk/government/uploads/system/uploads/attachment_data/file/524750/GSR_Social_Media_Research_Guidance_-_Using_social_media_for_social_research.pdf%0Ahttp://dera.ioe.ac.uk/26600/)

- Steensen, S. (2018). Tweeting terror: an analysis of the Norwegian Twitter-sphere during and in the aftermath of the 22 July 2011 terrorist attack. In H. Hommen ve K. Backholm (Eds.), *Social media use in crisis and risk communication* (pp. 15-41). Bingley, UK: Emerald Publishing Limited. <https://doi.org/10.1108/978-1-78756-269-120181006>
- Stieglitz, S., Mirbabaie, M., Ross, B. ve Neuberger, C. (2018). Social media analytics-challenges in topic discovery, data collection, and data preparation. *International Journal of Information Management*, 39, 156-168. <https://doi.org/10.1016/j.ijinfomgt.2017.12.002>
- Strategies for managing social media information. (y.y.). Retrieved from <https://www.records.nsw.gov.au/recordkeeping/advice/strategies-for-managing-social-media-information>
- Taylor, J. ve Pagliari, C. (2018). Mining social media data: How are research sponsors and researchers addressing the ethical challenges? *Research Ethics*, 14(2), 1-39. <https://doi.org/10.1177/1747016117738559>
- Thomson, S. D. ve Kilbride, W. (2015). Preserving social media: The problem of access. *New Review of Information Networking*, 20(1-2), 261-275. <https://doi.org/10.1080/13614576.2015.1114842>
- Torres, J., Anu, V. ve Varde, A. S. (2021). Understanding the information disseminated using twitter during the COVID-19 pandemic. In *2021 IEEE International IOT, Electronics and Mechatronics Conference, (IEMTRONICS)* Piscataway, USA: IEEE. doi:10.1109/IEMTRONICS52119.2021.9422523
- Townsend, L. ve Wallace, C. (2016). *Social media research: A guide to ethics*. The University of Aberdeen. Retrieved from [https://www.gla.ac.uk/media/Media\\_487729\\_smxx.pdf](https://www.gla.ac.uk/media/Media_487729_smxx.pdf)
- Türkiye Bilimler Akademisi. (y.y.). *Türkçe bilim terimleri sözlüğü*. <https://terim.tuba.gov.tr/> adresinden erişildi.
- Tufekci, Z. (2014). Big questions for social media big data: Representativeness, validity and other methodological pitfalls. In E. Adar ve P. Resnick (Eds.), *Proceedings of the 8th International Conference on Weblogs and Social Media, ICWSM 2014* Retrieved from <http://arxiv.org/abs/1403.7400>
- van Rooyen, G. (y.y.). What social media can tell us about a city? Retrieved from <https://blog.pagefreezer.com/what-social-media-can-tell-us-about-a-city>
- Vis, F. (2013). A critical reflection on big data: Considering APIs, researchers and tools as data makers. *First Monday*, 18(10). <https://doi.org/10.5210/fm.v18i10.4878>
- Walker, S. (2017). *The complexity of collecting digital and social media data in ephemeral contexts*. (Doctoral thesis). Retrieved from [https://digital.lib.washington.edu/researchworks/bitstream/handle/1773/40612/Walker\\_washington\\_0250E\\_17763.pdf?sequence=1&isAllowed=y](https://digital.lib.washington.edu/researchworks/bitstream/handle/1773/40612/Walker_washington_0250E_17763.pdf?sequence=1&isAllowed=y)
- Wang, S.-Y. N. ve Inbar, Y. (2021). Moral-language use by U.S. political elites. *Psychological Science*, 32(1), 14–26. <https://doi.org/10.1177/0956797620960397>
- Weller, K. ve Kinder-Kurlanda, K. E. (2016). A manifesto for data sharing in social media research. In W. Nejdl ve W. Hall (Eds.), *WebSci '16: Proceedings of the 2016 ACM Web Science* (pp. 166–172). New York, USA: Association for Computing Machinery. <https://doi.org/10.1145/2908131.2908172>
- Ye, Q., Ozbay, K., Zuo, F. ve Chen, X. (2021). Impact of social media on travel behaviors during the COVID-19 pandemic: Evidence from New York City. *Transportation Research Record*, 1-20. <https://doi.org/10.1177/03611981211033857>
- Ye, Y., Ye, D., Zeljak, C., Kerchner, D., He, Y., Ye, Y. ;, ... Littman, J. (2017). Web-archiving Chinese social media: Final project report. *Journal of East Asian Libraries*, 2017(165), 1-22. Retrieved from <https://scholarsarchive.byu.edu/cgi/viewcontent.cgi?article=2701&context=jeal>
- Young, K. (2013). Online and internet-based technologies: Social networking. In C. J. ve B. B. Sara Price (Ed.), *The SAGE handbook of digital technology research* (pp. 427-442). London: SAGE Publications. <https://dx.doi.org/10.4135/9781446282229>

- Young, S. D., Rivers, C. ve Lewis, B. (2014). Methods of using real-time social media technologies for detection and remote monitoring of HIV outcomes. *Preventive medicine*, 63, 112-115. <https://doi.org/10.1016/j.ypmed.2014.01.024>
- Zelenkauskaitė, A. ve Bucy, E. P. (2016). View of a scholarly divide: Social media, big data, and unattainable scholarship. *First Monday*, 21(5). Retrieved from <https://firstmonday.org/article/view/6358/5511>
- Zeller, F. (2017). Analyzing social media data and other data sources: A methodological overview. In A. Q.-H. Luke Sloan (Eds.), *The SAGE Handbook Of Social Media Research Methods* (pp. 386-403). <https://dx.doi.org/10.4135/9781473983847.n23>
- Zhang, L., Hall, M. ve Bastola, D. (2018). Utilizing Twitter data for analysis of chemotherapy. *International Journal of Medical Informatics*, 120, 92–100. <https://doi.org/10.1016/j.ijmedinf.2018.10.002>

