

Büyük Veriye Dayalı Araştırmalarda Yöntemsel Sorunlar: Güncel Kadın Araştırmaları
Methodical Issues in Big Data-Based Researches: Current Women's Studies

Burcu Gümüş¹

¹ Doğuş Üniversitesi, İstanbul, Türkiye, burcuqumus@dogus.edu.tr, ORCID ID: 0000-0003-2643-2744.

**Article Info/Makale
Bilgisi**

Received/Geliş:
24.07.2023

Accepted/Kabul:
03.12.2023

Online/Çevrimiçi:
30.12.2023

Öz: Web 2.0 teknolojisinin ortaya çıkması ve sosyal medya mecralarının yaygınlaşması sonucu büyük verinin farkına varılması, bu veri türünü kullanarak yapılan analizlerin hızla popülerlik kazanmasını sağlamıştır. Bu durum, araştırmacılar için yeni araştırma fırsatlarının ortaya çıkmasına neden olurken çeşitli zorlukları da beraberinde getirmiştir. Bu çalışmanın amacı, sosyal medyadan elde edilen veriler ile yapılan çalışmaların geleneksel yöntemlerle yapılan çalışmalara kıyasla yeni bir anlayış ortaya koyup koymadığını değerlendirmektir. Çalışma kapsamında Türkiye'deki kadın araştırmaları konusunda 2013 – 2022 yılları arasında sosyal bilimler alanında yayınlanan ve araştırma verileri büyük veri olan 47 makale dikkate alınmıştır. Söz konusu bu çalışmaların veri setlerini hangi mecralardan nasıl elde ettikleri, veri setinin ne denli büyük olduğu, hangi analiz yöntemlerinin kullanıldığı, çalışmaların hangi konu başlıkları etrafında toplandığı, kısıtların neler olduğunu belirlenerek büyük veriye dayalı yapılan çalışmalara daha geniş bir çerçeveden bakabilmek hedeflenmiştir. Böylece büyük veri kullanarak yapılan kadın çalışmalarındaki uygulamalar hakkında bir değerlendirme ve tartışma sunulmuştur.

Anahtar Kelimeler: Kadın, Kadın Çalışmaları, Sosyal Medya, Büyük Veri, Sosyal Medya Araştırmaları

Cite this article/Atf:
Gümüş B, Büyük Veriye
Dayalı Araştırmalarda
Yöntemsel Sorunlar:
Güncel Kadın
Araştırmaları, Saü Kadın
Araştırmaları Dergisi,
2(2), (2023), 90-105.

Abstract: The emergence of Web 2.0 technology and the widespread adoption of social media platforms have led to the recognition of big data. This, in turn, has rapidly popularized the use of analyses conducted with this type of data. While this situation has created new research opportunities for researchers, it has also brought various challenges along with it. The aim of this study is to evaluate whether research conducted using data obtained from social media presents a new understanding compared to studies conducted using traditional methods. Within the scope of the study, 47 articles published in the field of social sciences in Turkey between 2013 and 2022, with research data derived from big data, have been examined. The goal of this study is to examine research conducted based on big data from a broader perspective by identifying how these studies obtained their datasets from which platforms, the size of the dataset, the analysis methods employed, the subject areas around which the studies revolved, and the limitations they encountered. Thus, an evaluation and discussion have been presented regarding the applications in women's studies conducted using big data.

Keywords: Woman, Women's Studies, Social Media, Big Data, Social Media Studies.

Extended Abstract

Methodical Issues in Big Data-Based Researches: Current Women's Studies

Burcu GümüŖ¹

¹*Doğuş University, Istanbul, Türkiye, burcugumus@dogus.edu.tr*

Social media enables the production of big data, which is a huge source in terms of speed, volume, diversity and scope. Big data by its very nature appeals to different disciplines and branches of research. However, while this situation offers opportunities for scientists and researchers, it also causes difficulties. One of the most important challenges experienced in the research process is the lack of production of big data for research hypotheses. This requires a re-evaluation of the steps related to the classical research methods adopted to date. Within the scope of this study, it was first tried to evaluate whether big data poses a threat to classical research stages and methods. The aim of this study is based on revealing the method followed in the study design and data collection in the women's studies conducted between 2013 and 2022. For this purpose, it was examined how the data was obtained, the duration of the data collection, the size of the data set collected, how it differed in terms of data collection tools, what the analyzes of the collected data were, whether the constraints based on the data set in the studies were similar, and how much the collected data explained the research questions. Thus, it was tried to evaluate whether there was a lack of standardization or the comparability of the studies in the studies. In addition, the size of the data obtained from social media was tried to be evaluated within the scope of the study.

In the first part of this study, it is evaluated as much as possible what research methods can be used and whether big data is a threat to classical research methods. In the second phase of the study, 47 academic and Turkish articles published in the field of social sciences between 2013 and 2022 on women's research in Turkey and using big data as research data were examined. These 47 studies were determined through Dergi Park and Google Scholar by selecting only the keywords social media, alternative media, new media, women's and women's studies. The main reason for this, as mentioned in the literature section, is that there are many studies using big data, especially in recent years. It is not possible in terms of time to evaluate all of these studies. For this reason, the number of samples was tried to be made reasonable. Since the oldest article in the search engines with the keywords determined in the process of creating the sample starts from 2013, the sample set covers the years 2013 -2022.

In the articles examined, it was seen that the data collected from social media, the analysis of text content or multimedia contents, user profile pages, network data or interactions that can be followed such as likes and shares consisted of. The gathered data may originate from various web pages, as well as social media platforms such as Facebook, Twitter, YouTube, or Instagram. A limited number of studies have shown that the data is obtained using Twitter's programming interface (API).

In many of the studies conducted, it was determined that social media data was obtained in a less structured format. Some researchers are thought to have created a data set by manually copying and pasting selected text content from social media platforms, users' or followers' shares. Some researchers appear to have taken into account the content of users' profile pages or shared photos and collected screenshots to archive images within their research to create a

dataset. This shows that researchers working in the field of social sciences are faced with a lack of technical knowledge in the approach they choose to collect data. The same deficiency is manifested in the analysis of the data. For this reason, it is seen that content analyzes cannot be gone beyond the studies carried out.

1.Giriş

Son on yılda sosyal medya ekonomik, politik ve toplumsal açılardan günlük yaşamın ayrılmaz bir parçası haline gelmiştir. Toplumsal yaşamda geleneksel medyanın etkisi azalırken sosyal medya platformlarının dünya çapında benzeri görülmemiş bir hızla gündelik yaşamın içerisinde kendine yer bulduğuna tanıklık edilmektedir. Boyd (2015), sırf bu nedenle bile sosyal medya olgusunu analiz etmenin zorunlu olduğunu belirtmiştir. Sosyal medya Twitter, Instagram ve Facebook gibi çeşitli sitelerden toplanabilen bilgilerin büyüklüğü, çeşitliliği ve elde edilme hızı nedeniyle genellikle büyük veri kavramı ile ilişkilendirilmektedir (Kitchin, 2014). Gönderiler, beğeniler, takip edilen hesaplar ve takipçiler aracılığıyla elde edilen bağlantılar, kullanıcı tarafından oluşturulan içerikten elde edilen veriler, siyaset sağlık, iş dünyası gibi birçok farklı alandan bilgiye erişilmesine imkân tanımaktadır. Bu kadar büyük miktarda verinin analizi emsalsizdir. Fakat bu denli büyük bir veri seti beraberinde etik hususları, donanım kısıtlamalarını, veri toplama ve analizi için yazılım geliştirme de dahil olmak üzere birçok zorluğu da beraberinde getireceği düşünülmektedir (McCay-Peet & Quan-Haase, 2016). Sosyal medyadan elde edilen verilerle büyük ölçekli analizlerin yanı sıra küçük ölçekli incelemeler de yapılabilmektedir. Şimdiye kadar sosyal medya araştırmaları bir disiplin olarak değerlendirilmemiştir (Mayr & Weller, 2016). Çeşitli disiplinlerdeki araştırmacılar, sosyal medya platformları ve kullanıcıları ile ilgilenmiş ve farklı araştırma sorularına odaklanarak neyin sosyal medya araştırması sayıldığı hakkında kendi tanımlamalarını yapmışlardır (Mayr & Weller, 2016). Bu durum, sosyal medya çalışmalarında bir dereceye kadar avantaj olarak değerlendirilmektedir (Kitchin, 2016). Zira bu durum yeni araştırma sorularına yönelik yaklaşımları keşfetmeye olanak sağladığı gibi farklı çalışma alanlarındaki araştırmacılar için sosyal medyayı heyecan verici bir konu /mecra haline getirmeye yardımcı olmaktadır (Kinder-Kurlanda & Weller, 2014). Ancak Mayr ve Weller (2017) sosyal medya verilerine dayalı yapılan çalışmaların konu ve yöntem açısından çeşitliliğe neden olmasının yanı sıra, yapılan çalışmalarda standardizasyon eksikliği doğurduğuna ya da çalışmaların karşılaştırılabilirliğine ilişkin zorlukları beraberinde getireceğine de dikkat çekmektedir.

Bu çalışmanın odak noktası Türkiye’de kadın çalışmaları alanında doğrudan sosyal medya platformlarından toplanan veri kümelerine dayanan araştırmalardır. Büyük verinin kullanılması kadın çalışmalarında da sık görülür hale gelmiştir. Ancak yapılan çalışmalarda sosyal medyadan elde edilen verinin nasıl elde edildiği, ne kadar büyük olduğu, nasıl analiz edildiği dolayısıyla büyük veri kullanılarak oluşturulan çalışmaların alana hangi araştırma soruları ve yöntemleri kapsamında nasıl bir katkı sağladığına dair bir araştırmaya bu çalışmanın hazırlanma aşamasında ilgili alan dizinde rastlanmamıştır. Çalışmanın amacı iletişim, yeni medya, sosyoloji ve psikoloji gibi sosyal bilimler alanında büyük veri kullanılarak yapılan ve kadına ilişkin çeşitli soruları merkezine alan çalışmalara ilişkin bir çerçeve oluşturmaktır. Bu sebeple de çalışmanın ilgili literatüre mütevazı bir katkı sağlaması hedeflenmektedir. Bu amaçla, 2013 – 2022 yılları arasında yapılan çalışmalarda, çalışma tasarımlarının ve veri setlerinin elde edilmesinde nasıl bir yöntem izlendiği ortaya çıkartılmaya çalışılmıştır. Bu çalışma kapsamında

belirlenen yıllar arasında yapılan çalışmalarda verinin nasıl elde edildiği, verilerin toplanma süresi, toplanan veri setinin boyutu, veri toplama araçlarının birbirinden nasıl farklılık gösterdiği, toplanan verilerle yapılan analizlerin neler olduğu, çalışmalardaki veri setine dayalı kısıtların benzerlik gösterip göstermediği, toplanan verilerin araştırma sorularını ne kadar açıkladığı incelenmiştir. Böylece Mayr ve Weller'in (2017) vurguladığı gibi yapılan çalışmalarda standardizasyon eksikliği ya da çalışmaların karşılaştırılabilirliğinin söz konusu olup olmadığı değerlendirilmeye çalışılmıştır. Seçilen çalışmalarda sosyal medyadan elde edilen verilerin ne denli büyük olduğu da çalışma kapsamında değerlendirilmeye çalışılmıştır.

2. Yazın İncelemesi

2.1. Büyük Veri Kavramı

Büyük veri kavramı, farklı bağlamlarda moda bir kelime haline gelmiş olup genellikle sosyal medya çalışmalarına atıfta bulunmak için kullanılmaktadır. Çeşitli sosyal medya mecralarındaki kullanıcı sayılarının artmasıyla paylaşılan içerik oranları da artmaktadır. Tweet'ler, yorumlar, gönderiler ve fotoğraflar gibi sosyal medya içerikleri, platform sağlayıcılardan veya farklı web sitelerinden kapsamlı bir şekilde büyük verinin oluşturulmasına katkıda bulunmaktadır (Kwon, Lee, & Shin, 2014; Lyu & Kim, 2016). Büyük veri, geniş ölçekli ve farklı veri kümelerinden bilgi toplama, yönetme ve üretmenin yeni yollarını geliştirmek için hem kamu hem de özel girişimler tarafından toplanıp değerlendirilmektedir (Weiss & Zgorski 2012). Her türlü bilimsel, sosyal ve coğrafi veri yeni bilgiler yaratmak ve farklı senaryolar geliştirmek için bir araya getirilmektedir. Örneğin, iklim değişikliğinin kaynak ekonomileri, insan göçü ve siyasi istikrarsızlıklar üzerindeki potansiyel etkisini anlamak için büyük veri kullanılmaktadır (Kitchin, 2016). Bu nedenle, büyük veri, hayatın herhangi bir alanındaki karmaşıklığı anlamak ve karar vermeye yardımcı olmak için farklı bilgi türlerinin birlikte değerlendirilmesi ile ilgilidir. Bu yeni paradigma ile, sosyal hayatın yönetiminin yeniden düzenlendiğine tanıklık edilmektedir. Diğer bir ifadeyle, büyük verinin analizi bir dizi gerçek dünya sorununu çözenin yolu olarak tanıtılmaktadır (Crawford, Milner & Gray 2014). Büyük veri analizinin çeşitli sorunları çözdüğüne inan araştırmacılar büyük veri analizinin devasa veri tabanları içindeki kalıpları ayırt ederek ve çeşitli veri tabanları arasındaki korelasyonları vurgulayarak karmaşıklığın anlaşılmasına yardımcı olduğunu belirtmektedir (Mayer-Schönberger & Cukier, 2013). Bu verilerin analiz edilmesi sayesinde geleceğe dair senaryolar görselleştirilirken çıkabilecek sorunlar önceden tahmin edilmeye çalışılmaktadır (Langlois, Redden & Elmer, 2015). Ancak, sosyal medya araştırmaları bağlamında neyin "büyük" sayıldığına dair ortak bir tanım bulunmamaktadır (Schroeder, 2014). Bir veri setini büyük yapan unsurun o veri setinin büyük kısmının kullanılmasına bağlı olduğu ileri sürülmektedir (Işıklı, 2014). Yapılan analizler için kullanıcı hesapları, bir ağdaki düğümler, tweet'ler veya Facebook, Instagram gönderileri gibi içerik birimleri, beğeniler veya görüntülemeler gibi eylemlere dair birim sayısının dikkate alınması, verilerin işlenmesi için gereken depolamanın boyutundan (örneğin gigabayt veya terabayt cinsinden) daha yaygındır (Mayer-Schönberger & Cukier, 2013). Sosyal medya verileriyle uğraşırken verilerin depolaması ve analiz edilmesine ilişkin soruların dikkatli düşünülmesi gerektiğinin altını çizmektedir. Her ne kadar bir veri setinin ne kadar büyük olduğuna ilişkin tartışmalar olsa da büyük verinin çeşitlilik (variety), hız (velocity), büyüklük (volume), doğrulama (verification), ve değer (value) gibi özelliklere sahip olması bu veri türünün birçok disiplin açısından çekici hale gelmesine neden

olmaktadır (Işıklı, 2014). Bu da araştırmacılar için büyük fırsatlar yaratmaklar birlikte yeni zorlukları da beraberinde getirmektedir (Kitchin, 2014).

2.2. Büyük Veri Klasik Araştırma Teorisinin Sonu mu?

Thomas Kuhn'un (1995) bilimin periyodik olarak baskın bir paradigmadan diğerine nasıl dönüştüğüne dair açıklamasında, yerleşik bir bilgi birikimine yeni bir dizi fikir seti tarafından meydan okunduğunu, istikrarsızlaştırıldığını ve sonunda da ikincisinin birincinin yerini aldığı bir devrilme noktasına ulaştığı anlatılmaktadır. Newton'un fizik yasalarından Einstein'ın görelilik teorisine geçiş bu açıklama için bir örnek niteliğindedir. Kuhn'un açıklamasında, bir paradigma kayması meydana gelmektedir. Zira, baskın bilim tarzı belirli fenomenleri açıklayamaz hale gelmekte veya kilit soruları cevaplayamamaktadır. Bununla birlikte, Herring (2004) paradigmlar arasındaki geçişlerin veri üretimindeki ilerlemeler ve yeni analitik yöntemlerin geliştirilmesi ile de ortaya çıkacağını belirtmektedir. Bu görüşün dayanağı ise "bilimdeki [d]evrimlerin genellikle ölçümdeki devrimlerden önce geldiği" gözlemdir (Mayer-Schönberger & Cukier, 2013). Böylece bilimin, büyük veri ve yeni analitiklerin artan kullanılabilirliğine dayalı dördüncü bir paradigmanın (keşif bilimi) ortaya çıktığını öne sürmektedir (Hey vd., 2009). Bu açıklamanın doğrultusunda, büyük verilerin akademik araştırmaların yürütülme biçiminde bir devrim başlatıp başlatmadığını düşünmenin faydalı olacağına inanılmaktadır. Bugün çok farklı alanlarda araştırmacılar ve akademisyenler için çekici hale gelen büyük verinin aslında klasik araştırma teorilerinin ve yöntemlerinin sonunu getirip getirmediği, verilerin kendi adına konuşabileceği yeni bir ampirizm biçimini mümkün kılıp kılmadığı, bilgi odaklı bilimden ziyade veri odaklı bilimin yaratılmasının aslında yeni bir bilgi üretimi çağını ve veri analitiği çağını işaret ettiği belirtilmektedir. Anderson (2008) ise büyük verilerin içerdiği örüntüler, ilişkiler, fenomenlerin anlamlı ve iç görülü bilgi üreten veri tufanının bilimsel yöntemi geçersiz kıldığını savunmaktadır. Prensky (2009) ise araştırmacıların artık hipotezler ve modeller oluşturmak ve bunları veriye dayalı deneyler ve örneklerle modelleyip tahminler yapmak zorunda olmadıklarını dile getirmiştir. Prensky 2009'daki çalışmasında, araştırmacıların daha fazla deney yapmadan bilimsel sonuçlar üretmek etkileri ortaya çıkaran kalıplar için tüm veri setini elde edebileceklerini belirtmektedir. Bu yeni veri kümesine ilişkin Dyche (2012) ise büyük veri madenciliğinin aramayı bile bilmediğimiz ilişkileri ve kalıpları ortaya çıkardığını dile getirmektedir.

Çoğu yeni olan ve şimdiye kadar erişilmesi zor olan fotoğraflar, paylaşımlar, çeşitli eserler, maddi nesnelere vb. ulaşılmasına imkân tanınması büyük verinin yarattığı fırsatlar olarak değerlendirilmektedir (Cohen, 2008). Bu da çok sayıda verinin işleyebilmesini, yeni veri analiz araçlarının ve yöntemlerinin gelişmesini sağlamıştır. Böylece çok sayıda ilgili yapıtı araştırmak ve aralarında bağlantı kurmak mümkün hale gelmiştir (Kitchin, 2016). Yeni teknikleri savunan araştırmacılar, (Kelling vd., 2009; Manovich, 2011) büyük verinin artması sonucu gelişen yeni araştırma yöntemlerinin mevcut beşerî bilimlere ilişkin yöntemleri güçlendireceğini ileri sürmektedir. Söz konusu yeni gelişmelerin geleneksel yorumlama ve teori oluşturma biçimlerini kolaylaştırdığı, çok daha geniş kapsamlı çalışmalara olanak tanıdığı ve hesaplama olmadan neredeyse imkânsız olan soruları yanıtlamaya olanak tanıdığı iddia etmektedir. Başka bir deyişle, büyük veri analitiğini reddetmeyen bunun yerine sosyal bilimcilerin büyük veriden değerli içgörüler elde etmelerini sağlayan farklı bir epistemolojik çerçeve içinde veriye dayalı bilimin metodolojik yaklaşımını kullanan yeni epistemolojiler üzerinde düşünmenin mümkün olduğu ileri sürülmektedir.

Bununla birlikte, büyük veriye dayalı bilim, bilimsel yöntemlerin ilkelerine bağlı kalmaya çalışmasının yanı sıra bir fenomenin anlaşılmasını ilerletmek için tümevarım ve tümdengelim yaklaşımlarının melez bir kombinasyonunu kullanmaya daha açıktır (Kitchin, 2016). Büyük veri kullanılarak yapılan analizler, deneysel tümdengelim tasarımından farklıdır. Çünkü "teoriden doğmak" yerine "verilerden doğan" hipotezler ve içgörüler üretmeye çalışılmaktadır (Kelling vd., 2009; Kitchin, 2016). Buraya kadar yapılan açıklamalar büyük veriye ilişkin madalyonun olumlu taraflarını ortaya koymaktadır. Fakat, söz konusu yeni veri kümesi ve buna bağlı araştırma yöntemlerine ilişkin kısıtların neler olduğu da konuyla ilgili alan dizinde yer almaktadır. Bu kısıtların neler olduğundan da bahsetmenin önemli olduğu düşünülmektedir. Örneğin, sosyal medya verilerine ulaşmada kullanılan teknoloji ve bunların çalışma parametreleri, verilerin doğasını etkileyebilmektedir. Diğer bir ifadeyle, sosyal medyada en çok hangi gönderilerin okunduğu veya paylaşıldığı, sadece ilgi değil, sıralama algoritmalarından da güçlü bir şekilde etkilenmektedir (Baym, 2013). Benzer şekilde, API'ler hangi verilerin çıkarılacağını yapılandırmaktadır. Örneğin Twitter'da ilgili tüm tweet'ler yerine yalnızca bir olayla ilişkili belirli etiketler (hashtag) incelenebilmektedir (Kitchin, 2016). Twitter verilerine erişmenin farklı yöntemlerinin (arama API'leri ile akış API'leri) oldukça değişik sonuç kümeleri ürettiği bulunmuştur (González-Bailón vd., 2012). Sonuç olarak, aynı verileri aynı anda toplamaya çalışan iki araştırmacının aynı veri kümesini elde edeceklerinin garantisi bulunmamaktadır (Bruns, 2013). Güvenilirlikle ilgili olarak ise sosyal medya paylaşımlarının insanların görüşlerini ne ölçüde temsil ettiği konusunda da soru işaretleri bulunmaktadır (Kitchin, 2016). Manovich (2011) paylaşılan içeriklerin, yüklenen fotoğrafların, yapılan yorumların ve diğer çevrimiçi katılım türlerinin şeffaf pencereler olmadığı konusunda uyarıda bulunmaktadır. Edwards ve arkadaşları (2013) Twitter'da paylaşımda bulunan kişilerin kimler olduğunun bilinmediğini belirtmektedir. Dolayısıyla yapılan araştırmalardaki seçilen paylaşımların belirli grupları doğru temsil edip etmediğinin net olmadığına çalışılmaktadır. Sistemler belirli türdeki verileri yakalamak için tasarlanmıştır ve kullanılan analitikler ve algoritmalar bilimsel akıl yürütmeye dayalıdır ve bilimsel testlerle iyileştirilmiştir (Berry, 2011). Sonuç olarak, veri elde etmek için kullanılan teknoloji ve platformlar, verileri şekillendirdiği için büyük veri kümesinin içinden seçilen örneklemin kapsam açısından yeterli olmadığı gibi örneklemin yanlı olduğundan da bahsedilmektedir (Crawford, 2013; Kitchin, 2013). Yukarıdaki açıklamalar doğrultusunda konunun daha anlaşılır olması için sosyal medyadan büyük verilerin elde edilme sürecinde dikkat edilmesi gereken noktalardan bahsetmenin faydalı olacağına inanılmaktadır. Aşağıda belirtilenlerle sınırlı olmamak üzere büyük verilerin toplamasına ilişkin dikkat edilmesi gereken noktalar sıralanmıştır.

Sosyal medya bağlamındaki büyük veri analiz teknikleri duygu analizi, sosyal ağ analizi, metin madenciliği gibi çeşitli yöntemlerle değerlendirilebilmektedir. Bunun için paylaşımda bulunan kişiler URL'ler aracılığıyla Twitter, Instagram, Facebook gibi farklı sosyal medya siteleri arasında bağlantı kurulmasını sağlamaktadır. Diğer bir ifadeyle Instagram gönderileri YouTube videolarına atıfta bulunabildiği gibi Wikipedia makaleleri blog gönderilerine atıfta bulunabilir ya da Meme'ler bir sosyal medya platformundan diğerine yayılabilir. Araştırılan birçok konu için büyük resim ancak birden fazla sosyal medya platformundan gelen veriler dahil edildiğinde görünür hale gelmektedir (Mayr & Weller, 2016). Öte yandan, sosyal medya kullanıcıları farklı ihtiyaçlar ya da

tercihlerden ötürü bir platformu diğerine tercih edebilmektedir (Quan-Haase & Young, 2010). Bu da farklı iletişim türleri için değişik sosyal medya platformlarının kullanılabilmesi ve bazı platformların belirli konuları çalışmak için diğerlerinden daha uygun olabileceği anlamına gelmektedir. Örneğin, Twitter'da ağırlıklı olarak düşünce, duygu vs. yazılı olarak ifade edilirken Instagram'da kişiler görseller aracılığıyla kendilerini ifade etmektedir.

Belirli bir olay (seçimler veya spor etkinlikleri gibi) veya bir grup insan tarafından tartışılan ve sosyal medyada en çok konuşulan konulara (*trend topic - tt*) dayalı olarak içeriklerin toplanması da sık kullanılan uygulamalar arasında yer almaktadır. Özellikle Twitter'da, güncel tartışmalar genellikle belirli etiketler (*hashtag*) kullanılarak yürütülmektedir. Ancak diğer platformlar da kullanıcıların etiketler veya anahtar kelimeler gibi içerik açıklayıcı meta verileri uygulamalarına olanak tanımaktadır. Bunlar, sosyal medyadaki belirli verilerin aranması ve toplanması için bir kriter olarak kullanılmaktadır. Ancak, çoğu zaman metin tabanlı veri toplama bütünlüğe ulaşmak zordur. Zira, insanlar aynı konuya atıfta bulunmak için farklı kelimeler kullanabilir veya konu, çok kısa gönderilerde açıkça belirtilmeyebilir. Örneğin, Twitter'da güncel bir olay hakkında yorum yapılırken bir veya daha fazla sayıda etiket kullanılabilir (ör. #FIFA, #Dünya Kupası, #WorldCup) ya da farklı etiketler kullanılabilir veya herhangi bir etiket kullanmadan doğrudan olaydan bahseden içerikler de paylaşılabilir. Bu sebeple anahtar kelimelere veya diğer tam metin aramalarına dayalı veri toplama yaklaşımı kullanılırken, arama yapılacak etiket(leri) ve alternatifleri dikkate almak önemlidir (Mayr & Weller, 2016). Bazı durumlarda veriler, metaveri olarak adlandırılan çeşitli yapısal kriterlere göre de toplanabilmektedir. Bu, anahtar kelimeler ya da etiketlere dayalı olmayan ama hesaplardaki içeriklere dayalı tüm bilgileri yansıtmaktadır. Örneğin, bir hesap içerisinde yer alan coğrafi konum, zaman dilimi (ör. pazar günü yayınlanan tüm durum güncellemeleri), dil gibi. Ancak hesap sahiplerinin bu bilgileri doğru girmemesi ya da hiç girmemesi durumunda çalışmada istenilen veya beklenen nitelikte sonuçlara ulaşılamamaktadır.

Belirli bir veri seti ile başlamak ve araştırma sorularını onun etrafında oluşturmak, bazı durumlarda çok anlamlı olabilmektedir (Hogan & Quan-Haase, 2010). Bu keşif tasarımı, belirli bir sosyal medya platformunun özelliklerinin haritasını çıkarmak için faydalı olabileceği gibi belirli bir platformdaki kullanıcı davranışlarının anlaşılabilmesi ile ilgili durumlarda da uygulanabilmektedir (Weninger vd., 2013). Ancak çoğu araştırma sürecinde spesifik araştırma sorusuyla başlanması ve ardından bu soruyu yanıtlamak için gerekli olan ideal veri kümesinin oluşturulması tavsiye edilmektedir (McCay-Peet & Quan-Haase, 2016). Zira veri toplama sürecinde ortaya çıkabilecek olan teknik, yasal veya etik sınırlamalar nedeniyle ideal veri seti ya da elde edilmek istenilen veri setine ulaşmak mümkün olmayabilir. Toplanacak olan verilerin hangi zaman dilimini kapsayacağı da sonuçları büyük oranda etkilemektedir (Jungherr, 2016). Bu sebeple toplanacak olan büyük verinin hangi zaman aralığını kapsamaması gerektiğinin belirlenmesi gerekmektedir. Çeşitli sosyal medya hesapları üzerinden elde edilmeye çalışılan verilerde sistemin istenilen tarihlerde bilgiye ulaşmaya izin vermemesi, çalışmanın sonuçlarını olumsuz etkileyecek unsurlar arasında yer almaktadır. İdeal veri kümesine ilişkin değerlendirmeler sosyal medyanın gerçekten mümkün olan en iyi veri kaynağını sağlayıp sağlamayacağını veya diğer veri toplama yöntemlerinin (örneğin deneyler, anket verileri, geleneksel kitle iletişim araçlarına dayalı ikincil kaynaklar vb.) daha uygun olup olmayacağı sorusu da önemli bir noktadır.

Klasik anlamda araştırma sorusunun oluşturulması, hipotezlerin belirlenmesi, yöntemin seçilmesi gibi klasik anlamdaki araştırma yöntemlerinin yanı sıra büyük veri kümeleri de artık teknolojik gelişmelerle birlikte hayatımızın içerisinde yer almaktadır. Bu çalışma kapsamında bu aşamaya kadar hızla gelişen büyük veri setlerinin araştırma yöntemleri arasında yer almasını sağlayan ama bir taratan da kısıt olarak değerlendirilen noktalarından bahsedilmiştir. Bundan sonraki bölümde ise büyük veri ile yapılan ve merkezinde kadınlara ilişkin çeşitli sorunları analiz eden makalelerin, sosyal medya başta olmak üzere çalışma alanları ile ilgili elde ettikleri verileri analiz etmede ne denli etkin olduğu içerik analizi kullanılarak değerlendirilmeye çalışılmıştır.

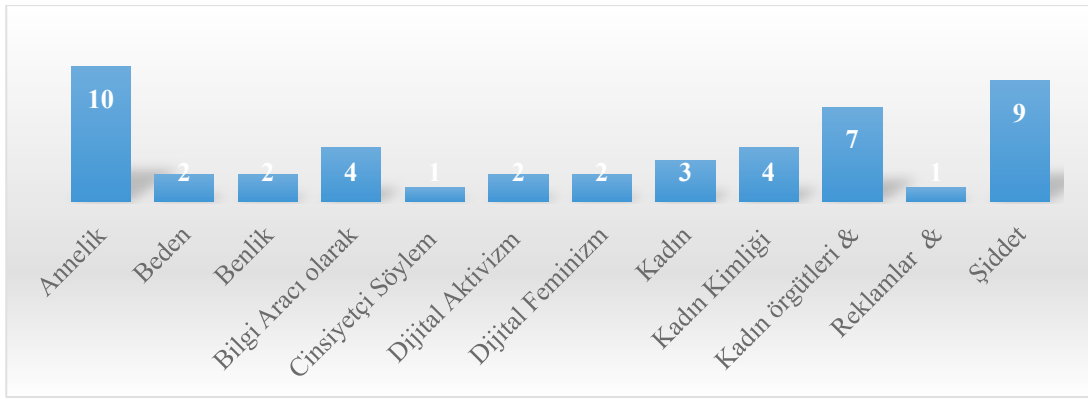
3.Yöntem

Bu çalışmanın araştırma sorusu Türkiye'deki sosyal bilimler alanında akademik dergilerde yayınlanan, kadını merkeze alan ve veri seti olarak büyük veriyi kullanan çalışmaların büyük veri setine nasıl ulaştıklarını, buradan elde ettikleri verileri nasıl analiz ettikleri, çalışmalarının geçerlilik ve güvenilirliklerini nasıl sağladıkları gibi noktaları ortaya çıkarmaya dayanmaktadır. Bu çalışma kapsamında değerlendirilen makaleler tarafsız şekilde, doğrudan araştırma sorusuna yönelik olarak belirlenmiştir. Çalışma kapsamında TR dizin ya da yabancı indekslerde taranan dergilerde yayınlanan makaleler dikkate alınmıştır. Örneklem hacmini sınırlandırmak için "sosyal medya, alternatif medya, yeni medya, kadın, kadın çalışmaları" anahtar kelimelerinin birkaçını birlikte kullanan makaleler dikkate alınmıştır. Söz konusu anahtar kelimeler Dergi Park ve Google Scholar'da yayınlanan makaleler arasından seçilmiştir. Örneklemenin oluşturulması sürecinde belirlenen anahtar kelimelerle söz konusu arama motorlarında en geri tarihli makale 2013'ten başladığı için örneklem kümesi 2013- 2022 yılları arasında kapsamaktadır. Yapılan literatür taraması sonucunda belirlenen anahtar kelimelerle 2013 – 2022 yılları arasında iletişim, yeni medya, sosyoloji ve psikoloji alanlarında yayınlanan ve araştırma verisi olarak büyük veri kullanan 47 akademik ve Türkçe makale olduğu tespit edilmiştir. Böylece örneklem hacmi makul bir büyüklüğe getirilmeye çalışılmıştır.

Bu çalışma kapsamında oluşturulan örneklemi analiz etmek için içerik analizi yönteminden yararlanılmıştır. İçerik analizi, değişik formatlarda (örneğin, yazılı, işitsel veya görsel gibi) oluşturulmuş metinlerden anlamlar çıkarmak için kullanılan bir yöntemdir. İçerik analizi yönteminden, özellikle, öne çıkan kalıpları nesnel, sistematik, tümdengelim dayalı bir şekilde önceden belirlenen kriterler dikkate alınarak değerlendirmek için faydalanılır (Metin & Ünal, 2022). Söz konusu bu yöntemdeki temel amaç elde edilen veri setindeki bilgiyi sistematik olarak azaltmak olup (Tavşancıl & Aslan, 2001) bilginin içeriğine odaklanılır. Yapılan bu çalışma kapsamında belirlenen temel sorular dikkate alınarak, elde edilen veri seti ayrı ayrı değerlendirmiş ve ortaya çıkan kodlar karşılaştırmıştır. İçerik analizlerinde veri setinin içerisinde belirlenen önemli temalardan kategoriler oluşturularak araştırma sorusuna ilişkin çeşitli betimlemeler yapılabilmektedir (Metin ve Ünal, 2022). Çalışma kapsamında genel değerlendirmelerde bulunulmuştur. Yapılan analizden elde edilen sonuçların, gelecekte yapılacak çalışmalara yön göstermesi hedeflenmektedir.

4.Bulgular ve Tartışma

Seçilen anahtar kelimelerle incelenen makalelerin toplandığı konu başlıkları Şekil 1’de gösterilmiştir. Dikkate alınan 47 makalenin 12 farklı başlık etrafında toplandığı belirlenmiştir. Buna göre ilgili dönem içerisinde en çok araştırılan konunun annelik başlığı olduğu görülmektedir. Annelik kavramından sonra en çok çalışılan konu başlığının ise kadına yönelik şiddet olduğu görülmektedir. Üçüncü en çok çalışılan başlığın ise Türkiye’deki kadın örgütleri ve derneklerinin sosyal medyayı ne denli etkin kullandığı konusunda yapılan çalışmalar olarak belirlenmiştir.

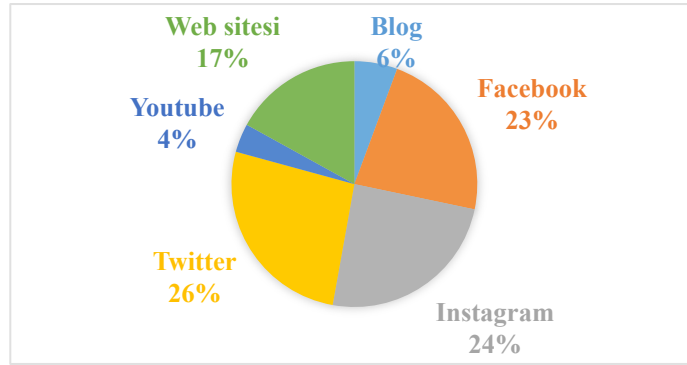


Şekil 1. Çalışılan Konu Başlıklarının Dağılımı

İncelenen makalelerin belli başlıklarda toplanmasından ziyade dikkati çeken önemli noktalardan ilki genelde belirlenen sosyal medya aracının nasıl ve ne denli etkin kullanıldığı yönündeki araştırma sorusunun çevresinde çalışmaların yapılmış olmasıdır. Özellikle kadın örgütleri ve dernekleri ile ilgili yapılmış çalışmaların sosyal medyayı ne kadar etkin veya hangi amaçlarla kullandıkları yönünde bir eksende toplandıkları görülmektedir. Şiddet başlığı altında yapılan çalışmaların da aslında sosyal medyanın hak arayışı için doğru bir mecra olup olmadığı sorgusuna dayalı olduğu görülmüştür. Yapılan incelemeler sırasında dikkat çeken ikinci nokta ise benzer konu başlıklarının benzer alan dizin taramasına dayanmasıdır. Her ne kadar bunda şaşırtıcı bir durum olmasa da çalışmaların literatür taramalarının genelde yabancı dizine atıfta bulunduğu ve benzer olmasına rağmen pek birbirlerinin farkına varmadığı görülmüştür. Çalışmaların birçoğunda gerek yabancı gerekse de yerli alan dizide yapılmış benzer çalışmalara ilişkin karşılaştırmalı bir değerlendirmeye yer verilmediği, dolayısıyla çalışmaların durum değerlendirmesi ile sınırlı kaldığı gözlemlenmiştir.

Dikkate alınan çalışmalarda Twitter, Instagram ve Facebook’un en çok veri elde edilen sosyal medya mecraları olduğu belirlenmiştir. Şekil 2’de de görüldüğü üzere söz konusu sosyal medya araçlarının kullanım dağılımları birbirine çok yakındır. Ancak bazı çalışmalarda her iki ya da üç sosyal medya sitesi veri elde etmek adına birlikte kullanılmıştır. Yapılan analizler sonucunda dikkat çeken bir nokta ise 2020 ya da 2021 çalışmalarında hala Facebook’un veri seti oluşturmada tek sosyal medya mecrası olarak kullanılması ve bunun nedeninin çalışmalar içerisinde belirtilmemiş olmasıdır.

Şekil 2’de “web siteleri” başlığında kadın derneklerinin web sitelerinin yanı sıra Ekşi sözlük gibi sitelerin ya da incelenen reklamların yer aldığı sitelerin bulunduğu görülmüştür.



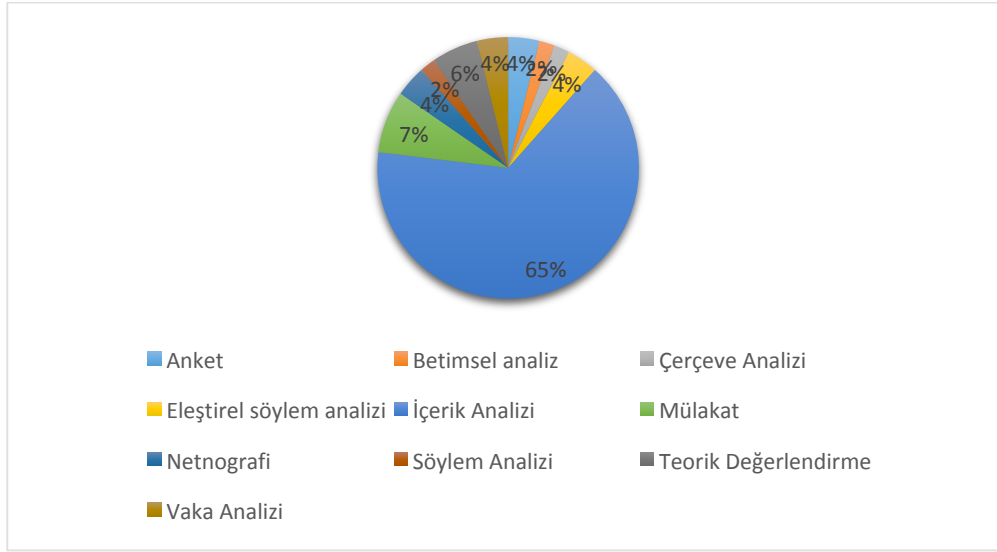
Şekil 2. Sosyal Medya Veri Kaynakları

Özellikle sosyal medyanın gelişmesi ve buna bağlı olarak annelik kavramının nasıl değiştiği çalışmaların içerisinde yer alan konu başlıkları arasındadır. Bu bağlamda blogger annelerin açtıkları hesaplar ve paylaştıkları bilgiler alanda çalışan kişiler için bilgi toplanacak kaynaklar arasında yer almıştır.

Önemli bir başka nokta ise örnekleme yer alan kimi makalelerde sosyal medya kavramı ile anlatılmak istenilenin ne olduğunun anlaşılmasındadır. Örneğin, Ersoy (2018) “Sosyal Medya Kullanımının Annelerin Annelik Rolü Üzerindeki Etkisi” adlı çalışmasında anne olan kadınların sosyal medya kullanımlarının annelik kimlikleri üzerine etkisini araştırmıştır. Çalışma her ne kadar sosyal medya kullanıcısı olan ve olmayan 5’er anne ile yapılan mülakata dayansa da söz konusu çalışmada sosyal medya ile kastedilenin ne olduğunun örneklem için netleştirilmemiş olmasıdır. Diğer bir ifadeyle mülakat yapılan kadınlar hangi sosyal medya mecralarını, ne sıklıkla kullanıyor açıklanmamıştır. Her ne kadar bu çalışmalar da alanın gelişmesine katkı sağlayan önemli ve değerli işler olsalar da okuyucunun kafasında soru işareti oluşmaktadır.

Sosyal medya verisine dayalı yapılan çalışmalardaki en önemli noktalardan biri verinin çok hızlı şekilde eskiebilmesidir. Zira gündem çok hızlı değişmekte, öne çıkan paylaşımlar her gün her saat farklılaşmaktadır. Bu sebeple kullanılan verilerin ve verilerin elde edildiği mecraların eski olmaması konuların güncelliği açısından önem kazanmaktadır. Huzziyas ve Nizam’ın 2021’de yayınlanan “Özgecan Aslan Cinayeti Sonrası Sosyal Medyada Dayanışma Olgusu” adlı makalesi Türk halkının hafızasından hala silinmemiş olan Özgecan Aslan cinayeti ile ilgili önemli bir çalışma olmakla birlikte, kullanılan sosyal medya verisi 2015 – 2016 yıllarında Facebook üzerinden elde edilmiş olan hesap değerlendirmelerine dayanmaktadır. Bugün Facebook’un güncelliğinin hangi yaş grubu için geçerli olduğunun sorgulandığı bir dönemde Facebook yerine farklı bir mecra kullanılsa nasıl olurdu ya da farklı sosyal medya araçlarından elde edilen veri setleri kullanılarak aynı başlık değerlendirilse nasıl farklılıklar tespit edilirdi sorusu akıllara gelmektedir. Örneğin, bugün Instagram ve Tik Tok’un da toplumsal hareketler açısından bir araç haline geldiğine tanıklık edilmektedir. Bu mecralarda yapılan içerik paylaşımlarının, geçmişten bugüne,

dilinin, şeklinin nasıl değiştiğine dair bir değerlendirme daha mı güncel bir yorum elde edilmesine neden olur muydu diye düşünülmektedir.



Şekil 3. Kullanılan Analiz Türleri

Örneklem olarak incelenen çalışmalarda içerik analizinin en çok tercih edilen yöntem olduğu görülmektedir. İçerik analizinin İngilizce dizindeki karşılığı metin madenciliği (text mining) olarak geçmektedir. Yapılan çalışmaların %47'si bu analiz türü ile değerlendirilmiştir. İçerik analizinden sonra mülakat da en çok kullanılan yöntem olarak görülmektedir (Bakınız Şekil 3). Fakat incelenen çalışmalar içerisinde mülakat araştırma tekniği tek başına kullanılmamıştır. Aynı durum teorik değerlendirme için de geçerlidir. Özellikle kimlik, benlik gibi konu başlıklarının öncelikle teorik bir değerlendirme ile başladığı sonrasında bunu destekler nitelikteki ikinci bir analiz kullanıldığı görülmüştür. Genellikle içerik analizi ve mülakat şeklinde hem nitel hem de nicel çalışmalar birbirini destekler nitelikte kullanılmıştır. Aynı durum anketler içinde geçerlidir. Zira seçilen anahtar kelimelere göre dikkate alınmış ancak büyük veriyi kullanmadan sadece anket ya da mülakat yapmış olan çalışmalar örneklem içerisine dahil edilmemiştir (Örneğin, Çetin & Edis, 2022; Tekin & Turan, 2020; Köse & Akyazı, 2020). Şekil 3'e bakıldığında zaman anket, vaka analizi, netnografi ve eleştirel söylem analizi yapılmış olan çalışmalarının yüzde dağılımlarının aynı olduğu görülmektedir. Yüzde dağılımları betimsel analiz, söylem analizi ve çerçeve analizi için de aynı olarak görülmektedir. Söz konusu bu analiz türleri ise büyük veriden elde edilen veri kümelerinin analiz teknikleri arasında sıklıkla yer alan duygu analizleri (sentiment ve emotion) sosyal ağ analizi gibi yöntemlerle değerlendirilmediklerini göstermektedir.

Seçili makalelerden sadece birisi 2002 – 2018 yılları arasında Instagram'daki kadın temsillerini incelemiştir. Ancak Instagram'da dikkate alınan 16 yıllık süre içerisinde incelenen içerik sayısı 56'dır. İncelenen makalelerden bir diğeri 11 aylık Twitter verisini dikkate almıştır ve toplam 3070 tweet incelemiştir. Çalışmanın verisinin, R programlama dili ile elde edildiği makalede belirtilmiştir. Değerlendirilen makaleler içerisinde sadece ikisi Twitter'dan elde ettikleri veri setini R ile topladığını açıklamıştır. Diğer makalelerin hiç birisinde verilerin nasıl elde edildiği belirtilmemiştir. Verilerin nasıl elde edildiğinin çalışmaların içerisinde yer almaması, toplanan verilerin araştırmacılar tarafından manuel şekilde oluşturulduğunu düşündürmektedir. Her ne kadar bunda yanlış

olan bir taraf olmasa da bu durum beraberinde başka soru işaretlerini doğurmaktadır. Örneğin değerlendirilen birçok çalışmada etkileşim oranları tekrar tweet (retweet), beğeni (like) ve yapılan yorum sayılarına göre değerlendirilmiştir. Verilerin manuel olarak toplanması yapılan paylaşımların düzenli olarak kontrol edilmesini gerekmektedir. Bu da çalışmanın güvenilirliği açısından soru işareti yaratacak bir etkene dönüşmektedir.

Dikkate alınan makalelerden üçünün veri setinin bir yıllık bir süreyi kapsadığı görülmektedir. Araştırmacıların belirlediği bu süre içerisinde incelenen paylaşım sayılarının 1382, 764 ve 192 olduğu belirlenmiştir. Bu da aslında elde edilen veri kümesinin ne kadar büyük olduğu sorgulamasını yapmamıza neden olmaktadır. Seçili makalelerden birisi Twitter'dan elde ettiği 3 günlük veri içerisinde 90 cinsiyetçi söyleme dayalı tweet içeriği incelediğini belirtirken, bir diğeri gene sosyal medyada 8 gün boyunca yapılan 189 paylaşımı inceleyerek çalışmanın analizini tamamlamıştır. Bu iki çalışma incelenen zaman aralığı açısından en az zaman dilimini içeren çalışmalar olarak belirlenmiştir. Değerlendirilen çalışmalardan üçünün, bu çalışma kapsamında belirlenen anahtar kelimeleri kullanarak sadece teoriye dayalı değerlendirmede buldukları ve herhangi bir veri analizinin yapmadıkları görülmüştür. İncelenen makalelerden 11'inin veri setini 1 ay, birinin 2 ay, ikisinin 3 ay, birinin 4 ay, birinin 8 ay ve ikisinin 7 ay içinde verilerini topladıkları belirlenmiştir. Bu açıklama doğrultusunda incelenen makaleler özelinde elde edilen veri kümelerinin genellikle bir aylık bir zaman dilimini içerdikleri görülmektedir. Bunların dışında kalan makalelerin ise sadece inceledikleri video, hesap ya da paylaşımlar üzerinden çalışmalarının yöntemlerini şekillendirdikleri görülmüştür. Bu çalışmalarda dikkate alınan içeriklerin hangi zaman aralığını kapsadığına yer verilmemiştir. İncelenen içerik sayılarının genelde çok az olması dikkat çekmektedir. İncelenen makalelerin ne yazık ki %65'inde seçilen örneklemin neden dikkate alınan sosyal medya hesabından seçildiği, incelenen tarih aralığının neden o aralık olarak belirlendiği gibi detaylara da yer verilmemiştir. Örneklem olarak dikkate alınan çalışmaların %40'unda amaçlı örneklem kullanıldığı görülmüştür. Her ne kadar bu yanlış bir durum olmasa da çalışmanın tarafsızlığı konusunda soru işareti oluşturduğu düşünülmektedir. Zira incelenen görseller ya da analiz edilecek olan açıklamalar doğrudan araştırma sorusuna uygun olarak seçilmektedir. Bu noktada büyük verinin kullanılmasının anlamı ortadan kalkmaktadır. İncelenen makalelerden üçü dışında kalan tümünün büyük veriyi elde etmeden analiz etmeye kadar birçok aşamada yetkin olmadığını düşünülmektedir. Bu sebeple yapılan çalışmalarda söylem ve içerik analizlerinden öteye gidilemediği görülmektedir. Bütün bunlarla birlikte yukarıda bahsedilenlerin dışında gerek kullanılan veri setinin büyüklüğü ve kullanılan analiz yöntemleri gerekse de literatür araştırmaları açısından gerçekten büyük veri kullanılarak yapılmış olan çalışmaların da olduğu görülmüştür. Alikılıç ve Baş (2019), Dursun ve Yıldız (2020) bu makaleler arasında yer alırken, Koç, Bakır-Demir ve Öner-Özkan'ın (2019) makalesi farklı bir bakış açısının nasıl getirilebileceğini gösteren önemli makaleler arasında değerlendirilmiştir. İncelenen makalelerin 37'sinde çalışmaların kısıtlarının neler olduğu ve ilgili konuyla ilişkin ileride yapılabilecek çalışmaların neler olabileceğine ilişkin açıklamaların ise sınırlı olması dikkat çekicidir. Yeri gelmişken, bu araştırmaya ilişkin en temel kısıtlardan biri incelenen örneklem sayısı ve alanı ile ilgilidir. Belirlenen makale sayısının artması daha farklı şekilde incelenmiş makalelerin bulunmasını sağlayabilir. Bununla birlikte çalışma sadece iletişim, yeni medya, sosyoloji ve psikoloji ile sınırlandırılmıştır. Ancak seçilen anahtar kelimelerle iktisat, pazarlama, reklamcılık gibi alanlarda da çalışılmaktadır. Çalışma süresine ilişkin kısıt yüzünden bu alandaki

çalışmalar dikkate alınmamıştır. Dolayısıyla bu çalışma kapsamında belirlenen araştırma ve analiz kısıtlarının farklı disiplinlerde de gözlemlenip gözlemlenmediği bilinmemektedir. Özetlemek gerekirse, incelenen makalelerde sosyal medyadan toplanan veriler, metin içerikleri veya multimedya içeriklerinin analizi, kullanıcı profili sayfaları, ağ verileri ya da beğeniler, paylaşımlar gibi takip edilebilen etkileşimlerden oluştuğu görülmüştür. Elde edilen veriler farklı web sayfalarından, Facebook, Twitter, YouTube, Instagram gibi pek çok platformdan gelen veriler olabilmektedir. Sınırlı sayıdaki çalışmada verilerin, Twitter'ın programlama arabirimi (API) kullanılarak elde edildiği görülmüştür.

Yasal veya teknik kısıtlamalar nedeniyle ideal veri setini elde etmek mümkün olmayabilir. Bu durumda da araştırmacıların ikame bir veri kümesi ile çalışmaktan başka seçeneği bulunmamaktadır. Bu durumun sorun olarak değerlendirilmemesi gerektiği düşünülmektedir. Ancak araştırmacıların, veri toplama yaklaşımlarının potansiyel sınırlamaları konusunda çok daha net olmalarının gerekliliğine inanılmaktadır. Söz konusu kısıtların elde edilen bulgular, sonuçlar açısından neden olabileceği durumlar ana hatlarıyla belirtmelidir. Veri toplama sürecine ilişkin kısıtları bilmek okuyucu açısından önemlidir. Fakat elde edilen veri setine ilişkin araştırmacının sınırların ne olabileceğinin farkında olmasının da bir o kadar önemli olduğuna inanılmaktadır. Örneğin, incelenen çalışmalardan bazılarında Twitter'dan veri elde eden araştırmacılar veri kümelerini belirledikleri etiketlere göre oluşturmuştur. Bu, etiketleri kullanarak spesifik bir olay için veri kümesi oluşturmanın iyi bir yolu olarak değerlendirilebilir. Bununla birlikte bu durumun, belirli kullanıcıları sistematik olarak oluşturulan veri kümesinin dışında tutacağı unutulmamalıdır. Daha önce de belirtildiği gibi etiket kullanımına aşina olmayan veya popüler olan etiketleri kullanmadan seçilen konuda paylaşımda bulunan kişiler örneklemin dışında kalacaktır. Elde edilen veri kümesi ile gerçekleştirilebilecek analiz türlerinin neler olabileceği ve veri setinin çalışmanın üzerindeki etkilerinin bilinmesi de önemli olduğuna inanılmaktadır. Yapılacak çalışma için en uygun sosyal medya kanallarının hangileri olduğu, veri kümesinin hangi zaman aralığında yer alacağına belirlenmesi, yeni oluşturulan kullanıcı hesaplarının takip edilmesi veya belirlenen etiketlerin sürekli kontrol edilmesi, farklı etiketlerin neler olabileceği çalışmaların içerisinde dikkate alınması gereken ve metin içerisinde de okuyucuya aktarılması gereken noktalar olarak öne çıktığı düşünülmektedir.

5.Sonuç

Her geçen gün farklı araçların toplumsal hayatın rutini içerisinde kabul görmesi ve yaygınlaşması nedeniyle daha da zenginleşen büyük veri birçok bilim dalı gibi sosyal ve beşerî bilimcileri epistemolojiyi yeniden değerlendirmeye zorlamaktadır. Birçok şekilde elde edilebilen verilerin farklı disiplin ve alanları kapsar nitelikteki büyüklüğü, klasik araştırma yöntemlerinin sonunu mu getirdi sorusundan ziyade bu veriye nasıl ulaşılır ve elde edilen veri nasıl işlenmelidir sorularının daha önemli olduğu düşünülmektedir. Bu da disiplinler arası aktif bir bilgi akışının ve öğrenme sürecinin ortaya çıkmasına neden olmaktadır. Bugün sosyal bilimciler de büyük veriye ulaşabilmek adına R, Python gibi yazılım dillerini kullanarak veri elde etme ve analiz yöntemlerini öğrenmeye çalışmaktadır. Tabii ki klasik araştırma yöntemleri ile elde edilen veri kümeleri, ki bunlar daha küçüktür, anlamsız ya da önemsiz değildir.

Hatta belki de büyük verinin elde edilmesi ya da analizindeki eksikleri nedeniyle değerlidir ve değerli olmaya devam edecektir. Bu açıklama için demografik verilerin elde edilmesi belki de iyi bir örnek niteliğindedir. Sosyal medya hesaplarına demografik bilgilerin hesap kullanıcıları tarafından, bilinçli ya da bilinçsiz şekilde, keyfi olarak girildiği bilinmektedir. Örneğin, erkek bir kullanıcı cinsiyetini kadın olarak seçip, yaşadığı yeri Mars olarak belirtebilmektedir. Klasik veri toplama yöntemlerinde ise demografik verilerin yanlış olması olasılığı çok daha düşük olup elde edilen demografik veriler tamamlayıcı niteliğe sahiptir.

Bu araştırma kapsamında görülmüştür ki sosyal medya aracılığı ile elde edilerek iletişim, yeni medya, sosyoloji ve psikoloji alanlarında yapılan Türkçe çalışmaların birçoğu büyük veri setlerini elde etmek, analizini yapmak ve aktarmak konusunda büyük eksiklikler içermektedir. Her ne kadar bu çalışma kapsamında iğne, örneklem olarak seçilmiş araştırmacılara batırılıyor gibi görünse de büyük veri ile araştırmalar yapmayı deneyen bu çalışmanın yazarı aslında çuvaldızı kendine batırmaktadır. Zira bu çalışma kapsamında özetlenen birçok nokta yazarın çalışmalarında da bulunmaktadır. Bu amaçla bu çalışma kapsamında aslında sosyal bilimlerde yapılan araştırmaların kalitesinin artırılmasına ilişkin temel eksikliklerin altı çizilmeye çalışılmıştır. Sosyal bilimlerdeki kadın çalışmaları başlığı altında büyük veriyi dikkate alarak yapılan akademik çalışmaların niceliğinden çok niteliğinin artması, Türkiye’de yaşayan kadınlara ilişkin incelenmek istenilen herhangi bir konu başlığının çok daha anlamlı ve değerli yorumların yapılmasına neden olacaktır. Bunun da ötesinde kişilerin kendi rızalarına dayalı paylaşımlarından elde edilen veriler sayesinde daha geniş kadına ilişkin birçok konuda çok daha geniş bir bakış açısı elde edilmiş olacak ya da sesi çok fazla duyulmayan, belki de her zaman görmediğimiz “ötekilerin” seslerinin duyulmasına neden olunacaktır. Bu da içinde yaşadığımız toplumu daha derinlemesine tanımamızı ve anlamamızı sağlayacaktır.

Kaynakça

- Alikılıç, Ö. & Baş, Ş. (2019). Dijital feminizm: Hashtag’in cinsiyeti. *Fe Dergi*, 11 (1), 89-111.
- Anderson, C. (2008). The end of theory: The data deluge makes the scientific method obsolete. *Wired*, Retrieved from, http://www.wired.com/science/discoveries/magazine/1607/pb_theory.
- Baym, N. K. (2013). Data not seen: The uses and shortcomings of social media metrics. *First Monday* 18(10), Retrieved from, <http://firstmonday.org/ojs/index.php/fm/article/view/4873/3752>
- Berry, D. (2011). The computational turn: Thinking about the digital humanities. *Culture Machine* 12, Retrieved from, <http://www.culturemachine.net/index.php/cm/article/view/440/470>.
- Boyd, D. (2015). Social media: A phenomenon to be analyzed. *Social Media + Society*, 1(1), 1–2. doi.org/10.1177/20563051155580
- Bruns, A. (2013). Faster than the speed of print: Reconciling ‘big data’ social media analysis and academic scholarship. *First Monday* 18(10), Retrieved from, <https://firstmonday.org/ojs/index.php/fm/article/view/4879/3756>
- Cohen, D. (2008). Interchange: The promise of digital history. *Journal of American History*, 95(2), 452–491. doi.org/10.2307/25095630
- Crawford, K. (2013). The hidden biases of big data. *Harvard Business Review Blog*. Retrieved from, <http://blogs.hbr.org/2013/04/the-hidden-biases-in-big-data/>
- Crawford, K., Milner, K. & Gray, M. L. (2014). Critiquing big data: Politics, ethics, epistemology. *International Journal of Communication*, 8, 1663 – 1672. doi.org/1932/8036/20140005

- Çetin, A. ve Edis, K. (2022). Ev kadınlarının sosyal medya deneyimleri: Batman örneği. *Hafıza*, 4 (2) (Özel Sayı), 12-35. doi.org/10.56671/hafizadergisi.1136508
- Dyche, J. (2012). Big data 'Eureka!' Don't just happen. *Harvard Business Review Blog*, Retrieved from, http://blogs.hbr.org/cs/2012/11/eureka_doesnt_just_happen.htm.
- Dursun, O. & Yıldız, F. (2020). Kamu adaletinin büyük veri aracılığı ile inşası: Şule Çet davası üzerine bir twitter analizi. *Connectist: Istanbul University Journal of Communication Sciences*, (58), 291-328. <https://doi.org/10.26650/CONNECTIST2020-0009>
- Edwards, A., Housley, W., Williams, M., Sloan, L. & Williams, M. (2013). Digital social research, social media and the sociological imagination: Surrogacy, augmentation and re-orientation. *International Journal of Social Research Methodology*, 16 (3), 245–260. doi.org/10.1080/13645579.2013.774185
- Ersoy, A. F. (2018). Sosyal medya kullanımının annelerin annelik rolü üzerindeki etkisi. *The Journal of Academic Social Sciences*, 6 (81), 14 -20. doi.org/10.16992/ASOS.14279
- González-Bailón, S., Wang, N., Rivero, A. Borge-Holtoefer, J. & Moreno, Y. (2012). Assessing the Bias in Communication Networks Sampled from Twitter. Retrieved from, <https://arxiv.org/ftp/arxiv/papers/1212/1212.1684.pdf>
- Herring, S. C. (2004). Computer-mediated discourse analysis: An approach to researching online behavior içinde *Designing for virtual communities in the service of learning*. Editörler: S.A. Barab, R. Kling ve J.H. Gray: New York: Cambridge University Press.
- Hey, T., Tansley, S. & Tolle, K. (2009). Jim Grey on eScience: A transformed scientific method içinde *The fourth paradigm: Data-intensive scientific discovery*. Microsoft research. Editörler: T. Hey, S. Tansley ve K. Tolle Washington: Redmond.
- Hogan, B. & Quan-Haase, A. (2010). Persistence and change in social media. *Bulletin of Science, Technology & Society*, 30 (5), 309–315. doi.org/10.1177/027046761038001
- Huzziyas, Ş. & Nizam, F. (2021). Özgecan Aslan cinayeti sonrası sosyal medyada dayanışma olgusu. *Jass Studies-The Journal of Academic Social Science Studies*, 14(86), 389-420. doi.org/10.1016/j.chb.2018.08.039
- Işıklı, Ş. (2014). Büyük veri, epistemoloji ve etik tartışmalar . *AJIT-e: Academic Journal of Information Technology*, 5 (17), 89-122. doi.org/10.5824/1309-1581.2014.4.006.x
- Jungherr, A. (2016). Twitter use in election campaigns: A systematic literature review. *Journal of Information Technology & Politics*, 30(1), 72–91. doi.org/10.1080/19331681.2015.1132401
- Kelling, S., Hochachka, W., Fink, D., Riedewald, M., Caruana, R., Ballard, G. & Hooker, G. (2009). Data-intensive science: A new paradigm for biodiversity studies. *BioScience*, 59 (7), 613–620. doi.org/10.1525/bio.2009.59.7.12
- Kinder-Kurlanda K. & Weller, K. (2014, Haziran). I always feel it must be great to be a hacker! : The role of interdisciplinary work in social media research. Proceedings of the 6th ACM Conference on Web Science. New York: ACM Press.
- Kitchin, R. (2013). Big data and human geography: Opportunities, challenges and risks. *Dialogues in Human Geography*, 79 (1), 1–14. doi.org/10.1177/2043820613513388
- Kitchin, R. (2016). Big data - hype or revolution? İçinde *The SAGE Handbook of social media research methods* T. Hey, S. Tansley ve K. Tolle: London: SAGE Publications.
- Koç, G., Bakır-demir, T. & Öner-özkan, B. (2019). Türkiye'de çalışan annenin sosyal medyadaki temsili. *Karadeniz Sosyal Bilimler Dergisi*, 11 (21), 401-421.
- Köse, S. B. & Akyazı, E. (2020). Kadın cinayetlerinin alternatif medyadaki algısı. *İletişim Çalışmaları Dergisi*, 6 (1), 47-64. doi.org/10.17932/IAU.ICD.2015.006/icd_v06i1003
- Kwon, O., Lee, N., & Shin, B. (2014). Data quality management, data usage experience and acquisition intention of big data analytics. *International Journal of Information Management*, 34 (3), 387–394. doi.org/10.1016/j.ijinfomgt.2014.02.002

- Kuhn, T. (1995). *Bilimsel Devrimlerin Yapısı*. İstanbul:Alan Yayıncılık.
- Langlois, G., Redden, J., & Elmer, G. (2015). Introduction: Compromised data—from social media to big data içinde *Compromised data: From social media to big data*. G. Langlois, J. Redden ve G. Elmer: New York: Bloomsbury Academic.
- Lyu, K., & Kim, H. (2016). Sentiment analysis using word polarity of social media. *Wireless Personal Communications*, 89(3), 941–958. doi.org/10.1007/s11277-016-3346-1
- Manovich, L. (2011). Trending: The Promises and the Challenges of Big Social Data. Retrieved from, http://www.manovich.net/DOCS/Manovich_trending_paper.pdf.
- Mayer - Schönberger, K. & Cukier. (2013). *Büyük veri: Yaşama, çalışma ve düşünme şeklimizi dönüştürecek bir devrim*. İstanbul: Paloma Yayıncılık.
- Mayr, P., & Weller, K. (2016). Think before you collect: setting up a data collection approach for social media studies içinde *The SAGE Handbook of social media research methods*. Sloan, L., & Quan-Haase, A. London: SAGE Publications
- McCay-Peet, L., & Quan-Haase, A. (2016). What is social media and what questions can social media research help us answer içinde *The SAGE Handbook of social media research methods*. Sloan, L., & Quan-Haase, A. London: SAGE Publications.
- Metin, O. ve Ünal, Ş. (2022). İçerik analizi tekniği: iletişim bilimlerinde ve sosyolojide doktora tezlerinde kullanımı. *Anadolu Üniversitesi Sosyal Bilimler Dergisi, AÜSBD Sosyal Bilimlerde Araştırma Yöntemleri Özel sayısı*, 273-294.
- Prensky, M. (2009). H. sapiens digital: From digital immigrants and digital natives to digital wisdom. *Innovate* 5(3). Retrieved from, <http://www.innovateonlineinfo/index.php?view=article&id=705>
- Quan-Haase, A. & Young, A. L. (2010). Uses and gratifications of social media: A comparison of Facebook and instant messaging. *Bulletin of Science, Technology & Society* 30(5), 350–361. doi.org/10.1177/02704676103800
- Schroeder, R. (2014). Big Data and the brave new world of social media research. *Big Data & Society* 1 (2). Retrieved from, <http://bds.sagepub.com/lookup/doi/10.1177/2053951714563194>
- Tavşancıl, E. ve Aslan, E. A. (2001). *Sözel, yazılı ve diğer materyaller için içerik analizi ve uygulama örnekleri*. İstanbul: Epsilon.
- Tekin, F. & Turan, A.T. (2020). Çalışan kadınların sosyal medya kullanım karakteristikleri. *Sakarya Üniversitesi İşletme Enstitüsü Dergisi*, 2 (1), 27-32.
- Weiss, R. & Zgorski, L.J. (2012). Obama administration unveils ‘Big Data’ Initiative: Announces \$200 Million in New R&D Investments. *Office of Science and Technology Policy*. Retrieved from, http://www.whitehouse.gov/sites/default/files/microsites/ostp/big_data_press_release.pdf
- Weninger, T., Zhu, X.A. ve Han, J. (2013). An exploration of discussion threads in social news sites: A case study of the Reddit community. Retrieved from, <http://dl.acm.org/citation.cfm?doid=2492517.2492646>

Makale Bilgi Formu
Yazar(lar)ın Katkıları: Makale tek yazarlıdır.
Çıkar Çatışması Bildirimi: Çıkar çatışması yoktur.
Destek/Destekleyen Kuruluşlar: Bu çalışma herhangi bir kurum tarafından desteklenmemiştir.
Etik Onay ve Katılımcı Rızası: Bu çalışmada etik kurul belgesine gerek yoktur.
Not: Bu makale iThenticate sistemi tarafından taranmıştır.