

Büyük Veri ve Gazetecilik İlişkisi Bağlamında Veri Gazeteciliği

Bilge NARİN, Gazi Üniversitesi, İletişim Fakültesi, narinbilge@gmail.com

Bahar AYZA, Gazi Üniversitesi, İletişim Fakültesi, bharayaz@gmail.com

Feyyaz FIRAT, Gazi Üniversitesi, İletişim Fakültesi, feyyazfirat@gmail.com

Demet FIRAT, Gazi Üniversitesi, İletişim Fakültesi, demetfirat14@gmail.com

ÖZ

Büyük veri iletişim disiplini genelinde gazetecilik özelinde etkili olan toplumsal, kültürel, ekonomik ve teknolojik bir fenomen olarak karşımıza çıkmaktadır. Verideki artış, çağdaş gazeteciliğin haber rutinlerinde de değişime neden olmuş, yeni ve etkili bir anlatı türü olarak veri gazeteciliği yaygınlık kazanmaya başlamıştır. Bu araştırmada alana ilişkin literatürün kapsamlı bir taraması ile durum tespiti yapılmasının yanı sıra; yeniliklerin yayılımı kuramı bağlamında #datajournalism etiketi ile paylaşılan tweetler gönderildikleri ülkelere, konularına ve bu etiket altında paylaşılan veri gazeteciliği türlerinin temalarına göre kategorileştirilerek içerik analizi ile incelenmiştir. Araştırma sonucunda yeniliklerin yayılımında coğrafyanın etkisi kısmen doğrulanmış, Twitter'ın veri gazeteciliği ile ilgili teknolojileri dolaşıma sokup yaygınlaştıran bir araç olduğu bulgulanmıştır.

Anahtar Kelimeler: *Büyük veri, veri gazeteciliği, veri görselleştirme, yeni medya, sosyal medya*

Data Journalism in the Context of Big Data and Journalism

ABSTRACT

The big data is thought to be a social, cultural, economic and technological phenomenon that is effective generally in communication discipline and journalism in particular. The increase in data had also caused a change in the news routines of contemporary journalism, and data journalism had become prevalent as a new and effective narrative type. In this research, in addition to the determination of the situation by a comprehensive literature review, the tweets which were shared with the #datajournalism hashtag are categorized by content analysis according to their sending locations, subjects and subjects of data journalism types with the theory of diffusion of innovations. As a result of the research, the impact of geography on the spread of innovations has been partially confirmed, and it has been found that Twitter is a tool that enables the circulation of technology related to data collection operations.

Keywords: *Big data, data journalism, data visualization, new media, social media*

Giriş

Diğer tüm sektörler ve kurumlar gibi gazetecilik pratikleri de veri bolluğu altındaki internet ortamında anlamlı verileri depolama, değerlendirme ve onları analiz ederek haberleştirme konusunda çeşitli yollar ve yöntemler denemektedir. Medya sektörü için verinin doğru analizi her şeyden önce hedef kitlenin daha iyi anlaşılmasını, içeriğin kişiselleştirilmesini ve veri gazeteciliği gibi yeni anlatı türlerinin dolaşıma girmesini sağlamaktadır. Bu kapsamda araştırmamız büyük veri kavramının gazetecilik pratiklerinde yarattığı temel dönüşümü veri gazeteciliği bağlamında sorgulamayı amaçlamaktadır.

Son on yılda, bilgisayar veri tabanları ve algoritmalar haber kuruluşlarına hızla girmiş durumdadır (Gray, Bounegru ve Chambers 2012; Lewis, 2011). Her geçen gün daha fazla gazeteci, veritabanlarını ve algoritmaları soruşturma yapmak için elverişli birer araç olarak görmekte ve kullanmaktadır.

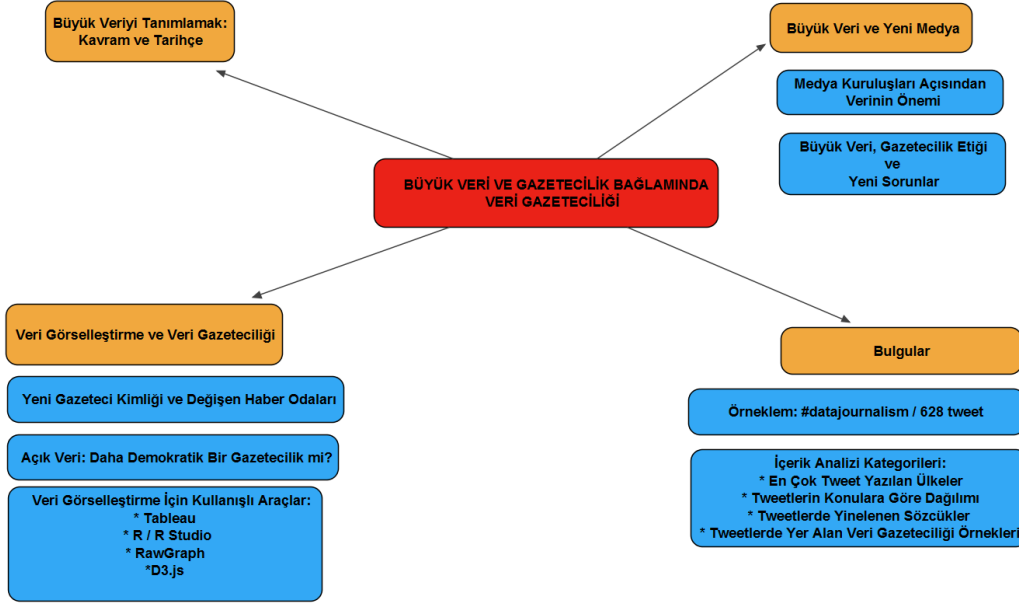
Ana damar medyanın New York Times ve The Guardian gibi haber kuruluşlarının yanı sıra, ProPublica gibi kâr amacı gütmeyen kurumlar ile daha az resmi olarak nitelendirilebileceğimiz araştırmacı gazeteci grupları veri işleme tekniklerine dayanan çalışmalar üretmektedir. Bu girişimler, genellikle "büyük veri" olarak betimlenen daha geniş bir fenomenden izole olarak değerlendirilmektedir. Bu popüler ifade hükümet kayıtları, genetik diziler, internet kullanıcılarının bıraktığı izler gibi türlü içeriklerden oluşan büyük miktardaki bilginin bilimsel araştırma, kamu politikaları veya işletmeler gibi çeşitli platform ve alanlarda işlenmesi anlamına gelmektedir (Parasie, 2014: 1).

Büyük veri tartışmalarıyla birlikte gazetecilik alanında en çok yaygınlık kazanan ifade yollarından biri veri görselleştirme olarak karşımıza çıkmaktadır. Veri görselleştirme hacmi büyük veriyi anlamlı bir hale getirerek açığa çıkarmanın ve karmaşık enformasyonu özetlemenin etkili bir yolu olarak, çağdaş gazeteciliğin önemli anlatı türlerinden biri haline gelmektedir.

Büyük veri, yeni bir gazetecilik türü olarak veri gazeteciliğini ön plana çıkarırken, haber odalarında da değişime neden olmaktadır. Örneğin, günümüzde Guardian gazetesinin iki interaktif ekibi, BBC'nin çevrimiçi haber uzmanı ekibi ve Times'ın görsel gazetecilik birimi bulunmaktadır. Yine Veri Gazeteciliği Ödülleri 2012 yılından bu yana The Global Editors ağı tarafından düzenlenmektedir. Öte yandan, medya sektörünün mevcut ödül kategorilerine dijital dünyada verinin işlenmesine dayalı yeni kategorilerin de eklendiği görülmektedir. Bu kapsamda, Uluslararası Amnesty Medya Ödülleri artık "dijital yenilik" kategorisine sahiptir. Benzer şekilde İngiliz Basın Ödülleri de "dijital ödül" kategorisi oluşturmuştur. Tüm bu gelişmeler veriye ve verinin işlenmesine dair potansiyel yeniliklerin artmasına ve alanın büyümesine hizmet etmektedir (Dick, 2014: 490).

Bu araştırmada büyük veri ve veri gazeteciliği kavramları yeniliklerin yayılımı kuramı bağlamında değerlendirilmiştir. Yeniliklerin yayılımı yeni fikirlerin ve teknolojinin nasıl, niçin ve hangi hızda yayıldığını saptamayı hedefleyen bir yaklaşımdır. Bir teknolojiyi ilk benimseyenlerin kimler olduğu, teknolojinin hangi iletişim kanalları ile ne kadar sürede yayıldığı, yayılım üzerindeki örgütsel ya da yönetsel etkilerin neler olabileceği soruları kuramın anahtar unsurlarını oluşturmaktadır (Roggers, 1983). Çalışmada bu bağlamda, öncelikle büyük veri ve gazetecilik ilişkisi irdelenerek kavrama ilişkin etik tartışmalar

serimlenmiş; verinin analizinde kullanılan yeni anlatı türlerinden biri olan veri gazeteciliğine ilişkin literatür taranmıştır. Veri gazeteciliğinde kullanılan yaygın ve kullanışlı araçların açıklanmasının ardından, çalışmanın ampirik bölümüne geçilmiştir. Araştırmada #datajournalism etiketi ile paylaşılan 628 tweet, içerik analizine tabii tutularak, veri gazeteciliği kavramına temel oluşturan teknolojilerin en çok hangi coğrafyalarda ve ne tür içeriklerle dolaşıma sokularak yayıldığı saptanmaya çalışılmıştır. Veriler, veri gazeteciliğinde kullanılan program ve araçlardan RStudio, Tableau, D3.js ve Raw Graphs aracılığıyla oluşturulan görseller ve tablolar aracılığıyla analiz edilmiştir.



Şekil-1: Araştırmanın Zihin Haritası

Makalenin temel başlıkları ve sunuş planı Şekil-1’de yer alan zihin haritasında gösterilmiştir. Makalede büyük verinin tanımına ilişkin tartışmaların ardından, kavramın yeni medya ile bağı kurulmuş; neden olduğu etik sorunlar sorgulanmış ve gazetecilik pratiklerinde meydana getirdiği önemli dönüşümlerden biri olarak veri görselleştirme ve veri gazeteciliğine ilişkin literatür özetlenmiştir. Veri görselleştirme için kullanışlı bazı araçların açıklanmasının ardından; bu araçlar kullanılarak gerçekleştirilen yeniliklerin yayılımına ilişkin ampirik çalışmanın bulguları sunulmuştur.

Büyük Veriyi Tanımlamak: Kavram ve Tarihçe

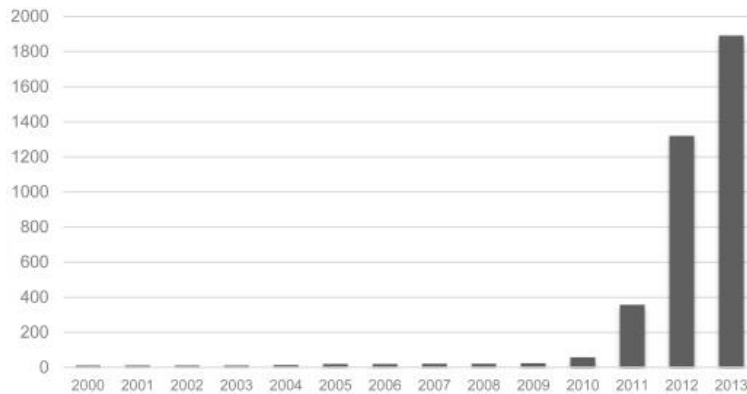
Geçmişte ölçülemeyen, saklanamayan, analiz edilemeyen, paylaşılabilen ve belki de “veri” olarak değeri anlaşılabilen türlü enformasyon parçaları, internet teknolojileri ile birlikte verileştirilmeye başlamıştır. Daha basit bir anlatımla, büyük veri sosyal medya platformlarında kullandığımız emojilerden gün içinde defalarca kez bastığımız beğen tuşlarına; her türlü kullanıcı türevli içerikten insansız hava araçlarının topladığı görüntülere dek uzanan çeşitlilikteki dağınık ve karmaşık enformasyon parçalarından oluşur. İşte bu devasa veri setlerini betimlemede kullanılan büyük veri kavramı, günümüzde modern bilimin ve işletmelerin temel enformasyon toplama ve analiz etme yollarından biri olarak karşımıza çıkmaktadır.

Veri kaydetme ve depolama teknolojisinde benzeri görülmemiş ilerlemeler sonucunda (Diebold, 2012; Schönberger ve Cukier, 2013: 25); ticari kurumlardan tüketicilere, bilimden devlet yönetimine kadar hayatımızın her boyutunda köklü değişiklikler meydana gelmektedir (Jagadish vd., 2014). Bu değişimle birlikte özellikle endüstrinin büyük veriyi benimsemesi, kavramın popüler bir söylem haline gelmesine sebep olmuştur. Akademik yayınlar da bu gelişmeyi takip etmeye başlamıştır (Gandomi ve Haider, 2015; Davenport, 2014: 13).

Büyük veri, geleneksel veri tabanı sistemlerinin işleme kapasitesini aşan hacimdeki veri setlerini tanımlamak için kullanılan bir kavramdır. Bu veri setleri çok büyüktür, çok hızlı hareket etmektedir ve çoğunlukla veri tabanlarındaki altyapı alanında saklanmaya ve incelenmeye uygun değildir. Başka bir anlatımla, bu veriden değer elde etmek için onu işlemeye yarayacak alternatif yol ve yöntemlere ihtiyaç duyulmaktadır (Dumbill, 2013). Bu nedenle, kurum ve kuruluşların büyük veriyi tespit edip depolayarak, düzenleyip analiz edebilmeleri için öncelikle bilişim teknolojisinde donanım, yazılım ve hizmetler için ek yatırım yapmaları gerekmektedir (Narayanan, 2014). Diğer bir anlatımla, büyük veri analizi yeni yöntemlerin geliştirilmesi ve öğrenilebilmesi için eğitime, verinin kayıtlanabilmesi için alt yapıya ve süreci yönetebilecek profesyonel iş gücüne ihtiyaç duymaktadır.

Gürsakağ'a (2014: 19) göre büyük veri kavramı ilk kez Ağustos 2000'de, Seattle'da düzenlenen 8. Dünya Ekonometri Kongresi'nde, Francis X. Diebold tarafından sunulan Makroekonomik Ölçümler ve Kestirim İçin Büyük Veri Dinamik Faktör Modelleri (Big Data Dynamic Factor Models for Macroeconomic Measurement and Forecasting) isimli bildiri kullanılmıştır. Francis X. Diebold (2012) ise, büyük veri kavramının ilk kez 90'lı yılların ortasında Silicon Graphics (SGI)' den John Mashey tarafından ortaya atıldığını belirtmektedir. Kavram, ilk kez 90'lı yıllarda kullanılmış olsa da, aşağıda yer alan Tablo-1'de görüldüğü üzere 2011 yılında yaygınlık kazanmaya başlamıştır.

Tablo-1: ProQuest Araştırma Kütüphanesinde "büyük veri" terimini içeren belgelerin frekans dağılımı (Gandomi ve Haider, 2015: 139).



Büyük verinin akademi, endüstri ve medya arasında paylaşılan ortak bir kaynak olması; her bir paydaşın çeşitli ve çoğunlukla birbiriyle çelişkili tanımlar sunmasına neden olmaktadır. Bu nedenle büyük verinin ne olduğu konusunda genel ve tek bir tanım bulunmamaktadır (Hoy, 2014; Schönberger ve Cukier, 2013: 14; Ward ve Barker, 2013). Bu

noktada alanda yapılmış farklı tanımları ele almak yararlı olacaktır.

Öncelikle kavramı tanımlarken en çok kullanılan üç sözcüğün hacim (volume), çeşitlilik (variety) ve hız (velocity) olduğu gözlenmektedir. Bu üç özellik ilk kez Laney (2001) tarafından dile getirilmiş, ardından çok sayıda yazar tarafından tekrarlanmıştır (Chen, Chiang ve Storey, 2012; Kwon, Lee ve Shin, 2014). Söz konusu özellikler verinin emsallerinden miktarca büyüklüğünü, heterojen doğası nedeniyle çeşitli türlerde olduğunu ve hızı nedeniyle akışkan bir doğası bulunduğunu ortaya koymaktadır.

Kavrama ilişkin literatür incelendiğinde, zaman içinde doğruluk (veracity) ile değişkenlik ve karmaşıklık (variability&complexity) sözcüklerinin de tanımlamada sıklıkla kullanılmaya başlandığı gözlenmektedir. Doğruluk, belirsiz verinin içindeki değerli bilginin veri madenciliği ve veri yönetimi için geliştirilen araç ve yöntemlerle değer kazanması olarak nitelenmektedir. Değişkenlik ise veri akış hızlarındaki değişimi ifade etmektedir (Gandomi ve Haider, 2015: 139).

Bazı tanımlarda geleneksel yöntemlerle toplanıp depolanan verinin hacminin artan doğası vurgulanırken (Kumar ve Pandey, 2013; Monino ve Sedkaoui 2016: XI); diğerlerinde hacimli verinin kurumlar, piyasalar, vatandaşlar ve hükümetler arasındaki ilişki için içgörü sağlayarak karar verme davranışını etkileyebileceği belirtilmektedir (Beyer ve Laney, 2012; Schönberger ve Cukier, 2013: 14). Sonuç olarak büyük veri, çoğunlukla internet ortamında yayıldığından ve büyük ölçüde yapılandırılmamış olduğundan, değeri ilk bakışta anlaşılabilen hacimli enformasyon parçaları olarak tanımlanabilir. Büyük veri analizi ise miktarca çok, dağınık ve karmaşık olan bu verilerin yeni yöntemlerle anlamlı hale getirilerek; mevcut durumun saptanması ve geleceğe yönelik kestirimlerde bulunulması sürecidir.

Büyük Veri ve Yeni Medya

Bir önceki bölümde de ifade edildiği gibi büyük veri, her ne kadar üzerinde uzlaşılmış tek bir tanıma sahip olmasa da genel olarak, gelişen ve yaygınlaşan internet, bilgisayarlar, akıllı telefonlar gibi yeni iletişim teknolojileri üzerinden; Facebook, Twitter, YouTube, websiteleri ve bloglar gibi sosyal medya uygulamaları aracılığıyla üre-tüketiciler (pro-consumer) tarafından üretilen ve aynı zamanda tüketilen metinlerin, görsellerin ve videoların oluşturduğu verilerin toplamı olarak tanımlanabilir (Parks, 2014; Mahrt and Scharrow, 2013; Stone, 2014). Sosyal medyanın bireyler tarafından yaygın kullanımı, muazzam büyüklükteki verinin ortaya çıkmasını sağlayan en önemli itici güç olmuştur (Gürsaka, 2014: 7). Görüldüğü üzere, büyük verinin oluşturulduğu, depolandığı ve dağıtım sokulduğu kanalların yeni medya kanalları oluşu, kavramın iletişim disiplini ile olan organik bağını ortaya koymaktadır.

Diğer yandan, Gürsaka'ya göre (2014: 8) Facebook, Microsoft, Twitter, LinkedIn gibi şirketlerin sahip olduğu ortak özellikler; kalabalığın gücünden yararlanmaları, ortaya çıkan büyük veriyi depolayacak veri merkezleri kurmaları, bilişim ve sosyal ağlarla ilgili teknolojik yenilik üretmeleridir. Bu şirketler, elde ettikleri verileri analiz ederken salt kendi ticari yararları açısından değil, aynı zamanda sosyal medya uygulamalarında da yenilikler yaparak kullanıcıların ya da müşterilerinin beklentilerine karşılık vermektedir. Sosyal ağ platformları kullanıcı bilgilerini verileştirmekte, bu verilerden yola çıkarak uygulamalarını bireyler açısından daha verimli (örneğin, LinkedIn'deki meslektaş önerileri) hale

getirmektedir (Schönberger ve Cukier, 2013: 99).

Yeni medya-büyük veri 'evliliğinin' diğer bir önemli yönü ise kriz zamanlarındaki işlevselliğidir. Belirli bir bölgede yaşanan doğal afet durumlarında Twitter gibi bir sosyal medya uygulaması hayat kurtarıcı olabilmekte; dahası afet bölgesinden atılan Tweetler bölgedeki genel durumun anlaşılmasında ve kurumların acil durum yönetiminde yararlı olabilmektedir. Örneğin büyük Japonya Depremi'nde 'j_j_helpme' etiketi kullanıma sunulmuş ve acil durum bölgesindeki bireyler tweetlerini bu etiketi kullanarak atmıştır. Böylece, atılan tweetlerin tek bir 'havuzda' toplanması sağlanmış, ortaya çıkan 'büyük veri' çeşitli araçlarla analiz edilerek; genel durumun anlaşılmasına dönük bilgi edinilmesine olanak sağlanmıştır (Wang, Wu, Yen, Guo ve Cheng, 2016). Sonuç olarak, yeni medya büyük veriyi sürekli bir biçimde besleyen önemli kanallara sahiptir. Büyük verinin analizi ise daha etkili ve doğru iletişim süreçlerinin geliştirilmesinde önemli öngörüler sağlamaktadır. Bu bağlamda iki kavram birbiriyle etkileşim halindedir.

Medya Kuruluşları Açısından Büyük Verinin Önemi

Büyük veri, birbirinden farklı pek çok alanda faaliyet gösteren kuruluşlar açısından veri madenciliği ile doğru bir şekilde analiz edildiğinde son derece önemli olabilmektedir. Sağlık hizmetlerinden, endüstriyel ürün tasarımlarına, bankacılık hizmetlerinden pazarlama faaliyetlerine kadar pek çok alanda verinin gücünden hem kuruluşların kendisi hem de müşterileri faydalanabilmektedir. Medya kuruluşları da büyük verinin gücünü, veriyi toplamanın ve analiz etmenin önemini biraz yavaş ve geç olsa da anlamaya başlamaktadır.

Medya ve eğlence şirketleri, geçmişte genellikle kendi sezgileri doğrultusunda karar aldıkları için ve insanların paylaşmış oldukları içeriklerin bir tüketicisi olup olmadığını anlamaya yarayacak bir ölçümü nasıl yapacaklarını bilmediklerinden dolayı çok fazla başarılı olamamışlardır (Davenport, 2014: 52). Geleneksel medyada, örneğin televizyonlarda sunulan dizilerin, filmlerin izlenme oranlarını ölçümlemek çoğu durumda zordur ve hatalı sonuçlar verebilmektedir. Basılı gazetelerin tirajları ise çoğunlukla yalnızca satış rakamlarına bakılarak değerlendirilmektedir.

Bugün itibarıyla gelinen noktada ise dizilerin, filmlerin ve haberlerin internet üzerinden sunulması, onları sunan medya kuruluşlarına bir dizi avantaj sağlamaktadır. Hangi gazetenin ne kadar okunduğu –kaç tık aldığı-, okuyucuların sayfa üzerinde kalış süreleri, haberlerle ve gazeteyle ilgili olarak yapılan kullanıcı yorumları, hangi dizinin ya da filmin ne kadar izlendiği, dizideki karakterlerle ilgili olarak izleyicilerin düşünceleri ile rahatsız edici sahneler ve karakterlerle ilgili kullanıcı yorumları görülebilmektedir. Bu çok sayıdaki parametre sayesinde, izleyici/kullanıcı araştırmalarında geleneksel kitle iletişim araçlarının ölçülmesine kıyasla doğruya en yakın sonuçlar elde edilmektedir. Elde edilen tüm bu veriler, medya kuruluşlarına sundukları ürünlerin zayıf noktalarıyla ilgili fikir sunmakta; bu geri dönüşlerin veri madenciliği ile analizi sayesinde sunmuş oldukları ürünlerde bazı düzeltmeler yapmalarına olanak sağlanmaktadır. Dolayısıyla, büyük veri çağındaki medya kuruluşlarının elinde önemli bir fırsat bulunmakta; bu fırsatın somut bir başarıya dönüşmesi medya kuruluşlarının veri depolayabilme gücüne, verileri analiz etmelerini sağlayacak teknolojik araçlara sahip olmalarına ve belki de en önemlisi bu verileri doğru bir şekilde analiz edebilecek, veri analizi konusunda yetenekli insan kaynağına bağlı olmaktadır.

Büyük Veri, Gazetecilik Etiği ve Yeni Sorunlar

Büyük veri ve veri gazeteciliği temel olarak; kullanıcı gizliliğinin ihlal edilmesi, bilgi güvenliğinin sağlanamaması, büyük verinin gazetecinin failliğini ortadan kaldırması ve büyük veriye erişim olanaklarının herkes için eşit olmaması gibi çeşitli konularda sorunlarla karşılaşmaktadır. Kullanıcı gizliliği konusu, büyük veriye ilişkin olarak aşılması en zor etik sorunlardan biridir. danah boyd ve Kate Crawford (2015: 209), büyük verinin etik sorunları beraberinde getirdiğini söylerler; çünkü verilerin kamuya açık olması araştırmacıların kullanıcı türevli içeriği izin almaksızın kullanması gibi bazı sonuçlar doğurur. Eğer sorulacak olsa, kullanıcıların çoğu ürettikleri içeriğin başka yerlerde kullanılmasına izin vermeyecektir. Ayrıca, kullanıcıların çok büyük bir çoğunluğu ürettikleri içeriği depolayabilen algoritmalarından da habersizdir. Bu nedenle, araştırmacıların etik davranabilmeleri için, araştırmanın kaynaklarına hesap verebilirliklerini kavramış olmalıdırlar. Akademisyenler, insana dayalı çalışmalar yürütürlerken etik kurullardan onay almaktadır. Ancak pek çok etik kurul, büyük veri madenciliği ve anonimleştirme süreçlerinin, verilerin kişisel olarak tanınabilir hale gelmesi sorununa neden olabileceğini kavrayamamaktadır.

İnsanı konu alan araştırmalar ile gazetecilik mesleği arasında büyük veri kullanımına dair farklılıklar olsa da, insanı konu alan araştırmalarda uyulması gereken kurallar gazetecilik mesleği için de bir temel oluşturabilir. Çünkü gazeteciler de kitle iletişimi söz konusu olduğu zaman bilgilendirilmiş rıza edinme konusunda zorluklarla karşılaşmaktadır. Genellikle veriye katkı sağlayanların tamamıyla iletişime geçmek mümkün değildir; çünkü çok fazla veri vardır (Fairfield ve Shtein, 2014: 42). Büyük veriyi kullanan gazetecilerin ya da araştırmacıların, bu veriye dahil olan kişilere belli bir ücret ödemesi gerekirse, çok büyük bir verinin yalnızca çok küçük bir parçası olan bu kullanıcılara ödenecek ücret çok düşük olacaktır; ama örneğin bir kişinin tüm arama geçmişinin dünyaya açılmasının kişide yaratacağı hissin karşılığı çarpıcı olabilir. Ayrıca, gazeteciler çevrimiçi bir topluluk hakkında bir çalışma yürütüyorlarsa, onların yalnızca gözlem yapmaları bile topluluk ya da güvenlik hissini bozabilir (Fairfield ve Shtein, 2014: 43). Gazetecinin, sanal topluluğun tüm üyelerinden izin alarak araştırmasını yürütmesi gizliliğe dair etik sorunu ortadan kaldıracaktır; ancak gazetecinin tek tek bireylere ulaşması ve izin alması mümkün değildir. Bu nedenle gazetecilerin büyük veriyi kullanımını oldukça sorundur. Bu soruna neden olan faktörlerden birisi de, kullanım sözleşmeleridir.

Kullanım sözleşmeleri kullanıcıların tüm kişisel enformasyonunu depolamayı meşrulaştırmaya uygun zemin hazırlamaktadır. Sosyal medya uygulamalarına ya da çeşitli sanal cemaatlere üye olmak isteyen kullanıcılar, üyeliklerini tamamlayabilmek için son aşamada "kabul ediyorum" tuşuna basmak zorundadırlar. Genellikle kabul edilen kullanım sözleşmesi de gizlilik ihlali de dahil olmak üzere çeşitli sorunları beraberinde getirir. Kullanıcılar da bu sözleşmeleri aslında istedikleri için değil, bu platformlara dahil olmanın bir önkoşulu olduğu için mecburen kabul ederler. Kişisel verilerinin kontrolünü bir başkasına devretmek, bu platformlara dahil olmak için ödenmesi gereken bir bedeldir (Crawford vd., 2014: 1669). Kısacası, kullanım sözleşmeleri kişisel verilerin kontrolünün başkalarına devredilmesini yasal olarak mümkün hale getirmektedir.

Kullanıcıların yalnızca adı veya sosyal güvenlik numaraları gibi bazı kişisel bilgilerinden yola çıkılarak ve çeşitli kişisel enformasyonları birleştirilerek onlar hakkında bir dosya bile oluşturulabilir. Üstelik, kişisel verileri elinde tutanlar, kişilere dair çok daha özel

bilgilere ulaşabilirler. Örneğin, şirketler eskiden yaz aylarında hangi renkli otomobilin Teksas'ta en çok satın alındığını bilmekten heyecan duyuyorlarken, günümüzde ise geçen yıl ailenizin ne kadar dış macunu satın aldığını bilebilirler. Hatta kocanızdan önce hamile olduğunuzu öğrenebilirler (Mateosian, 2013: 60). Büyük veri üzerine çalışan kişiler, bir anlamda gözetim mekanizmasının kontrolünü de ellerinde tutuyor hale gelirler. Böylece hem kullanıcıların gizliliği ihlal edilmiş olur, hem de büyük verinin çok küçük bir parçası olan bu kişiler, kullanım sözleşmelerini mecbur oldukları için onaylayarak, aslında bir gazetecinin çok da dahil olmak istemeyecekleri araştırmasına katılmaya mecbur kalabilirler. Kısacası, kullanıcılar kullanım sözleşmelerini onaylayarak kişisel verilerini platform sahiplerine devrederler ve böylece mevcut güç ilişkilerinin sürmesine neden olurlar.

Ayrıca, büyük veri, deterministik bilginin insan davranışına uygulanmasından kaynaklanan insan özerkliğinin kaybolmasıyla ilgili sorunlar da ortaya çıkarmaktadır. Bu sorunlar, bilgi karşısında özgür irade ve insan failliği hakkındadır. Kısacası büyük veri, araştırmacıları ya da gazetecileri bireylerden uzaklaştırıyor gibi görünmektedir. Büyük veri, bilgiyi yeni alanlara (çevrimiçi etkinliklerden gelen davranışları tahmin etme gibi) genişletir; bireysel davranışları saptamada daha fazla doğruluğa erişir. Bilgiyi üretme kabiliyeti, yeni aktörler ve daha güçlü araçlar yaratır (Schroeder ve Cows, 2014). Tüm bu kabiliyetler de gazetecinin etkinliğini azaltmaktadır, bu durum da yeni sorunlara neden olmaktadır.

Bilgisayarlar ve veri analizi kapasiteleri oldukça gelişmiştir; ancak hala büyük verinin kendi başına bir gazetecilik değeri yoktur, bu nedenle de gazeteciler verinin içinde gizlenmiş hikâyeyi bulup çıkarmalıdır (Parasie ve Dagiral 2012, 859). Bu anlamda, büyük veriyi oluşturan, depolayan ve işleyen algoritmaların gazetecilerin yerine geçtiğini söylemek zordur. Yine de makineler, haber seçimi ve haber değerlendirmeleri gibi çeşitli konularda aktif olarak kullanılmaktadır. Algoritmalar henüz gazetecilere olan gereksinimi ortadan kaldırmasa da oldukça önemli roller üstlenmektedirler; ancak bu durum makinelerin etik olarak da hareket edebileceğini gösterir mi? Kraemer buradan yola çıkarak makinelerin etik davranıp davranamayacaklarını sorgular (Kraemer vd., 2011: 252). Matt Carlson da benzer şekilde "otomatize gazetecilik" kavramı ile birlikte gazetecinin etkinliğinin azaldığından bahseder. Bu durumun da gazetecilikte emek ve gazetecilik otoritesi gibi konularda sorunlara neden olduğundan söz eder ve algoritmalara dayalı değerlendirmelerin insan değerlendirmesinin otoritesinde bir düşüşe neden olabileceğini söyler (Lewis, 2015: 326). Gazetecinin failliğinin azalması da çeşitli etik sorunlara yol açabilir. Örneğin, bir gazetecinin kamu yararını gözeterek bir olayı haberleştirmemeye karar vereceği durumda bir algoritma nasıl davranacaktır?

Büyük veriye erişim olanaklarının tüm gazeteciler için eşit olmaması da büyük verinin temel sorunlarından birisidir. Herkesin büyük verilere ulaşma imkânı yoktur; çünkü büyük veri ile çalışabilmek bu veriyi işleyemeye uygun bilgisayarlar, depolama için geniş depolama kapasiteleri; kimi zaman da gerekli büyük veriye ulaşabilmek için ilgili endüstri ile bağlantıda olmak gerekir (Fairfield ve Shtein, 2014: 46). Bu nedenle, büyük veriye erişim imkânı da üzerine düşünülmesi gereken bir konudur. Sermaye sahiplerinin ve büyük şirketlerin, büyük veriye ulaşma, büyük veriyi depolama ve işleme imkânının daha fazla olması, mevcut medya yapılanmasının ve eşitsizliklerin sürdürülmesine neden olacaktır ve alternatif medya yapılanmaları büyük veriyi kullanma konusunda sıkıntılar yaşayacaktır.

Ayrıca, herkesin büyük veriye ulaşma imkânın eşit olmadığı ve algoritmaları oluşturanların güç sahipleri olduğu düşünüldüğünde, gazetecilerin büyük veriyi yalnızca güç sahiplerinin çıkarlarına hizmet edecek şekilde kullanabileceği tehlikesi de ortaya çıkmaktadır. Bunun da ötesinde, gazetecilerin karar verme görevini üstlenen algoritmaların toplumun çoğunluğunun yanı sıra azınlıkları da gerçekten temsil edip etmeyeceği sorusu da ortaya çıkmaktadır. Medya sahipleri, gazeteciler ve iletişim bilimciler, büyük veri, gazetecilik etiği ve gazeteciliğin geleceği hakkında düşünerek; mevcut etik sorunların nasıl üstesinden gelinebileceği hakkında değerlendirmelerde bulunmalıdırlar. Aksi halde büyük veri ile birlikte ortaya çıkan etik sorunlar geçmiş sorunlara eklenerek var olmaya devam edecektir.

Veri Görselleştirme ve Veri Gazeteciliği

Yeni iletişim teknolojileri dolayısıyla ortaya çıkan “büyük veri”, yalnızca içgörü ve tahmin sağlamamakta, aynı zamanda gazetecilik mesleği ve uygulamaları üzerinde de birtakım değişim ve dönüşümlerin yaşanmasına sebep olmaktadır. Teknolojik altyapının verinin daha hızlı toplanıp, kontrol edilmesi ve doğrulanması gibi avantajlar sağlaması; onun bir haber kaynağı olarak kullanılmasını da giderek daha önemli hale getirmektedir (Dağ, 2015: 1). Diğer bir anlatımla, dünya daha dijital hale geldikçe; daha fazla bilgi veri biçimlerinde depolanmakta, gazetecilerin verilere erişimi gittikçe kolaylaşmaktadır (Knight, 2015: 1). Bu verilerin belirli amaçlar doğrultusunda özetlenmesi amacıyla gerçekleştirilen veri görselleştirme, en basit haliyle ham verinin kullanıcılar için anlaşılır hale getirilecek şekilde görselleştirilmesi olarak tanımlanabilecektir. Veri görselleştirmeyi bir anlatı aracı olarak kullanan veri gazeteciliği ise dijital dünyanın araştırmacı gazetecilik türleri arasında son dönemde popülerleşen bir tür olarak karşımıza çıkmaktadır.

Veri gazeteciliği ve veri görselleştirme birbirleriyle ilişkili iki kavram olsa da aynı anlama gelmemektedir. Çünkü veri gazeteciliği yalnızca veri görselleştirmeden oluşmaz; ona indirgenemez. Veri analizine dayanan bir haber bazen en iyi imgeler ve infografi ile anlatılabilirken; bazense sözcüklerle anlatılabilmektedir. Bu iki türün ortak özelliği ise sayı ve istatistik içermeleridir (Rogers, 2014: 31-33-34). Özetleyecek olursak, veri görselleştirme veri gazeteciliğinin temel anlatı unsurlarından biri olmakla birlikte, veri gazeteciliğinde esnek anlatı biçimleri kullanılabilir.

Veri gazeteciliği kazalar ve basın bültenleri gibi rutin haber akışının hâkim olduğu haber odasındaki günlük çalışmaların aksine, hiç kimsenin hikayeleştirmediği bir alanda haber bulmakla ilgilidir (Appelgren ve Nygren, 2014: 401). Bu anlamıyla veri gazeteciliği, geleneksel gazeteciliğin yerine geçen değil aksine onu tamamlayan ve belki de derinleştiren bir gazetecilik türü olarak değerlendirilebilecektir. Çünkü veri gazeteciliği haber metni, fotoğraf ve video sunumundan oluşan bir haberciliğin çok daha ötesinde bir habercilik biçimini ifade etmektedir.

Zotto, Schenker ve Lugmayr’a göre (2015), verinin haber içeriklerini geliştirme ve habere konu olan bilgiyi zengileştirme anlamında kullanılması aslında 1800’lü yıllara dayanmaktadır. Bu kapsamda, gazetecilerin zaten senelerdir bazı haberleri oluştururken veri, görselleştirme ve istatistik kullandıklarını düşündüğümüzde; onların çalışma pratikleri açısından neyin değiştiğini ve neyin “yeni” adlandırıldığını sorgulamak yerinde olacaktır.

Bazı haber editörlerine göre “veri gazeteciliği” sözü anlamsızdır. Çünkü gazeteciler, gazetecilik tarihi boyunca dijital araçları kullanmasalar da elle veri analizi yapmışlardır. Bugün değişen şey ise, daha büyük veri kümelerinin dijital olarak ve daha hızlı analiz edilebilmesidir. Yine geçmişte, izleyicilerin basılı bir haber ürününde yer alan verilere kendi hikayelerini katma, haberleri filtreleme ve seçme olasılıklarının sınırlı olması sebebiyle veri gazeteciliğinin Web'e bağlı bir doğası bulunmaktadır (Appelgren ve Nygren, 2014: 402). Yani iki asır öncesinde veriye erişim, veri kaynağının doğrulanması ve kullanıcı türevli içeriğin aktarımı bugüne kıyasla çok daha zor olduğu için veri gazeteciliği tanımı, içinde bulunduğumuz teknolojik çağın gazetecilik anlayışıyla daha uygun düşmektedir.

Veri gazeteciliği en basit haliyle, geleneksel gazetecilik çalışma yöntemlerinin veri analizi, programlama ve görselleştirme teknikleriyle harmanlandığı haber anlatım biçimi olarak tanımlanmaktadır (Nygren, Appelgren ve Hüttenrauch 2012). Haberin sayılarla anlatılması, sayılardan haber bulunması ve veriye kaynak olarak yaklaşılmasıdır (Dağ, 2015: 2). Daha genel bir tanımlama yapmak gerekirse, veri gazeteciliği, dağınık halde bulunan büyük orandaki veri içerisinden birbiriyle anlamlı ilişkiler sunan verilerin görülebilmesi; bu verilerin kümelenebilmesi ile istatistikler, infografikler ve etkileşimli veri görselleştirmeleri yoluyla bir haber öyküsü oluşturulabilmesidir (Howard 2014; Smit, Yael ve Laura, 2014).

“Veri gazeteciliği” teriminin ilk kez 2008 yılında Guardian gazetesinde kullanıldığı görülmektedir. Simon Rogers, yayınladığı blogda kavramı şöyle tanımlamıştır:

“Dün itibarıyla, geliştirme ekibimiz ham verileri alıp düzenlenebilir bir haritaya dönüştüren bir uygulama geliştirdi. Bu da, bu rakamlara dayalı fantastik bir etkileşimli grafik üretebileceğimiz anlamına geliyor. Veri gazeteciliği ile editör ve geliştiriciler, teknik açıdan ilginç bir şeyler geliştirdiler. Bu, veriyle nasıl çalıştığımız ve onu nasıl gördüğümüzü değiştiriyor” (Rogers, 2008).

Veri gazeteciliğinin yaygın olarak gündeme gelmesi ise özellikle 2010 yılında Irak ve Afganistan’la ilgili 92.000 sayfalık gizli hükümet yazışmalarının Wikileaks aracılığıyla sızması ile olmuştur (Zotto vd., 2015). Bu olay, gazetecilerin ortaya çıkan büyük orandaki veriyi haber öyküsüne dönüştürme konusunda yaşadığı zorlukları ve mesleki yetersizlikleri göstererek; veri gazeteciliğinin ciddi anlamda ele alınması gerektiğini ortaya koymuştur.

İlgili literatürden hareketle, veri gazeteciliğinin temel yapısı ve özellikleri şöyle özetlenebilecektir:

- Rutin haber akışından farklı olarak; daha çok zaman isteyen, daha hacimli ve çoğunlukla web temelli verinin analizine dayanan bir araştırmacı gazetecilik türüdür.
- Görselleştirme veri gazeteciliğinin yaygın yöntemlerinden biri olsa da esnek formatlar bir arada kullanılabilir.
- Kullanılan kaynaklar çoğunlukla halka açık veriler, kullanıcı türevli içerik ya da kitle kaynaktan oluşmaktadır.
- Veri gazeteciliği yeni araçlara ve yöntemlere ihtiyaç duymaktadır. Bu durum gazetecilerin bu yeni teknikler konusunda eğitilmelerini ve yeni yetiler kazanmalarını zorunlu kılmaktadır.

Açık Veri: Daha Demokratik Bir Gazetecilik mi?

Veri gazeteciliğinde dikkat çeken unsurlardan biri de kaynağın ne olduğudur. Rogers, verilere ve sayılara artık yalnızca uzmanların değil sıradan tüm kullanıcıların sahip olduğunu söyleyerek, bu durumu “verinin demokratikleştirici gücü” olarak nitelemektedir (2014: 31-33). Veri gazeteciliği haberleri de çoğunlukla kamuya açık verilerden veya genel halkın yardımıyla toplanan kitle kaynaklı büyük veri setlerinden oluşturulmaktadır (Appelgren ve Nygren, 2014: 394). Bu kapsamda açık veri kavramının ne anlama geldiğinin açıklanması yerinde olacaktır.

OpenDefinition.org’a göre açık veri, herkes tarafından özgürce ve ücretsiz bir biçimde erişilmesi, kullanılması ve yayılması mümkün olan veri türü olarak tanımlanmaktadır (Dağ, 2015: 29). The Open Knowledge Foundation (2012) ise açık verinin, mevcudiyet ve erişim, yeniden kullanım ve dolaşım ile evrensel katılım unsurlarından oluştuğunu belirtmektedir. Özellikle, ülkeler sahip oldukları pek çok veriyi şeffaflık, hesap verilebilirlik ve yolsuzlukla mücadele adına açık hale getirmektedir. Türkiye de Açık Devlet İş birliği (Open Government Partnership) oluşumuna 2012 yılında imza atmıştır (<http://www.opengovpartnership.org/country/turkey>). Dolayısıyla, veri gazeteciliği ile hükümetlerin paylaştığı veriler istatistikî, görsel unsurlarla birlikte haber öyküsü oluşturularak okuyuculara sunulmaktadır. Eğitim, sağlık ve ekonomi gibi pek çok konuda verilerin kolayca ulaşılabilir olması veri gazeteciliğinin daha demokratik bir toplum oluşturmada bir köprü görevi üstlenmesini sağlamaktadır.

Literatürde, veri gazeteciliğinin demokratikleştirici gücünü fazlaca iddialı bulan araştırmalar da yer almaktadır. Örneğin, İngiltere’de yayınlanan “nitelikli” üç gazete grubunda yer alan veri gazeteciliği türlerinin analiz edildiği bir çalışmada, veri kaynakları incelenmiştir. Veri gazeteciliği ile oluşturulmuş haberlerde devlet kurumları, tüzel kişiler, araştırma enstitüleri (kâr amacı gütmeyen gruplar ve akademik kurumlar), pan-ulusal örgütler (Birleşmiş Milletlerin çeşitli ajansları gibi), anketler ve kendiliğinden üretilen haberler (örneğin haber kuruluşunun kendisi tarafından toplanan veriler) de dahil olmak üzere birçok jenerik kaynağın kullanıldığı gözlenmiştir. Ancak çalışmanın bulguları, örneklem olarak seçilen gazetelerde sunulan veri gazeteciliği örneklerinin kurumsal kaynaklara ve çoğunlukla da devlet kurumlarına dayandığını ortaya koymuştur (Knight, 2015). Bir başka anlatımla, bazı veri gazeteciliği örnekleri resmi kaynaklara bağımlı doğaları nedeniyle araştırmacı gazeteciliğinin yeni, demokratikleştirici ve dijital bir türü olmaktan çok; derinlikten uzak, resmi kaynak bağımlı ve yüzeysel birer anlatı olarak karşımıza çıkmaktadır.

Yeni Gazeteci Kimliği ve Değişen Haber Odaları

Veri gazeteciliğinin daha demokratik bir habercilik anlayışı olup olmadığı ile ilgili bir önceki başlıkta yer alan tartışmalar sürerken; bu yeni tür haber odalarında giderek daha fazla yaygınlık kazanmaktadır. Bu durum hem haber odalarındaki iş gücü ve iş bölümü üzerinde hem de gazetecinin sahip olması gereken yetiler hakkında yeniden düşünmemizi gerekli kılmaktadır.

Günümüzde profesyonel haber odalarında gazetecilik ve görsel tasarımın iç içe geçtiği yeni bir haber üretim süreci söz konusudur. Enformasyon görselleştirme, veri görselleştirme ve infografik gibi adlarla anılan (Smith vd., 2014: 345) bölümler, çağdaş

gazeteciliğin haber merkezlerine hızla eklenmektedir. Gündelik haber pratiklerinde, hem okur hem de üretici kitlesini daha çok kendi kendine öğrenebilen gençlerin oluşturduğu yeni bir dalgaın başladığı gözlemlenmektedir (Rogers, 2014: 33-34).

Veri gazeteciliği haber odalarına gazetecilerin yanı sıra istatistik, enformasyon tasarımı ve bilgisayar mühendisliği formasyonuna sahip profesyonellerin de dahil olmasına neden olmaktadır (Dick, 2014: 491). Bir başka ifadeyle, veri gazeteciliği ile ilgili teknolojik alt yapı ve becerilerin haber odalarına transferi yalnızca gazeteciler tarafından gerçekleştirilmemektedir. Profesyonel gazetecilik istatistik, enformasyon tasarımı ve bilgisayar mühendisliği alanları ile artan oranda yakınsamaktadır.

Farklı uzmanlık kollarının haber odalarına girişlerinin temel nedenlerinden biri haber kuruluşlarının pek çoğunun, veriyle çalışma konusunda uzmanlaşmış gazeteciden yoksun olmasıdır (Zotto vd., 2015). Bu açığın nasıl kapatılacağı ile ilgili iki farklı anlayış bulunmaktadır.

İlk görüş gazetecilerin mesleki yeterliliklerini gözden geçirerek veri gazeteciliğinin gerektirdiği becerileri kazanarak geliştirmeleri gerektiğini savunmaktadır. Bu yaklaşıma göre, gazetecilerin sayısal verilerle ayrıntılandırılmış, veri görselleştirme ile özetlenip anlaşılır kılınmış haber içeriklerini üretebilmesi için, şu zamana kadar yazılımcılar ve tasarımcıların iş yükü gibi gözükken konularla ilgili teknik donanımlarını geliştirmesi zorunluluktur.

Her ne kadar gazeteciler ve bilişim teknolojisi uzmanları arasındaki mesleki beceri açığı kolay kapanacak gibi gözükmeseyse ve gazetecileri teknik özelliklerle donatmak kısa vadede mümkün olmasa da (Howard, 2014; Lewis ve Usher, 2014); gazeteciler temel düzeyde de olsa yazılım, algoritma ve kod bilgisine sahip olursa veri gazeteciliği açısından önemli bir avantaja sahip olacaklardır (Stavelin, 2014). Kirk (2012)'e göre gazeteciler artık yalnızca haber toplayıp yazmamalı aynı zamanda şu özelliklere de sahip olmalıdır: öncü, veri bilimci, gazeteci, bilgisayar bilimci, tasarımcı, bilişsel bilimci, iletişimci ve proje yöneticisi...

Üstelik mevcut teknolojik gelişmelerle bu becerileri kazanmak artık çok maliyetli de değildir. Taşımacılık sübvansiyonları ile ilgili çalışmaları ile veri gazeteciliği ödülü alan Arjantinli veri takımı La Nacion'dan Florencia Coelho, yazılımcı ve tasarımcı görevlendirecek bütçeleri olmadığından gazetecileri nasıl eğittiklerini şöyle özetlemiştir: "Haber odamızda bir web programcımız ya da bilgisayar destekli raporlama sorumlusu yoktu. Biz interaktif bir tasarımcı bulduk ve Buenos Aires'de bir alışveriş merkezindeki kahvecide kendimizi Tableau'nun ücretsiz eğitim videolarını izleyerek eğittik" (Rogers, 2014: 33). Avrupa Gazetecilik Merkezi (European Journalism Centre) tarafından, 200 kişiyle yapılan bir anketin sonuçlarına göre de gazetecilerin %70'i veri gazeteciliğinin çok önemli olduğuna inanmakta ancak aynı zamanda kendilerinin konuyla ilgili yetersiz bilgiye sahip olduklarını hissetmektedirler (Bradshaw, 2011). Gazetecilerin basit yöntemler ve araçları kullanmak için hızlı ve pratik kurslara ihtiyaç duydukları ve zaman baskısı hissettikleri başka araştırmalar (Appelgren ve Gunnar, 2014) tarafından saptanan bir bulgu olarak da karşımıza çıkmaktadır.

Diğer yaklaşım ise, gazetecilerin veri görselleştirme programları öğrenmeleri yerine,

farklı meslek kolları ile koordineli bir iş birliğini önermektedir. Çünkü bu yaklaşıma göre, haber odalarında bütün çalışanlar tüm yetilere sahip olmayabilir. Ancak gazeteci, tasarımcı ve programcı gibi farklı uzmanlar arasındaki iş birliği çok önemlidir (Beak, Liebowitz ve Lewis, 2000, Mora, 2012). Örneğin The New York Times gazetesi, gazeteciler ile yazılımcıları aynı departmanda buluşturarak geçici de olsa bir çözüm üretmiştir (Doğu, 2015: 193).

Gazetecilerin veri gazeteciliği becerilerini kendilerinin mi kazanacağı yoksa profesyonel iş birliği ve yardım mekanizmaları mı geliştirecekleri konusundaki tercihte, haber organizasyonunun yapısı da etkili bir faktör olarak karşımıza çıkmaktadır. Çalışma hızı, üretim süreci, kuruluşun büyüklüğü, enformasyon görselleştirme için verilen yönetsel destek de görsel öğelerin nasıl ve kimler tarafından hazırlanarak dolaşıma sokulacağı konusunda etkili olmaktadır. Örneğin büyük kurumlarda, bağımsız ve büyük bir grafik bölümü bulunurken, küçük kurumlarda ise çoklu yeteneğe sahip gazeteciler enformasyon görselleştirme işini üstlenmiş durumdadır (Smith vd., 2014: 350).

Öte yandan, haber kuruluşlarındaki günlük işleyiş; zamanın hayati önemde olduğu, hızlı hareket edilmesi gereken bir işleyiş mantığına dayanmaktadır. Görsel üretim ise, daha fazla zaman isteyen, daha derin art alan araştırması gerektiren bir süreçtir. Bu farklılık günlük üretim sürecinde çatışmalara neden olabilmektedir (Weber ve Rall, 2012). Veri görselleştirme ile ilgili çalışma pratiği yürüten diğer sektörlerle kıyaslandığında; gazeteciliğin hız baskısı nedeniyle farklı bir görünüm arz ettiği gözlenmektedir. Örneğin beyin fırtınası ve görselleştirmenin amacına yönelik tartışma gibi önemli basamaklar, çoğunlukla hız baskısı nedeniyle atlanmaktadır (Smith vd., 2014: 352).

Özetleyecek olursak, veri gazeteciliğinin yaygınlık kazanıp olgunlaşmasının önündeki en büyük engel internet gazeteciliğinin hız baskısıdır. Eğitimli ve donanımlı çalışan sayısındaki eksiklik de diğer bir engel olarak mevcudiyetini korumaktadır. Bu bilgi eksikliğinin nasıl ve hangi araçlarla giderilebileceği dikkate değer bir konudur.

Veri Görselleştirme İçin Kullanılı Araçlar

Veri gazeteciliği pratikleri ülkeden ülkeye farklılık gösterse de çalışma yöntemleri ve araçları büyük oranda benzerlik göstermektedir. (Appelgren ve Nygren, 2014: 394). Etkili veri görselleştirme için gazetecilerin yeteneklerinin neler olması gerektiği konusunda literatür tarandığında veri madenciliği, programlama, yazma ve tasarlama becerileri genel öneriler olarak karşımıza çıkmaktadır (Giardiana ve Medina, 2013; Kirk, 2012; Segel ve Heer, 2010; Weber ve Rall, 2012). Bu bölümde gazetecilerin hangi araçlarla ve yöntemlerle bu becerilerini geliştirilebileceği konusu üzerinde durulacaktır.

Tableau: Veri görselleştirmeye yarayan bir programdır. Hızlı, pratik bir araç olup yazılım bilgisi gereksiz eldeki verilerden basitçe veri görselleştirme yapabilmeye olanak sağlamaktadır.

RawGraps: Karmaşık verilerin görsel sunumunu herkes için kolaylaştırmak amacıyla inşa edilmiş, açık kaynaklı bir veri görselleştirme çerçevesidir. E-tablo uygulamaları (örneğin Microsoft Excel, Apple Numbers, OpenRefine) ve vektör grafik düzenleyicileri (örneğin, Adobe Illustrator, Inkscape, Eskiz) arasında bir bağlantı sağlamayı amaçlamaktadır.

R/RStudio: Otomatize gazetecilik ve veri gazeteciliği için en kullanışlı programlama

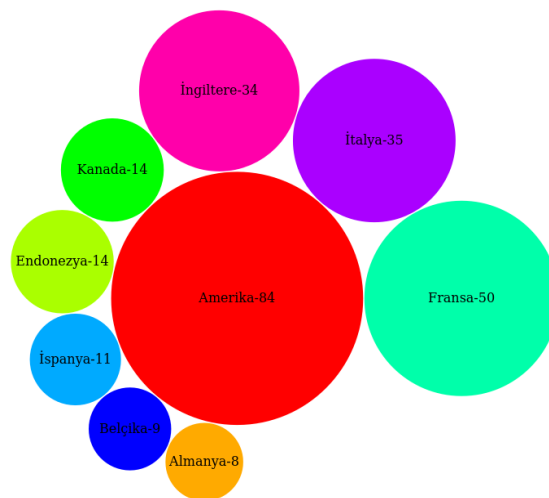
dillerinden biridir. R veri toplama, veri temizleme, veri analizi ve görselleştirmede kullanılabilir. Açık kaynak bir yazılımdır ve ücretsizdir. R'ın kaynak kodları GNU Genel Kamu Lisansı altındadır. Kullanımı bir miktar yazılım bilgisi gerektirmektedir. İstatistiksel hesaplama ve görselleştirme için uygun bir araçtır. R için geliştirilmiş olan RStudio adlı tümleşik geliştirme ortamı (integrated development environment) özellikle yeni başlayanlar için kullanım kolaylığı sağlamaktadır.

D3.js: Veri görselleştirme amacıyla kullanılan bir JavaScript kütüphanesidir. Açık kaynak kodlu olup, JavaScript dilini temel düzeyde bilmeyi gerektirmektedir.

Araştırmanın bulgular bölümünde yukarıda önerilen bu araçlar kullanılarak #datajournalism etiketi ile paylaşılan tweetler analiz edilmiştir. Böylece araçların uygulama örnekleri gösterilmiş ve veri gazeteciliğinin uluslararası alanda kullanım pratikleri ile yaygınlaşma seyri Twitter paylaşımları üzerinden saptanmaya çalışılmıştır.

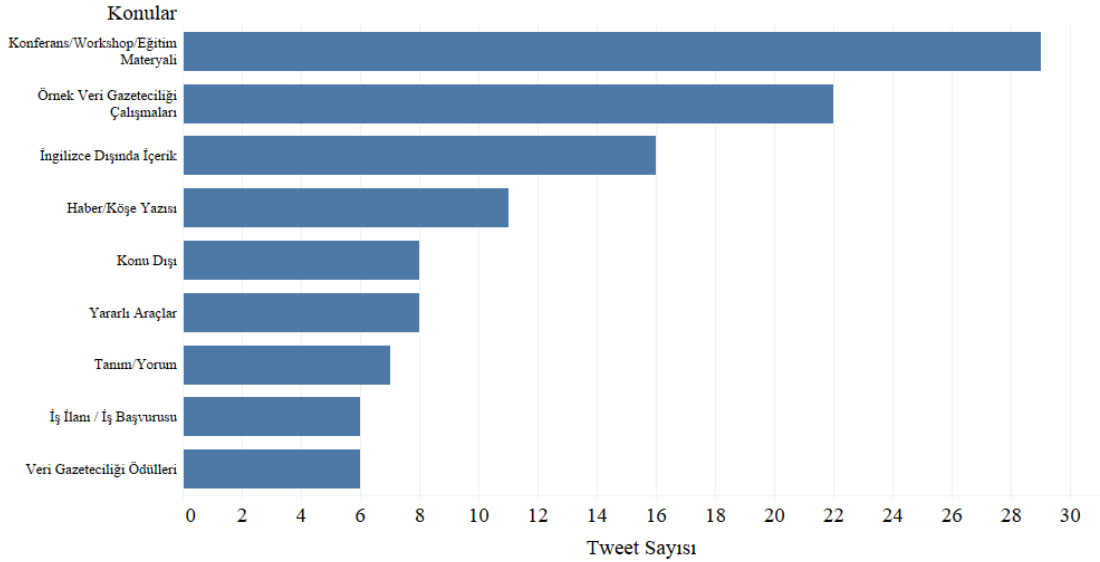
Bulgular

Çalışmanın ampirik bölümünde veri gazeteciliği alanındaki güncel tartışma konularını saptamak, alanın coğrafi gelişim haritasını ortaya çıkarmak ve örnek uygulamaları kategorize etmek için Twitter'da #datajournalism etiketi ile paylaşılan tweetler içerik analizi yöntemi ile analiz edilmiştir. Süre kısıtı nedeniyle tüm tweetler analiz edilememiş; araştırma örnekleme NodeXL programının tesadüfi seçimle sağladığı tweetlerle sınırlı tutulmuştur. Bu kapsamda, ilgili programda 12.02.2017 tarihinde #datajournalism etiketi için arama yapılarak; programın sağladığı 628 tweet kayıt altına alınmıştır. Bu tarihin seçilme nedeni veri gazeteciliği kavramının tanınırlığının artması, örnek çalışmaların yaygınlaşması, ödül törenleri, dernekler ve atölyeler aracılığıyla bu gazetecilik türünün dünya geneline yayılmasıdır. Örneklem olarak seçilen tweetler, oluşturulan içerik çözümlemesi kodlama cetvelindeki kategorilere göre kodlanmıştır. Veriler tweetlerin atıldığı ülkeler, tweetlerin konu dağılımı ve tweetlerde paylaşılan veri gazeteciliği örneklerinin temaları başlıkları altında; yaygın kullanılan farklı veri gazeteciliği araçları ile görselleştirilmiştir.



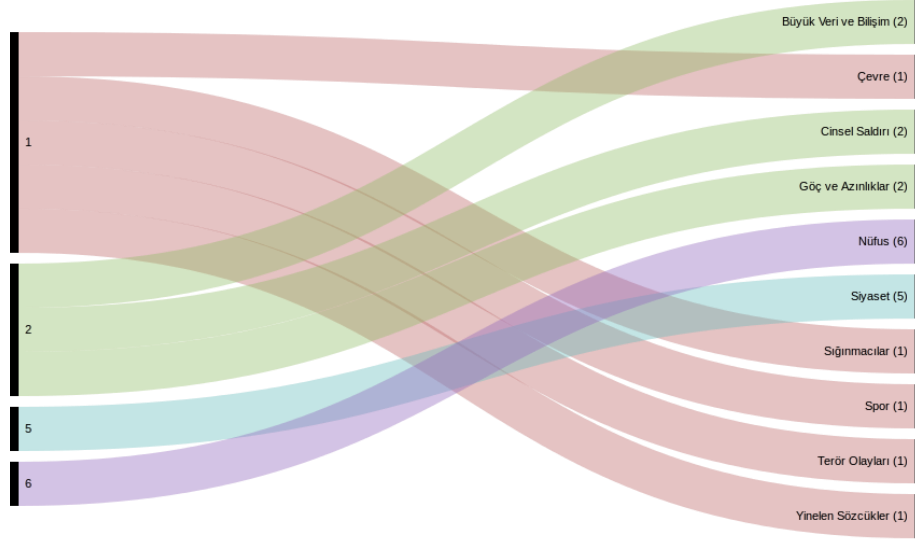
Görsel-1: #datajournalism etiketi ile en çok tweet atılan ülkeler (RStudio'da oluşturulmuştur)

Araştırmada #datajournalism etiketi kullanılarak atılan toplam 628 tweet analiz edilmiş, 479'unun yer bildirimi içerdiği gözlenmiştir. Bu bildirimlerden 77'si, coğrafi konum yerine "somewhere near food" (yiyeceğe yakın bir yerler), "right behind you" (hemen arkadayım) gibi ifadelerden oluşmaktadır. Bu bildirimler temizlendikten sonra, kalan 402 yer bildirimli içerik incelenerek; en çok tweet atılan ülkeler görselleştirilmiştir. Tweetlerdeki yer bildirimleri analiz edildiğinde, veri gazeteciliği fenomeninin Amerika Birleşik Devletleri, Avrupa ve Kanada merkezli olarak gündeme taşındığı söylenebilecektir. Veri gazeteciliğinin Wikileaks, Swissleaks, Luxleaks ve Afrikaleaks gibi yenilerinin eklendiği büyük sızıntı verilerin artması ile birlikte daha fazla uygulama alanı bulduğu ve özellikle Amerikan ve Avrupa medyasında önemli bir yere oturduğu literatürde yer alan çalışmalarla da doğrulanmaktadır (Dağ, 2015: 4). Rogers, yeniliklerin yayılımında coğrafyanın önemli bir etkisi olduğunu belirterek, bir teknolojiyi ilk benimseyenlerin "komşu etkisi" (*neighborhood effect*) ile teknolojiye yakın olanlar olduğunu söylemektedir (Rogers, 1983: 77). Bu yaklaşım, veri gazeteciliği örneğinde kısmen doğru kabul edilebilecektir. Ancak, bu ülkelerin dışında Meksika (7), Hindistan (6), Hollanda (7), Avustralya (6), İsviçre (6), Arjantin (5), Brezilya (5), Türkiye (5), Kenya (5), Portekiz (5), Venezuela (4), İsveç (3), Pakistan (3), Tanzanya (2), Rusya (2), Singapur (2), Slovakya (2), Myanmar (2), Nijerya (2), Güney Afrika (2), Costa Rica (2), Danimarka (2), El Salvador (2), Şili (2), Katar (2), Japonya (2), İrlanda (2), Gana (2), Filipinler (2), Finlandiya (2), Lübnan (2), Mali (2), Yemen (1) ve Yunanistan (1) gibi dünyanın farklı bölgelerinden çok sayıda ülkeden de konu ile ilgili tweet atılmış olması; veri gazeteciliğinin tüm dünyada tartışılmaya başlayan bir olgu olduğunu ortaya koymaktadır. Bu bulgu teknolojilerin sosyal medya üzerinden coğrafi bağımlılıkları aşarak yayılmasını göstermesi açısından dikkate değerdir.



Görsel-2: Tweetlerin Konulara Göre Dağılımı (Tableau ile oluşturulmuştur)

Görsel-2'de #datajournalism etiketi kullanılarak yazılan tweetlerin konulara göre dağılımı görülmektedir. Bu kategoride, mention ve yanıt içeren tweetler temizlenerek; kalan 113 tweet konulara göre kategorize edilmiştir. İngilizce dışındaki içerikler, araştırmacıların dil engeli nedeniyle incelemeye tabii tutulamamıştır. Veri gazeteciliği ile ilgili tweetlerin %25,6'sının konferans, workshop ve çeşitli eğitim materyalleri ile ilgili olduğu görülmüştür. Örnek veri gazeteciliği çalışmaları ile ilgili paylaşımların %19,4, İngilizce dışındaki



Görsel-4: #datajournalism etiketinde yer alan veri gazeteciliği örneklerinin temaları (RawGraphs ile oluşturulmuştur)

#datajournalism etiketi altında paylaşılan ve veri gazeteciliği örnekleri içeren 22 tweet temalarına göre incelendiğinde, en çok nüfus (6) ile ilgili görselleştirmelerin yapıldığı saptanmıştır. Görselleştirmeye sıkça konu olan ikinci tema ise siyasettir (5). Bu durumun temel nedenlerinden biri, Amerikan başkanlık seçimlerinin yakın bir zamanda gerçekleşmiş olmasıdır. Büyük veri ve bilişim, cinsel saldırılar, göç ve azınlıklar, çevre sorunları, sığınmacılar, spor, terör olayları ve yinelenen sözcükler de veri görselleştirme amacıyla kullanılan diğer temalar olarak karşımıza çıkmaktadır. Bu kapsamda, veri gazeteciliği ile ilgili çalışmaların konu çeşitliliğinin giderek arttığı ve farklı haber kategorilerine uyarlandığı gözlenmektedir. Teknolojiden siyasete, spordan gündelik yaşam haberlerine uzanan bir çeşitlilikte, hemen her kategoride veri gazeteciliği örneğine rastlanmaktadır.

Sonuç

İnternet teknolojileri ile birlikte gazetecilik ve görsel düşünme yetilerinin yakınsadığı bir habercilik türü olarak veri gazeteciliği yaygınlık kazanmaya başlamıştır. Bu gazetecilik türünü destekleyen en önemli etmenlerden biri ise büyük veri olgusudur.

Büyük veri kişisel bilgilerin güvenliğinin sağlanamaması, gazetecinin failliğinin ortadan kalkması, veriye ulaşmadaki eşitsizlikler ve kaynak olarak resmi kanallara fazlaca bağımlılık gibi bir dizi sorunu da beraberinde getirmektedir. Algoritmaları oluşturanların güç sahipleri olması nedeniyle gazetecilerin büyük veriyi güç sahiplerinin çıkarlarına hizmet edecek şekilde kullanabileceği ve büyük verinin toplumun çoğunluğunun yanı sıra azınlıkları da gerçekten temsil edip etmeyeceği sorunları da üzerinde düşünülmesi gereken konulardır. İnternet ortamında gazetecilerin hız baskısı altında çalışması, eğitilmiş ve donanımlı çalışan sayısındaki eksiklik gibi faktörler de veri gazeteciliğinin gelişimi ve yaygınlık kazanması üzerinde olumsuz etkiye neden olmaktadır.

Yukarıda sıralanan sorunlara karşın, veri bolluğu altındaki gazetecilik faaliyetlerinde değişim gözlenmektedir. Bir yandan veriyi temizleme, işleme ve ondan anlamlı sonuç elde

etme süreçlerinde etkili olabilecek diğer iş kolları haber odalarına dahil olurken; öte yandan profesyonel gazeteciler de bu iş kollarının yaptığı işleri günlük iş yapış pratiklerinin bir parçası olarak öğrenmeye ve kullanmaya çabalamaktadır.

Bu araştırmada büyük veri ile birlikte gazetecilik pratiklerinde yaygınlık kazanan veri gazeteciliğinin bir yenilik olarak yayılımı, Twitter’da #datajournalism etiketi örneğinde değerlendirilmeye çalışılmıştır. Veri gazeteciliği ile ilgili konferans/workshop ve eğitim materyalleri ile veri gazeteciliği örneklerinin sıklıkla paylaşıldığı görülmüştür. Ayrıca Twitter’ın veri gazeteciliği ile ilgili örneklerin ve kullanılabilir yararlı araçların duyurulduğu; konferans çağruları ve iş ilanlarının dolaşıma sokulduğu bir mecra olarak türün yaygınlaşmasında etkili bir iletişim kanalı olduğu saptanmıştır. Bu etiket ile paylaşılan tweetlerde en çok yinelenen sözcükler incelendiğinde de, tweetlerin konu dağılımını desteklediği, eğitim ve konferans konulu içerikleri destekleyecek şekilde “eğitim”, “ders vermek”, “sınıf”, “öğrenmek” gibi sözcüklerin sıklıkla yinelenildiği görülmüştür. Hem tweetlerin konuları hem de yinelenen sözcükler incelendiğinde, veri gazeteciliği ile ilgili konferans ve eğitimlerin önemsendiği söylenebilecektir.

Araştırma ile yeniliklerin yayılımında teknolojiye yakınlığın etkisi kısmen doğrulanmış, konu ile ilgili tweetlerin en çok Amerika Birleşik Devletleri, Avrupa ve Kanada’dan yazıldığı bulgulanmıştır. Ancak sosyal medya gibi iletişim araçlarının da etkisiyle yeni teknolojilerin coğrafi sınırları aşarak, kademeli de olsa çok sayıda dünya ülkesine yayıldığı gözlenmiştir. Ayrıca Twitter’ın veri gazeteciliği ile ilgili örneklerin ve kullanılabilir yararlı araçların duyurulduğu; konferans çağruları ve iş ilanlarının dolaşıma sokulduğu bir mecra olarak türün yaygınlaşmasında etkili bir iletişim kanalı olduğu saptanmıştır.

Kaynakça

- Appelgren, E. ve Gunnar, N. (2014). Data Journalism in Sweden, *Digital Journalism*, 2:3, 394-405.
- Beak, S., Jay L., ve Marshall L. (2000). “An Exploratory Study: Supporting Collaborative Multimedia Systems Design.”, *Multimedia Tools and Applications*, 12, 189–208.
- Beyer, M.A. ve Laney, D. (2012). The Importance of ‘big data’: A definition, *Gartner Report*, Erişim Tarihi: 18.12.2017 from <https://www.gartner.com/doc/2057415/importance-big-data-definition>.
- boyd, d., Crawford, K. (2015). Büyük Veri Üzerine Eleştirel Sorular: Kültürel, Teknolojik ve Bilimsel Bir Olgu Hakkında Eleştirel Sorgulamalar, Çev: Savcı, B. *Folklor/Edebiyat*, 21 (83), 199-215.
- Bradshaw, Paul. (2011). Data Journalism Survey: A Mixed Picture, Erişim Tarihi: 21.02.2017. http://datadrivenjournalism.net/news_and_analysis/data_journalism_survey_analysis,
- Chen, H., Chiang, R. H. L., & Storey, V. C. (2012). Business intelligence and analytics: From big data to big impact. *MIS Quarterly*, 36 (4), 1165–1188.
- Crawford, K., Miltner, K., & Gray, M. L. (2014). “Special Section Introduction”, *International Journal of Communication*, 8, 1663-1672.
- Dağ, P. (2015). Gazeteciliğin Geleceği: Veri Gazeteciliği, Erişim Tarihi: 17.02.2017, https://www.academia.edu/16286677/GAZETECİLİĞİN_GELECEĞİ_VERİ_GAZETECİLİĞİ_-

Pınar_Dağ_Ekitap.

Davenport, T. (2014). Big data @ work, çev. M. Çavdar, İstanbul: Türk Hava Yolları Yayınları.

Dick, M. (2014). "Interactive Infographics and News Values", *Digital Journalism*, Vol. 2, No. 4, 490-506.

Diebold Francis X. (2012). A Personal Perspective on the Origin(s) and Development of Big Data: The Phenomenon, the Term, and the Discipline, Erişim Tarihi: 17.02.2017, http://economics.sas.upenn.edu/~fdiebold/papers/paper112/Diebold_Big_Data.pdf,

Doğu, B. (2015). Veri Haberciliği: Demokratik Medya İçin Olanaklar, *Folklor/Edebiyat*, 21(83), 181-197.

Dumbill, E. (2013). Making Sense of Big Data, *Big Data*, Vol.1, Issue:1, 1-3.

Fairfield, J., & Shtein, H. (2014). Big data, big problems: Emerging issues in the ethics of data science and journalism, *Journal of Mass Media Ethics*, 29(1), 38-51.

Gandomi, A. ve Haider M.(2015). Beyond the hype: Big data concepts, methods, and analytics. *International Journal of Information Management*, 35(2), 137-144

Giardina, Marco, and Pablo Medina. 2013. Information graphics design challenges and work flow Management. *Online Journal of Communication and Media Technologies*, 3(1), 108-124.

Gray, J., L. Bounegru ve L. Chambers (2012). *The Data Journalism Handbook*, Sebastopol, CA: O'Reilly Media.

Gürsakal, N. (2014). *Büyük Veri*, 2. Baskı, Bursa: Dora.

Howard, A. (2014). The Art and Science of Data-driven Journalism. *Tow Center for Digital Journalism*, New York, Columbia University.

Hoy, B. (2014). Big data: An Introduction for Librarians, *Medical Reference Services Quarterly*, 33(3), 320-326.

Jagadish, H.V., Gehrke, J., Labrinidis, A., Papakonstantinou, Y., Patel, J., Ramakrishnan, J.M.R. ve Shahabi, C. (2014). Big Data and Its Technical Challenges, *Communications of the ACM*, 57 (7), 86-94.

Kirk, A. (2012). *Data Visualization: A Successful Design Process*, Birmingham-Mumbai: Packt Publishing.

Knight, M. (2015). Data Journalism in the UK: A Preliminary Analysis of Form and Content, *Journal of Media Practice*, <http://dx.doi.org/10.1080/14682753.2015.1015801>, 1-18.

Kraemer, F., Van Overveld, K., & Peterson, M. (2011). Is there an ethics of algorithms?, *Ethics and Information Technology*, 13(3), 251-260.

Kumar P. ve Pandey K. (2013). Big Data and Distributed Data Mining: An Example of Future Networks, *International Journal of Advance Research and Innovation*, 1 (2), 36-39.

Kwon, O.; Lee, N. ve 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.

Laney, Doug (2001). 3D Data Management: Controlling Data Volume, Velocity, and Variety, *Application Delivery Strategies*, 6 February 2001, File: 949, 1-4.

- Lewis, S. C. (2015). Journalism in an Era of Big Data: Cases, concepts, and critiques, *Digital Journalism*, Vol.3, 321-330.
- Lewis, S., ve Usher, N. (2014). Code, collaboration, and the future of journalism: a case study of the hacks/hackers global network, *Digital Journalism*, 2(3), 383–393.
- Lewis, S. C. (2011). Journalism Innovation and Participation: An Analysis of the Knight News Challenge, *International Journal of Communication*, 5, 1623–1648.
- Mahrt, M. ve Scharkow, M. (2013). The value of big data in digital media research, *Journal of Broadcasting & Electronic Media*, 57(1), 20-33.
- Mateosian, R. (2013). Ethics of Big Data, *IEEE Micro*, 33(2), 60-61.
- Monino, J. ve Sedkaoui S. (2016). Big Data, Open Data and Data Development, Volume 3, London, Wiley.
- Mora, G. (2012). Journalists vs Infographic designers-I, Erişim Tarihi: 01.03.2017, <http://visualoop.com/blog/3134/journalists-and-infographic-designers-a-love-hate-relationship-part-one>.
- Narayanan, V. (2014). Using big-data analytics to manage data deluge and unlock real-time business insights, *Journal of Equipment Lease Financing*, 32 (2), 1-7.
- Nygren, G., E. Appelgren, ve H. Hüttenrauch (2012). Datajournalistik-Ett växande område, *Nordicom Information*, 34 (3–4), 1–19.
- Open Government Partnership, Erişim Tarihi: 20 Şubat 2017, <http://www.opengovpartnership.org/country/turkey>
- Parasie, S., ve Dagiral, E. (2012). Data-driven journalism and the public good: “Computer-assisted-reporters” and “programmer-journalists” in Chicago, *New Media & Society*, 15 (6), 853-871.
- Parasie, S. (2014). Data-Driven Revelation?, *Digital Journalism*, DOI:10.1080/21670811.2014.976408, 1-17.
- Parks, R. M. (2014). Big data in communication research: its contents and discontents, *Journal of Communication*, 64(2014), 355-360.
- Rogers, E. (1983). *Diffusion of Innovations*, Third Edition, London, The Free Press.
- Rogers, S. (2008). Turning Official Figures into Understandable Graphics, at the Press of a Button, *Inside the Guardian Blog*, Newspaper, Erişim Tarihi: 18.12.2008, <http://www.Guardian.co.uk/help/insideGuardian/2008/dec/18/unemploymentdata>
- Rogers, S. (2014). Data Journalism is the New Punk, *British Journalism Review*, 2014/6, 25:2, 31-34.
- Schönberger M.V. ve Cukier K. (2013). Büyük Veri: Yaşama, Çalışma ve Düşünme Şeklimizi Dönüştürecek Bir Devrim, çev. B.Erol, İstanbul: Paloma Yayınevi.
- Schroeder, R., & Cows, J. (2014). Big data, ethics, and the social implications of knowledge production, *Data Ethics Workshop*, KDD@ Bloomberg, August (Vol. 24).
- Segel, E. ve J. Heer (2010). Narrative Visualization: Telling Stories with Data, *Journal IEEE Transactions*

on *Visualization and Computer Graphics*, 16 (6), 1139–1148.

Smit, G., Yael de H., ve Laura B. (2014). Visualizing News: Make it Work, *Digital Journalism*, 2 (3), 344–354.

Smit, Gerard, Yael de Haan ve Laura Buijs (2014). Visualizing News, *Digital Journalism*, Vol. 2, No. 3, 344–354.

Stavelin, E. (2014). Computational Journalism: When Journalism Meets Programming, Yayınlanmamış Tez, Bergen: University of Bergen.

Stone, L. M. (2014). Big data for media (Report), University of Oxford: Reuters Institute for the Study of Journalism.

The Open Knowledge Foundation. (2012). Erişim Tarihi: 20 Şubat 2017, <https://okfn.org>.

Wang, J., Wu, Y., Yen, N., Guo, S. ve Cheng, Z. (2016). “Big data analytics for emergency communication networks: a survey”, *IEEE Communications Surveys & Tutorials*, Vol. 18(3), 1758–1778.

Ward, S. J. ve Barker A. (2013). “Undefined By Data: A Survey of Big Data Definitions”, Erişim Tarihi: 18.02.2016 <https://arxiv.org/abs/1309.5821>.

Weber, Wibke, and Hannes Rall (2012). Data Visualization in Online Journalism and Its Implications for the Production Process, *Paper for the 16 International Conference on Information Visualization*, Erişim Tarihi: 18.02.2016, <http://ieeexplore.ieee.org/document/6295837/>

Zotto, C., Schenker, Y. ve Lugmayr, A. (2015). “Data journalism in News Media Firms – The role of Information Technology to Master Challenges and Embrace Opportunities of Datadriven Journalism Projects”, Twenty-Third European Conference on Information Systems (ECIS), Münster/Germany, 1-10.