



Büyük Verinin Sağlık Hizmetlerinde Kullanımında Epistemolojik ve Etik Sorunlar

Epistemological and Ethical Issues of Big Data Use in Healthcare

Abdullah Uçar¹ , İlhan İlkılıç² 

öz

Büyük Veri (*big data*) dendiğinde geleneksel bilişim sistemlerinin kapasitesinin çok üstünde olan verilerin depolama ve analiz edilerek bilgi üretme süreçlerinin tümünü anlamak mümkündür. “Büyük” ifadesi verinin hacim büyüklüğünden ziyade 5 temel özelliğın büyüklüğünü ifade eder: Hacim, Hız, Çeşitlilik, Doğruluk, Değer. Büyük veri analizleri ile mevcut bilgi ve verilerin ilişkileri tespit edilebilmekte, yeni değişkenler arasındaki bağ hızla tespit edilebilmekte ve böylece gelecekle ilgili sağlık alanında kuvvetli tahminler yapılabilmektedir. Bu bağlamda şimdiye kadar görülmemiş büyüklükteki verinin halk sağlığından klinik bilimlere kadar birçok sağlık hizmeti alanında yeni kullanım ve uygulama imkanları oluşmaktadır.

Yeni imkanların yanında tıbbi bilginin oluşmasında epistemolojik olarak önemli farklılıklar ortaya çıkmakta ve bunlar uygulamada bazı sorunları beraberinde getirmektedir. Ayrıca çok sayıda insanın oluşturduğu verilerin güvenli şekilde saklanması gerekmektedir. Burada ise tıbbın kadim etik ilkelerinden olan mahremiyet ve hekimin sır saklama mükellefiyeti ilkelerinin anlamı ve içeriği değişmektedir. Bu bağlamda makalemizde büyük verinin tarihi gelişimi, temel özellikleri ve önemli kavramları açıklanmıştır. Sağlık alanında mevcut ve muhtemel kullanım alanları tartışılarak bu alanlardaki bilgi üretimi ile ilgili epistemolojik sorunlar tespit edilerek eleştirel bir perspektiften ele alınmıştır. Yine bu meyanda ortaya çıkabilecek etik sorunlar ortaya konmuş ve analiz edilmiştir. Meta düzlemde ise bu uygulamaların insan anlayışımız, insan özgürlüğü ve sorumluluğu açısından ne anlama geldiği sorgulanmıştır.

Anahtar Kelimeler: Büyük veri, sağlık hizmetleri, etik, epistemoloji, veri hakimiyeti

ABSTRACT

The concept of “Big Data” covers all processes of storing and analyzing information that is beyond the capacity of traditional information systems. “Big” describes the magnitude of five major characteristics of data and not volume alone: Volume, Velocity, Variety, Veracity, Value. With the help of big data analysis, it is possible to identify the relationships between existing medical information and data, to determine the link between new variables and to make strong predictions about the future healthcare. In this context, new usage and application opportunities arise in various healthcare services ranging from public health to clinical sciences.

In addition to new opportunities, big data analysis also presents new epistemological differences in the formation of medical knowledge, possibly engendering new controversies. Furthermore, the data generated by a large number of people need to be safely stored. The challenge of data security is set to alter the meaning and content of two well-established principles of medicine, “patient-privacy” and “physician-patient confidentiality”. In this context, we explain the historical development of big data, its basic features and important related concepts. The article further treats the current and potential uses of big data analysis in the field of health, and identifies and critically analyzes epistemological problems related to medical knowledge production. We also discuss the ethical problems that may arise. At the metalevel, we ask how the potential applications of big data analysis can impact our understanding of humanhood, human freedom and responsibility.

Keywords: Big data, healthcare, ethics, epistemology, data sovereignty

¹ Dr., İstanbul Üniversitesi, Sağlık Bilimleri Enstitüsü, Halk Sağlığı Doktora Programı öğrencisi, İstanbul, Türkiye

² Prof. Dr., İstanbul Üniversitesi İstanbul Tıp Fakültesi, Tıp Tarihi ve Etik Anabilim Dalı, Öğretim Üyesi, İstanbul, Türkiye

ORCID: A.U. 0000-0002-0220-3720;
İ.İ. 0000-0002-4250-8676

Sorumlu yazar/Corresponding author:

Abdullah Uçar,
İstanbul Üniversitesi, Sağlık Bilimleri Enstitüsü,
Halk Sağlığı Doktora Programı öğrencisi,
İstanbul, Türkiye
E-posta: abduhahucar@gmail.com

Başvuru/Submitted: 06.05.2019

Revizyon Talebi/Revision Requested: 16.05.2019

Son Revizyon/Last Revision Received: 16.05.2019

Kabul/Accepted: 27.06.2019

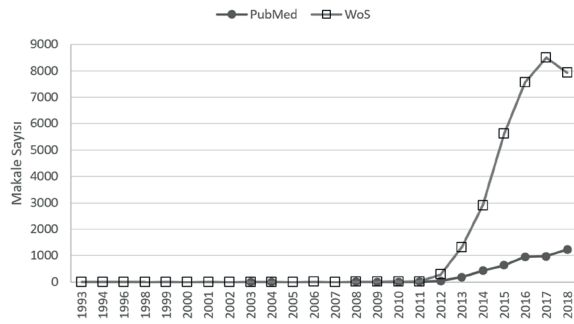
Atıf/Citation: Ucar A., İlkılıç I. (2019):

Epistemological and ethical issues of big data use in healthcare, *Sağlık Bilimlerinde İleri Araştırmalar Dergisi*, 2(2): 80-92.

<https://doi.org/10.26650/JARHS2019-616389>

GİRİŞ

Büyük Veri (Big Data), bir kavram olarak ilk kez 1997'de bir makalede kullanılmıştır.¹ Kavramın bilimsel bir sunum başlığı olarak ilk kullanımı ise 1998'dir.² Teknik bir kavram olmaksızın kullanımına dair erişebildiğimiz en eski literatür kaydı ise 1993 tarihli bir makalede bulunmaktadır.³ Kavramın PubMed aramasındaki 5.034 ve Web of Science (WoS) veri tabanı aramasındaki 36.025 sonucun zamana dağılımı Şekil 1'de görülmektedir. (Şekil-1)



Şekil 1. PubMed ve Web of Science (WoS) veri tabanlarında 29.04.2019 tarihli, ("big data" OR bigdata) kodlu arama sonuçlarının zamana göre dağılımı

Bu bilgilere göre kavram olarak büyük verinin 2000'li yılların başında şekillendiği söylenebilir. Grafik üzerinde de görüldüğü gibi 2012 yılından sonra yayın sayısında ivmeli bir yükseliş söz konusudur. Şüphesiz bu değişime sebep olan en önemli etkenlerden birisi de 2012 yılındaki Davos Zirvesi'nde "veri"nin ekonomik bir değer olduğunun ifade edilmiş olmasıdır.⁴ Büyük veri, ilk başta verinin hacimsel büyüklüğünü ifade etmek için kullanılan bir kavramken, daha sonra verinin depolanmasından bilgiye dönüşmesine kadarki tüm süreçlerin büyüklüğünü ifade eden bir anlam kazanmıştır. Büyük veri kavramı yerine "büyük veri analizi" (Big Data Analytics) kavramı da sık kullanılmaktadır.

Büyük veri analizinin temelleri geleneksel istatistik yöntemlere dayanır. Kavram yeni bir istatistik yöntem sunmaktan ziyade, çok büyük ve çeşitli veri setlerini yüksek hızla analiz edebilen algoritmalar içermektedir. Veri setlerinin çeşitlerine ve veriden elde edilmek istenen bilginin türüne göre seçilen algoritma veri setine uygulanmakta, böylece veri-

nin içinde saklı olan desenler, değişkenler arasındaki korelasyonlar, geleceğe yönelik tahminler elde edilmektedir. Geleneksel istatistik uygulamalarında yaygın olarak ana yığından seçilen örneklem üzerinde belirli hata payı ile işlem yapılırken büyük veri analizi doğrudan ana yığınla çalışmayı ve gerçek zamanlı analizleri mümkün kılmaktadır.

Büyük veri analizinin özellikle veri toplayan, depolayan ve bu veriyi işleme ihtiyacı duyan Google, Amazon, Twitter, Facebook, LinkedIn gibi büyük şirketlerin çok büyük veri setleriyle baş edebilmek için geliştirdikleri sistemlerin içinde doğduğu söylenebilir. Verinin depolanması, işlenmesi gibi süreçlere dair teknolojik gelişmelerle artan devasa veri miktarı artık bir "dijital dünya" içinde yaşadığımızı göstermektedir. Üstelik bu dijital dünyanın boyutlarındaki büyüme, büyük bir hızla gerçekleşmektedir. Bir araştırma şirketine göre 2020'de dijital dünyanın boyutu, 2009 yılına göre 44 kat daha büyük olacaktır.⁵ Bu bilgiler büyük veri kavramının günümüzdeki ve yakın gelecekteki önemini göstermektedir.

Kavramsal Çerçeve

Tekili "datum", çoğulu "data" kavramı ile ifade edilen veri kavramına dünyanın yeni petrolü gözüyle bakılmaktadır.⁶ Bu yeni petrol, birçok faktör arasında insan zekasının tespit edemediği korelasyonları ve nedensellikleri tespit etme imkanı sunan ve geleceği tahmin etmeye kapı aralayan büyük veri analizinin temel girdisini oluşturmaktadır. Veri kavramının ilk kez ortaya çıkışı ise 1646 yılına kadar uzanır.⁷

Tarih boyunca insanlığın mal ve hizmet üretiminin miktarında çeşitli kırılma noktaları olduğunu görmekteyiz. Yeryüzünde insan varlığının başladığı günden bugüne kadar sayısız devrimsel gelişme söz konusu olmakla birlikte günümüzde teknolojiye başı çeken ülkeler kendi gelişmelerine bakan yönüyle bu devrimleri kategorize etmektedirler. Almanya'nın Endüstri 4.0'ı, Japonya'nın Toplum 5.0'ı, İngilizlerin Catapult programı buna örnek verilebilir. Endüstri 4.0 yaklaşımına göre sanayi devrimiyle insanlık buharlı makine gücüne geçiş yapmış ve üretim miktarı büyük bir artış göstermiştir. Elektrik

ve seri üretim bandının icadı ikinci, programlanabilir elektronik devrelerin ve bilgisayarların icadı ise üçüncü kırılma noktasıdır. Günümüzde veriyi önemli yapan, büyük veri analizini gündeme taşıyan sebep ise bu yazının yazılması sırasında yaşanmakta olan dördüncü büyük kırılmanın (endüstri 4.0) sunduğu perspektiftir.⁸ Endüstri 4.0 çağının içerdiği büyük veri analizi, makine öğrenmesi, doğal dil işleme, nesnelerin interneti, bulut bilişim, yapay zeka, artırılmış zeka gibi birçok yeni kavramla ifade edilen sistemler temelde “veri” kullanmakta ve karar süreçlerinin çok daha doğru yapılmasını sağlamaktadırlar. Büyük hacimli ve çok çeşitli şekillerdeki verinin hızlı ve doğru analizi, dünyadaki tüm karar alma süreçlerine yeni bir bakış açısı sunmuş ve endüstriyel değer üretiminde patlayıcı bir büyüme oluşturmuştur.

Verinin işlenmesiyle bilgiye dönüşümü karmaşık bir süreç içerisinde gerçekleşmektedir. Bu süreç, “veri hiyerarşisi” olarak tanımlanır ve bu kavram ilk kez 1982 yılında kullanılmıştır. “Veri (data) – Enformasyon (information) – Bilgi (knowledge)” şeklinde sıralanan bu süreçteki ilk basamağı verilerin düzenlenmesi ile enformasyon oluşturur. Enformasyonların mevcut tecrübelerle yorumlanması ile bilgi ortaya çıkar. Bilgisayar bilimleri bu sürecin en başına “sinyal” kavramını da ekler.⁹

Büyük Verinin Temel Özellikleri

Büyük veri, genel olarak geleneksel bilgisayar sistemlerinin kapasitesini aşan veridir. Büyüklükten kasıt, mevcut işlem kapasitesinin aşılmış olmasıdır. Örneğin 1 megabyte boyutundaki bir yazı dosyasını kolaylıkla açabilen bir bilgisayar, 100 megabyte boyutundaki yazı dosyasını açmakta zorlanmakta, daha yüksek boyutlarda ise hata vermektedir. Bu veri, bilgisayar için büyük veridir. Ancak 100 megabyte video kaydı bilgisayar için büyük veri değildir. Bir kanser hastası için uygun tedavi aranırken kişinin DNA dizilenmesi sırasında oluşan veriyi işlemek ve aranan mutasyonları hasta kötüleşmeden önce bulmak hızlı analiz yapmayı gerektirmektedir. Bu hız, geleneksel sistemler için çok uzun süreler almakta ve veri bilgiye -dolayısıyla değere- dönüşmemektedir.

Büyük veri dendiğinde ilk olarak verinin miktarındaki büyüklük akla gelmektedir. Ancak buradaki büyüklük daha da derin bir anlam içermektedir. Yapılan bir çalışma, Büyük Veri kavramı ile birlikte kullanılan en sık ve öne çıkan kavramları “hacim” (volume), “hız” (velocity), “çeşitlilik” (variety), “karmaşıklık” (complexity), “depolama” (storage), “enformasyon” (information), “işlem” (process), “veritabanı” (database), “karar” (decision making), “değer” (Value) şeklinde sıralamaktadır.¹⁰ Literatürde büyük verinin genellikle beş temel özelliği ifade edilmektedir. Bunlar hacim, hız, çeşitlilik, doğruluk ve değer özellikleridir.

Hacim (Volume)

Büyük veri hacim olarak çok büyük olup, lineer veya geometrik bir artışla değil, artık üstel (x^n) bir artışla yani patlayıcı şekilde büyümektedir.¹¹ Tüm SMS’ler, e-postalar, fotoğraflar, ses kayıtları, videolar, sağlık kayıtları, DNA dizileme çalışmaları, tıbbi görüntülemeler, çağrı merkezi aramaları, makine verileri sürekli depolanmaktadır. IBM’e göre 2013 yılında dünyadaki tüm verinin %90’ı son iki yılda üretilmiştir.¹² 2010 yılında ilk kez veri miktarı zeta-byte düzeyine erişmiş, 2016 yılında ise 16 zettabyte ulaşmıştır. Seagate firmasının 2015 yılında yayınladığı raporda dünya üzerindeki büyük veriden “küresel datasfer” olarak bahsedilmektedir.¹³ Sadece Facebook üzerinde aylık aktif kullanıcı sayısı 2 milyar 230 milyondur ve kullanıcılar bir günde 350 milyon fotoğraf yüklemektedir.^{14,15}

Bu büyüklükler, içinde yaşadığımız dijital dünyanın da hızla genişlediğini göstermektedir. Bu kadar büyük verinin depolanması, şu ana kadar kullanılan geleneksel yöntemleri aşmaktadır. Verilerin depolanması için veri tabanları ve veri ambarları kullanılmaktadır. Veri tabanları genellikle yapılandırılmış verilerin satır ve sütun şeklinde düzenlendiği ve bir yazılımla kullanıcıya sunulduğu araçlardır. Veri ambarları ise çok farklı yapıdaki verilerin depolandığı ve anlık olarak işlenebildiği araçlardır. Verilerin yazılımlar tarafından tanınmasını sağlayan ve verinin kimliği olarak adlandırılan veriler ise metaveri kavramı ile ifade edilmektedir. Örneğin bir hastanın baş X-ray grafisinde hangi tarafın

sağ veya sol olduğu bir metaveridir. Buna dayanarak klinisyen filmi yorumlayacaktır. Metaveriler de veri miktarındaki artışın önemli bir bileşenidir.

Büyük miktardaki verinin üretilmesinde öne çıkan bir kavram da nesnelerin interneti (IoT) kavramıdır. Bu kavram, birbirine bağlanabilen elektronik cihazların ve cihazlar arası bağlantıların sayısını, oluşturdukları ağı yapısını, ağlar arası kurgulanan ekosistemlerin ürettiği verileri ifade etmek için kullanılmaktadır. Bir araştırma şirketinin raporuna göre dünya üzerinde 2012 yılında 8,7 milyar bağlanabilir cihaz varken, bu rakamın 2020 yılında 40 milyar cihaz olması beklenmektedir.¹⁶ Giyilebilir dijital aksesuarlar, akıllı ev aletleri, arabaların dijital aksamaları ve geliştirilmekte olan insansız arabalar ve daha bir çok akıllı sistem sürekli veri üretmekte, bu veriler işlenmeyi ve bilgiye dönüşmeyi, böylece değer üretilmesini gerektirmektedir.¹⁷

Hız (*Velocity*)

Büyük veriyi kayıt altına alıp depolamak tek başına bir anlam oluşturmamaktadır. Verinin üretildiği hızda işlenmesi ve bilgiye dönüşmesi gerekmektedir. Bir kalp hastasının tüm sağlık kayıtlarının analizi yapılarak kalp krizi geçirme riskinin hesaplanması, kalp krizi geçirmeden yapıldığında anlamlıdır. Diğer taraftan bir gebeliğin erken doğum riskini tahmin etmek, tüm kayıtların belirli bir hızda işlem görmesini zorunlu kılar. Üstelik tüm bu hesaplamaların ve analizlerin milyarlarca insan için kısa bir süre içinde yapılması büyük bir hız ve işlem kapasitesi gerektirir. Böylece devasa verilerin hızlı bir şekilde işlenmesi de büyük verinin önemli özelliklerinden biridir.

Hız sorununa çözüm olarak önemli teknolojik gelişmeler yaşanmıştır. Eskiden büyük miktarda verinin hızlı işlenebilmesi geleneksel işlemcilerle mümkün olmadığından her kapasite aşımında veri sahipleri daha gelişmiş işlemcilerle kendilerini yenilemek durumunda kalmışlardır. Ancak her işlemci yenilenmesi tüm sistem mimarisinin de değişmesini zorunlu kılmakta ve bu sürdürülemez bir ortam oluşturmaktadır. Bir işlemci yetmediğinde süper işlemci almak, o da yetmeyince daha süper işlemci almak gelecek vadetmeyen bir yaklaşımdır. Bunun

yerine, artan ihtiyaca göre modüler olarak tasarlanmış, işlemcilerin birbirine eklenmesiyle devasa işlem kapasitesine sahip sistemler geliştirilmiştir. MapReduce ve Hadoop isimli bu öncü sistemler, büyük veri analizlerinde bir kırılma noktasıdır. Bu sistemler sayesinde veri küçük parçalara ayrılabilir ve eş zamanlı çalışan birçok işlemcinin veriyi işlemesi mümkün hale gelmiştir.¹⁸

Çeşitlilik (*Variety*)

Verinin depolanması ve hızlı işlenmesi sorununa ek olarak büyük verinin heterojen yapıda olması bir başka sorundur. Zettabyte düzeyinde ve büyük bir hızla üretilen veriler yazılı olabilmekte veya göze ve kulağa hitab edebilmektedir. Örneğin tıp alanında bir hastanın anamnezinde elle tutulan notlardan bir yenidoğanın APGAR skoruna, boy ölçümlerinden konsültasyon raporlarına kadar geniş bir çeşitlilikte dağılmış verinin bütüncül bir sistem içinde yorumlanması bir sorun teşkil etmektedir. Geleneksel sistemlerde yapılandırılmış, yani satır ve sütunlara ayrılmış verilerde işlem yapılması zor değildir ancak yapılandırılmamış, yani düz metinler, e-postalar, konsültasyon notları gibi satır ve sütunlara ayrılmamış verilerde işlem yapmak çok daha zordur. Mevcut verilerin %13'ünün yapılandırılmış, %87'sinin ise yapılandırılmamış ve geleneksel metodlarla işlenmeyen veri olduğu da düşünüldüğünde bilgiye dönüştürülmede kullanılmayan / israf olan veri miktarının ne kadar büyük olduğunu tahmin etmek mümkündür. Sağlık sektöründe ise tüm verilerin %50'sinin yapılandırılmış olduğu ifade edilmektedir.¹⁹ Veri çeşitliliğinin fazla sayıda olması da büyük verinin temel özelliklerindedir.

Doğruluk (*Veracity*)

Geleneksel ve küçük veritabanlarında kullanım amaçlarına uygun olarak veriler şekillendirilmiş olduğundan hassas bilgiler yüksek doğrulukla elde edilebilmektedir. Ancak büyük verileri içeren veri ambarlarında veri belirli bir amaç doğrultusunda depolanmadığı için yapılan analizler hata payları içermektedir. Bu hata paylarını en aza indirmek için *veri madenciliği* başlığı altında birçok algoritma geliştirilmiştir. Büyük veri analizinde belirli hata

payları ile tahminler yapılmakta ve veri içindeki desenler tespit edilebilmektedir. Örneğin Visa firması, büyük veri imkanlarını kullanarak yaklaşık 73 milyar işlemi içeren 2 yıllık kayıtları için gereken işlem süresini yaklaşık bir aydan 13 dakikaya indirmiş ve doğru bir şekilde gerçekleştirmiştir.²⁰ Bu bağlamda verinin doğruluğunu artırmak için ön veri analizleri ve hazırlığı gerekmektedir.

Değer (Value)

Diğer 4 özelliğin bir bileşkesi olarak ortaya çıkan kavram değer kavramıdır. Her yönüyle büyük olan veriden elde edilecek değer de büyük olması beklenir. Veriden değere giden bu süreçte değer ortaya çıkmasını olumlu ve olumsuz etkileyen birçok bileşen bulunmaktadır. Bu bileşenler bu başlık altında tartışılmakta ve incelenmektedir.

Büyük Verinin Sağlık Hizmetlerinde Kullanım Alanları

Büyük veri sağlık sektöründe birçok uygulama alanı bulmaktadır. Özellikle büyük veri sahibi şirketlerin sağlıkla yakından ilgilenmeleri, sağlığın çok fazla ve karmaşık belirleyicilerden oluşması, dolayısıyla büyük veri analizinin en çok değer üretebileceği alanlardan biri olmasıyla yakından ilişkilidir. Tıptaki “hastalık yoktur, hasta vardır” şeklindeki klasik yaklaşım, her ne kadar hastanın tedavisi sırasında göz önünde bulundurulması gereken önemli bir ilke olarak ifade edilse de mevcut tıp uygulamalarında bu ilkenin sistematik olarak ihmal edildiğini ve uygulanmasında zorluklar olduğunu görüyoruz. Tek bir kanser hastasına özel olarak tedavi protokolü belirlenmesi, tıbbi literatürün, son tedavi rehberlerinin, hastanın tüm sağlık kayıtlarının, hastanın bireysel yaşam tarzının ve koşullarının eş zamanlı analiz edilmesini gerekli kılmaktadır. Literatürde günde 122 adet onkolojiyle ilgili bilimsel makalenin girildiği günümüzde ise bu yaklaşımın uygulanması devasa bir analiz yükü doğurmakta, hastanın uygun tedavi seçeneğiyle eşleştirilmesini, dolayısıyla tamamen hastaya özel tedavi uygulanmasını neredeyse imkânsız hale getirmektedir. Bu sebeple pratikteki uygulamada hasta, yapılan tetkikler ve hekim tecrübesi ile belirli sayıdaki tedavi protokolünden birine

göre gruplanmakta ve o protokol uygulanmaktadır. Ayrıca tıbbi uygulamadaki aşırı branşlaşma da hastanın bütüncül olarak ele alınmasını zorlaştırmaktadır.

Büyük veri analizinin gerçek zamanlı olarak yapılabilmesine imkân sağlayan teknolojik gelişmeler “yapay zekâ” (*artificial intelligence*) kavramını yeniden gündeme taşımıştır. Yeni bir kavram olmayan yapay zekânın günümüze kadar gündemde olmasının iki büyük sebebi, eskiden büyük verinin olmayışı ve bu veriyi işleyecek teknolojik donanımın mevcut olmayışıdır. Büyük veri analiziyle geliştirilen algoritmalar, depolama kapasitesindeki artış, işlemcilerdeki kapasite artışı, modüler işlem birimlerinin oluşturulması gibi yenilikler büyük veri analizlerinin gerçek zamanlı yapılmasını mümkün kılmıştır. Teknik donanımın yeterliliği, kurumsal zekâ / iş zekâsı (*business intelligence*) uygulamalarının sağlık sektöründe de uygulanabilmesine imkân tanımıştır. Özellikle pazarlama sektöründe kullanılan segmentleme yaklaşımının hastalıklara ve hasta gruplarına, ilaç endikasyonlarına, tedavi protokollerine uygulanması ile günümüzde tek bir hastaya özel olarak kişiselleştirilmiş tedavi protokolleri mümkün hale gelmiştir. Hatta henüz kişi hasta olmadan önce tüm sağlık kayıtlarının ve literatürün birlikte ve gerçek zamanlı analizi, kişiye özel sağlık belirleyicilerinin tespit edilmesi hastalıkları tahmin edebilmeyi mümkün kılmıştır. Bu uygulamalardan birkaçı aşağıda örnek olarak verilmiştir.

IBM şirketinin geliştirdiği Watson, bir yapay zekâ uygulaması olarak sağlık sektöründe önemli gelişmelere kapı aralamaktadır. Watson’ın ilk ticari uygulaması ABD’de bir kanser merkezi ve sigorta şirketinin ortaklığı ile sağlık sektöründe gerçekleşmiştir. Hindistan’daki bir uygulamada Watson, kanser kurulunun kararları ile %90 uyum içinde doğru teşhis koymuş ve tedavi protokolü oluşturabilmiştir. Watson, kanser hastaları için literatürü okuyabilmekte, güncel rehberleri ve hastanın sağlık kayıtlarını da birlikte analiz edebilmekte, kansere özgü DNA dizileme verilerini de hesaba katmakta ve sadece o hastaya özgü tedavi protokolü oluşturabilmektedir. Tedaviden fayda görmeyen ve uygulanacak ek

bir protokol olmayan kanser hastalarının deneysel aşamadaki binlerce çalışmadan en uygunu ile eşleştirilmesini sağlamaktadır. Yeni ilaç modellerinin geliştirilmesi ve eski ilaçların endikasyon gruplarının yeniden organizasyonunu sağlamaktadır. Watson sadece hastalar ve hastalıklarla değil, sağlık kurumlarının optimizasyonu, maliyet analizlerinin makine tarafından yapılabilmesi, tüm sağlık kayıtlarının tek bir platformda bütüncül olarak analiz edilmesi hizmetlerini de sunmaktadır. Ayrıca Watson bilimsel çalışmalara da yeni bir perspektif kazandırmakta, bilimsel çalışmaların dijital bir zeminde projelendirilmesi ve süreç takiplerinin tüm araştırmacılar için kolaylıkla yapılabilmesini sağlamaktadır. Sunduğu bu yenilikler sebebiyle Watson ile ilgili dört farklı literatürde (PubMed, Scopus, Web of Science, ScienceDirect) 2018 Şubat ayında 74 makale mevcuttur ve bu sayı giderek artmaktadır.²¹

Büyük veri analizi koruyucu sağlık hizmetlerine de yeni bir perspektif sunmaktadır. Yapılan bir çalışmada gebelikte erken doğuma sebep olan belirleyiciler veri madenciliği algoritmaları ile tespit edilmiş ve erken doğumlu gebeliklerde risk faktörleri tanımlanmıştır. Bu risk faktörleri ile yüksek riskli gebelerin tespit edilmesi mümkün olmuştur.²²

Büyük şirketler, artık verilerinin bir kısmını da kamuya sunmaktadırlar. Örneğin, Google şirketinin sunduğu Google Trends uygulaması ile toplumun grip salgınları konusundaki aramaları izlenebilmekte, ülke içinde hangi bölgelerin “grip” kelimesini daha çok aradığı ve aramaların zamana göre dağılımı anlık olarak izlenebilmektedir.

Veri analizindeki teknik gelişmeler büyük kolaylıklar doğurmuştur. Ancak özellikle sağlık konusunda verilerin kısıtlı olması, yasal ve etik sebeplerle oluşturulan veri gizliliği protokolleri yeni sivil oluşumların doğmasına sebep olmuştur. Örneğin, Wellcome Trust adlı kuruluşun UK biobank uygulaması 500.000 gönüllü insanın kişisel verilerini bir havuzda toplamakta ve araştırmacıları bu veri havuzuna davet etmektedir. Benzer misyonla kurulmuş olan European Genome-phenome Archive (EGA), ADNI, SageBionetwork gibi girişimler de diğer veri toplayan ve sunan programlardır. Bu programlar

kişilerden aldıkları verilerin karşılığında ödemeyi onlara kişisel verilerinin analiz edilmiş şeklini sunmak şeklinde yapmaktadır. Başka bir iş kolu olarak kurulan şirketler, farklı şirketlerin veri tabanlarındaki kişisel verilerin ortak havuzlarda toplanması ve paydaş şirketlere analizlerin sunulması modellerini kurgulamaktadır.

Dünyada hibeler artık sadece para olarak sunulmamakta, veriler de hibe olarak araştırmacılara bağışlanabilmektedir. Twitter, 2014 yılında Twitter Veri Hibeleri adlı projesiyle verilerini akademik araştırmalara açabileceğini duyurdu ve 6 proje bu hibe desteğini almaya hak kazandı. Bunlardan 4 tanesinin sağlık temalı projeler olması dikkat çekicidir.²³

İş sağlığı alanında bir şirket, kurumsal firmalara 20 yaş üzerindeki çalışanlarının sağlık kayıtlarını analiz ederek sağlık profillerini sunmaktadır.²⁴ Böylece bir işveren, işçilerinden hangisinin daha önce kalp krizi geçirebileceğini, hangisinin mesleki hastalık sınırında olduğunu tahmin edebilmektedir. Bu çok değerli bilginin işçinin menfaatlerine mi yoksa işverenin menfaatlerine mi kullanılacağı ise büyük bir etik ikilem olarak karşımızda durmaktadır. İş sağlığı ve güvenliği alanında önemli bir kavram olan “sağlıklı işçi etkisi”, bir işverenin hasta olma ihtimali olan işçileri önceden tahmin edip onları işten çıkarması ve yeni işçiler alması sayesinde görünürde tüm işçilerinin sağlıklı olmasını da kapsamaktadır. Büyük veri analizleri ile geleceği tahmin etmenin sağlıklı işçi etkisine nasıl bir boyut kazandıracığı ise nitelikli bir tartışma konusudur.

Sağlık sektöründeki büyük veri analizleri ve yapay zeka uygulamaları, “Hassas Tıp (*Precision Medicine*)” kavramının da doğmasına yol açmıştır. Bu kavram, halk sağlığı uygulamalarında toplumsal düzlemde ele alınan “sağlığın sosyal belirleyicileri” modelinin bireysel düzlemde uygulanmasını içermektedir. Bir kişinin gen dizilimi, çevresel etmenlere ait verileri ve yaşam tarzı verilerinin birlikte analiz edilmesi ile kişiye özel koruyucu ve tedavi edici sağlık hizmetinin mümkün olduğu bir yaklaşımı ifade etmektedir. Özellikle tedaviye cevapsız hastalıkları önleme, risk gruplarını belirleme

ve hastalıklar ortaya çıkmadan müdahale etmeyi kapsamaktadır. Bu alanda ABD'de Ulusal Sağlık Enstitüsü (NIH) bünyesinde önemli çalışmalar yürütülmektedir.²⁵

Büyük veri analizleri, sağlık alanında bir başka büyük soruna çözüm sunma potansiyeline sahiptir. Sağlık hizmetlerinde tıbbi müdahalelerde hata yapma (malpraktis), çıktıları açısından önemli bir yer teşkil etmektedir. ABD'de 2016 yılında yapılan bir çalışma tıbbi hatanın ABD'deki ölüm sebepleri arasında 3. sırada olduğunu ifade etmektedir.²⁶ Bu bilgi sağlık sektörü için sansasyon düzeyindedir. Tıbbi hatalar, sağlık kayıtlarının bütüncül bir yaklaşımla toplanması, analiz edilmesi ve olası ilaç yan etkilerinin ve alerjik reaksiyonların tüm hastalar için daha ortaya çıkmadan öngörülmesi ve müdahale edilmesi, hekimler ve branşlar arasındaki iletişim kopmalarına dijital ortamda çözüm sunulması büyük veri analizlerinin sunduğu yeniliklerdendir.

Epistemolojik Boyut

Bilim tarihinde verilerin toplanması, analiz edilmesi ve yorumlanarak bilgi haline getirilmesi yeni bir faaliyet değildir. Günümüzle karşılaştırılabilecek türden bu bağlamdaki faaliyetleri bilim tarihi içerisinde yeni çağa kadar götürmemiz mümkündür. Fakat bu alandaki çalışmalar günümüzde büyük veri kavramı çerçevesinde yukarıda kısaca açıklanan hacim, hız, çeşitlilik ve benzeri diğer parametreler göz önüne alındığında çok farklı bir boyuta erişmiştir. İşte gelinen bu nokta sadece bilginin oluşumu bağlamında epistemolojik açıdan değil, aynı zamanda etik açıdan da yeni soru ve sorunların ortaya çıkmasına yol açmıştır. Bu verilerin farklı alanlarda değişik amaçlarla kullanılması ve saklanması mahremiyet ve mağduriyet bağlamında değişik ve etik boyutu olan sorunları ortaya çıkarmıştır.²⁷

Büyük Veri dünyasında sağlık alanını ilgilendiren bilginin ortaya çıkması her ne kadar bilinen geleneksel araştırmalarla benzerlik gösterse de makro ve mikro düzeyde ciddi farklılıklar barındırmaktadır. Sağlık alanında belli bir konuda araştırma yapılacağı zaman veri toplanacak alan araştırma sorusuyla belirlenir ve bir hipotezle netleştirilir. Eğer bu araştırma insanlarla gerçekleştirilecekse o araş-

tırmanın gerçekleşmesini sağlayacak bireyler sürece belli prosedürlerle müdahil olurlar. Yine bu araştırmalar çerçevesinde verilerin toplanacağı alanlar daha önce belirlenmiş hastaneler, okullar, muayenehaneler gibi tespit edilmiş mekanlardır.

Sağlık alanında büyük verinin dünyasında bilgi üretmek söz konusu olduğunda ise bilginin oluşumunda ne zaman, ne de mekan olarak herhangi bir kısıtlama mevcuttur. Yine verinin bilgiye dönüşümünde klasik araştırmalarda olduğu gibi birbirini izleyen sabit bir süreçten bahsetmek mümkün değildir. Gündelik hayatımızda belli ihtiyaçlar için kullanılan cep telefonu ve bu telefondaki özellikle sağlıkla ilgili uygulamalar, bilgisayarla internet kullanımını ve belli konularda bilgilerin aranması sürekli olarak ileride sağlık alanında kullanılabilecek ve farklı çıkarımlara götürecek veriler üretmektedir. Hatta dijital ortamın imkanlarıyla kullanılan kredi kartı da aynı şekilde kişinin harcamaları hakkında bilgi vermekte ve dolayısıyla direkt ya da dolaylı olarak o kişinin sağlığıyla ilgili veriler üretmektedir. Diğer taraftan özellikle son yıllarda ingilizcede 'wearables' denilen kişinin sağlığıyla ilgili bilgi veren taşınabilir küçük aletler giderek yaygın bir şekilde kullanılmakta ve bunlar da yine yoğun bir şekilde veri üretmektedir. Hatta şu anda hızla gelişmekte olan 'smart home' (akıllı ev) konseptiyle evin içinde bulunan elektronik aletler o kişi evde olmadan da veriler üretmekte ve bu verilerin sağlıkla ilişkisi kurulabilmektedir. Yine günlük sosyal medyadaki paylaşımlar da kişinin sağlığı hakkında farklı değerler içeren veriler oluşturmaktadır.²⁷

Dijital teknolojinin yaygın kullanılmasıyla yukarıda bahsedilen alanların dışında özellikle sağlık alanında da yeni uygulamalarla birlikte tamamen bireyin sağlığına yönelik veriler üretilmektedir. Bunlar arasında tıbbi test sonuçları ve bu sonuçların ilgili veritabanlarına girilmesi, elektronik hasta dosyalarına yapılan diğer kayıtlar, hastanede toplanan diğer veriler, bakanlık düzeyinde yürütülen istatistikler, yeni aletlerle hastanın yakından ya da uzaktan takip verileri gibi alanları saymak mümkündür.²⁸

İşte bahsedilen bu alanlarda verilerin toplanıp bilgiye dönüşmesi tıp alanındaki geleneksel bilgi-

nin üretilmesiyle karşılaştırıldığında ikisinin de tümevarım metoduyla çalıştığı söylenebilir. Fakat geleneksel araştırmalar daha çok kausalite yani nedensellik üzerinde yoğunlaşmakta ve direk olarak nedene yönelik bir takım çıkarımlar yapıp onları sağlık alanında kullanmaktadır. Büyük veri bilgi üretimi ise verilerin oluşması çerçevesinde ve bu verilerin belli algoritmalar ve analizlerle daha çok belli olaylar arasında korelasyonu yakalamaya odaklıdır. A ürününü sıklıkla satın alan ve tüketen veya X davranışında sıklıkla bulunan insanların B hastalığına daha çok yakalanmaları gibi. Konuya epistemolojik olarak yaklaşıldığında A maddesini tüketmek B hastalığına yol açar şeklinde nedensel bir açıklamaya gitmek mümkün değildir. Böylece tespitler açıklanan nedenlerden dolayı daha çok korelasyon seviyesinde kaldığından bu iki fenomen arasındaki ilişkinin nedensel açıdan da bir bağının olup olmadığını tespit etmek daha çok bunun üzerine yapılacak hedefli çalışmalarla belli olacaktır.²⁹ Bu bağlamda medyamızda 'A maddesini tüketenlerin B hastalığına yakalandığını yada B hastalığından korunduğunu araştırmacılar buldu' ya da 'A maddesini tüketmenin B hastalığından korunmayı sağladığı bulundu' türünden haberlerin epistemolojik olarak ne kadar yanlış olduğu da ortadadır.

Diğer taraftan büyük veri alanındaki verilerin muazzam bir şekilde büyük olması, bir taraftan daha kesin bilgiye ulaşma ümidini beslerken, diğer taraftan da burada kullanılan algoritmalar ve teoriler bilginin şeklinin ve içeriğinin belirlenmesinde daha önemli rol oynamaktadır. Devasa verilerle yapılan analizlerin hastalığın etyolojisi hakkında daha kesin bilgiye, geliştirilecek tedavilerde daha bireysel konseptlerin gelişmesine ve hastalığın prognozunun daha kesin olmasına yol açması beklenmektedir. Analizin doğru ve bu yolla oluşturulan bilginin gerçek olması aynı zamanda bilgiyi oluşturacak olan verilerin de doğru ve gerçek olmasını gerekli kılmaktadır. Gerçek bilgiye erişim konusunda veriler ne kadar geniş bir ağ aracılığıyla toplanırsa toplanırsa bilgiye dönüşümün kalitesini belirleyen analiz, verinin gerçek veri olduğu varsayımına dayanmaktadır. Mesela insanın kullanımı için tasarlanan bir

küçük alet eğer ev hayvanına takılırsa sağlıklı veri üretilmeyecektir. Bu da analizde kullanılan verinin her zaman 'temiz veri' olmadığı anlamına gelmektedir. Onun için büyük veri özelliklerinden biri olan "verinin doğruluğu" bu bağlamda büyük bir sorun olarak karşımızda durmaktadır. Özellikle yapılandırılmamış verilerin analizinde mevcut algoritmalar belirli hata düzeylerinde işlem yapabilmektedir. Şu anda ve ileriki zamanlarda bahsedilen bu sorunlardan dolayı hatasız ve mükemmel bir analize ulaşmanın güçlüğü ortadadır.³⁰

Yukarıdaki farklılıkların yanında büyük veri bağlamında verilerin toplanması ve analiz edilmesi çok geniş bir ağ ve teknik alt yapı gerektirdiğinden bu alanda bilgi üretecek kurumların da bu ağa ve teknik imkanlara sahip olması gerkeir. Bu da günümüzde sadece internet dünyasında Amazon, Facebook, Google, Twitter gibi devasa şirketler ve karteller için mümkündür. Eğer bir kurum ya da şirket bu alanda bilgi üretmeyi planlıyorsa bu büyük şirketlerle birlikte çalışmak zorunda kalacaktır. Bu da bilim dünyasında yeni bir tekelleşmeyi beraberinde getirecektir. Artık bir tıp fakültesinde küçük bir grupla araştırma yaparak bilgi üreten 'butik araştırmalar' eski önemini yitirecek ve tarihe karışacaktır. Bu da ilgili bilimsel kurumların güçlü bilgi oluşturması için muazzam kaynaklara sahip olmasını gerekli kılacaktır. Yani 'bol para' yoksa 'iyi bilim' de yok olacaktır.

Etik Boyut

Yukarıda bahsedilen büyük veri bağlamında bilginin oluşumundaki temel değişiklikler ve sağlık alanındaki araştırma şekillerinin farklılaşması bir taraftan yeni ve daha iyi tedavi imkanlarının doğmasına yol açarken, bir taraftan da etik açıdan mevcut sorunların kompleksleşmesine ve aynı zamanda yeni soru ve sorunların ortaya çıkmasına yol açmaktadır. Yeni oluşan bu durumlarda farklı paydaşlar değişik karakterde olan ilişkiler içerisindedirler. Bütün bu gelişimlerin içerisinde direk olarak etkilenen hasta, büyük veri imkanlarının teşhis, tedavi veya hastalıkları önleme amacıyla kullanıldığı sağlık sistemi, araştırmaların yapıldığı ve verilerin bilgiye dönüştürüldüğü bilim dünyası ve verileri toplayan

teknik alt yapıyı sağlayan ve yöneten firmalar olmak üzere birçok paydaş bulunmaktadır.³¹ Bu paydaşlar birbiriyle olan ilişkilerinden dolayı birinin menfaatine olan bir durum diğerinin aleyhine olabilmektedir. Örneğin hastaya ait bazı bilgiler işverenin veya özel yaşam sigortalarının işine yarayabileceği gibi bu bilgilerin ait olduğu kişi bu durumdan zarar görebilir.

Büyük verinin sağlık alanında kullanılmasıyla ortaya çıkan etik açıdan anlamlı olan sorun alanlarından belki de en önemlisi kişinin sağlığıyla ilgili verilerinin dijital ortama aktarılmasından sonra bunların korunması ve kişinin tasarrufu ve rızası dahilinde kullanılmasıdır.³² Bu durum sağlık hizmetlerinden istifade eden herkes için geçerli olduğuna göre büyük veri kullanımı ile ilgili yeterli bilgi düzeyi ve hassasiyette olmayan insanlar için ciddi bir sorun teşkil etmektedir. Dolayısıyla bu alanda ilk karşımıza çıkan problem her yetişkin insanın bu konuda gerektiği kadar ve doğru bir şekilde bilgilendirilmesidir.³³ Mesela okuma yazma bilmeyen yaşlı bir Anadolu kadını bu konularda nasıl bilgilendirilecektir? Bu bilgilendirmenin sağlık hizmetleri sırasında hastanın rızasının bilgilendirme olmadan hukuki olarak geçersiz olduğu ve etik açıdan sorunlu olduğu göz önüne alınırsa ne kadar önemli olduğu anlaşılır. Bu bilgilendirme elbette sadece teknik ve formalite düzeyine indirgenmemelidir. Burada bilgilerin nerede ve nasıl kullanılacağı konusunda da açıklamalar yapmak gerekmektedir.

Bütün bu karmaşık prosedürler göz önüne alındığında sağlık hizmetlerinde çalışan her hekimin bu konuda bilgi sahibi olması, karşısındaki hastaya bunları anlatması, kendisine tanınan süre içerisinde neredeyse imkansızdır. Bu görevi başka makam ve kuruluşlara tevdi etmekte yine kendi içerisinde başka sorunlar içermektedir. Bunların başında teşhis ve tedavi ile ilgili konuların hekim olmayan kişiler yada kurumlar tarafından üstlenilmesi gelmektedir.

Büyük verinin sağlık alanında yaygın olarak kullanılmasında diğer bir sorun bu verilerin ve bilgilerin hasta veya kişi mahremiyeti bağlamında korunmasıdır.³⁴ Bilindiği gibi sağlık hizmetleri sırasında hekimin hastası ile olan ilişkisi özel olma duru-

mundadır ve dolayısıyla hem hukuki olarak hemde etik açıdan hekim sır saklama mükellefiyetindedir. Ancak büyük veri geniş alanlarda kullanıldığında bu mükellefiyetin gerçekleştirilmesi zorlaşmaktadır. Bu konudaki klasik problemlerden birisi kişinin sağlıkla ilgili bilgilerine başkaları tarafından ulaşılarak kişinin aleyhinde bir takım sonuçları doğurmasıdır. Mesela bu bilgilerin özel sağlık sigortaları ya da yaşam sigortaları tarafından bilinmesi ödeyeceği sigorta primlerinin yükselmesine yol açabilir. Bu bilgilerin izinsiz ve kontrolsüz bir şekilde internet ortamında, mesela sosyal medyada hızla yayılması yine o kişi için bilginin türüne göre ciddi ayrımcılığa, damgalanmaya ya da dışlanmaya yol açabilir. Bunlar için klasik örnek cinsel yollarla bulaşan hastalıkların ya da özel muayene ve test sonuçlarının sanal ortamda bulunması ve izinsiz kullanılmasıdır. Bu duruma vahim bir örnek olarak bir Alman kadın doğum uzmanının jinekolojik muayeneleri sırasında hastalarından izinsiz olarak 2008 yılından 2011 yılına kadar çekmiş olduğu 10.000'in üzerinde mahrem fotoğraf örnek olarak verilebilir. Bu durumun tespit edilmesinden sonraki mahkeme sürecinde doktor üç buçuk yıl hapis cezası almıştır. Her ne kadar resimler internet ortamında paylaşılmamış olsa dahi bu fotoğrafların dijital olması onları her an internette paylaşılabilir hale getirmektedir.³⁵ Burada bu ve benzeri davranışların hukuka aykırı olması sorunu basitleştirmemekte, tam tersine çok basit teknik imkanlarla gerçekleştirilebilir olması yasak olmasına rağmen bu konular üzerinde daha fazla durmamızı gerektirmektedir. Diğer taraftan hasta mahremiyetinin mevcut teknik imkanlarla bu derece kolay zedelenebilir olması yine hasta hekim ilişkileri açısından temel ilkelere olan güven konusunu da yeniden etraflıca ele almamızı gerektirmektedir.

Burada büyük veri ortamında klasik hasta hekim ilişkilerindeki mahremiyet kavramının da içeriğinin ve sınırlarının değişeceğinin altını çizmek gerekir. Önceden hastanın doktoru ile dört duvar arasında konuştukları ve muayene, teşhis ve tedavi sırasında söylenenler ve paylaşılanlar mahremiyete dahil olan ve korunması gereken bilgiye girerken,

artık sağlık hizmetleri sırasında bu kişi hakkında bilgisayara girilen her bilgi ve dijital veri mahremiyetle ilişkili hale gelmektedir. Yine bu kişinin iş hayatında ve özel hayatında kullandığı ve ürettiği bütün veriler şahısla ve dolayısıyla sağlığıyla ilişkilendirilebilir hale gelmektedir. O kişi artık sanal dünyada nerede olursa olsun X hastalığı olan bir veri üreticisidir. Bu verileri üreten şahıs artık herhangi bir aleti kullandığında 'büyük veri kosmosunda' pratik ve teknik olarak izlenilebilecek dijital ayak izleri ve parmak izleri bırakmaktadır. Hipokrat yemininde de yer alan hastanın sırlarını saklama mükellefiyeti artık hiçbir hekimin ne kadar isterse istesin bütün gayretlerine rağmen bireysel olarak yerine getirme konusunda garanti veremeyeceği bir durumdur.

Gerek tedavi hizmetlerinde gerekse klinik araştırmalarda verinin anonimleştirilmesi mahremiyetle ilgili sorunlara bir çözüm olarak teklif edilmektedir. Her ne kadar anonimleştirme birtakım faydalar sağlasa da kendi içerisinde paradokslar içermektedir. Bir kişinin verileri kimlik bilgilerinden bağımsız olarak işlendiğinde de o verilerin kime ait olduğunun tespiti büyük verinin sunduğu imkanlarla kolaylaşmıştır. Bu bağlamda kusursuz anonimleştirme uygulanması zor bir yöntem olarak karşımıza çıkmaktadır.

Bahsedilen sorun alanları yanında büyük verinin sağlık alanında kullanılmasının mevcut teknik faydaları yanında elbette başka yararlı olduğu konular olacaktır. Mesela daha fazla bilginin daha hızlı bir zaman içerisinde toplanıp veri tabanını genişleterek daha kesin ve kanıta dayalı bilgi oluşturulabilir. Bu da şu ana kadar sağlık ve tıp alanında bildiklerimizi arttırma veya mevcut yanlış olan bilgileri daha çabuk tashih imkanı sağlar.

Kişiselleştirilmiş Tıp (*individualized medicine*) alanına daha önceki zamanlardan çok daha fazla veri ve bilgi aktararak bu alanın hızlı bir şekilde gelişmesi sağlanabilir. Burada işte bu olumlu gelişmelere bu verilerin üreticisi durumunda olan bireylere belli sorumluluklar düşmektedir. Bu kişiler üretmiş olduğu verileri paylaşarak, bu verilerin belli alanlarda kullanılmasına izin vererek bu konulardaki

gelişmelere katkıda bulunmuş olurlar. Bu bağlamda klasik kan/organ bağışi yerine artık 'dijital veri bağışi'ndan, dijital bağışçılardan bahsetmemiz mümkün olacaktır.

Burada ahlaki olarak tıbbın gelişmesine katkıda bulunma mesuliyeti karşımıza çıkmaktadır. Takdir edilir ki bu alandaki bağışların çok ve cömertçe olması – ki istenen de budur – ancak insanların mahremiyet haklarına azami derecede riayet edilerek onlara bir güven oluşturmakla mümkündür. Diğer yandan bu konulardaki veri akışını yöneten resmi veya özel kurumlar ne kadar gayret ederse etsin bu bilgilerin teknik olarak ve illegal ulaşılabilir olması insanlarda haklı bir şüpheyi de beraberinde getirmektedir. Mesela bu konularda geçtiğimiz zamanlarda yaşanan negatif tecrübeler bu konudaki şüphe ve karamsarlığı haklı çıkarmaktadır. ABD'de bir kredi derecelendirme kuruluşu, 143 milyon ABD'li müşterisinin isim, soy isim, sosyal güvenlik numarası, adres, doğum tarihi, kimlik numarası, ehliyet belgesi verilerinin çalındığını duyurmuş olması bu alanlarda hataların ne kadar kolay gerçekleşebildiğini ve bunun ahlaki boyutunu gözler önüne sermesi açısından anlamlıdır.³⁶ Yine ABD'de bir sağlık şirketi 2014 yılında Çinli bilgisayar korsanlarının 4,5 milyon hasta verisini çaldığını duyurmuştur. Bunun üzerine FBI sağlık şirketlerine veri güvenliğine dair uyarı yayınlanmıştır.³⁷ Tarihin en büyük sağlık kayıtları hırsızlığı ise 2015 yılında yine ABD'de gerçekleşmiştir. Ülkenin ikinci en büyük sağlık sigorta şirketine ait 80 milyon müşterinin sağlık verileri çalınmıştır.³⁸

Bu kötü haberlere rağmen toplumun veri güvenliğine bakış açısını göstermesi açısından Vodafone'nun 8 Avrupa ülkesinde 8.256 kişiyle yaptığı çalışma önemli bilgiler sunmaktadır.³⁹ Çalışmada katılımcıların %53'ü sağlık ilişkili kayıtlarının başka insanlara faydalı olması durumunda anonimleşmeden dahi kullanımına rıza gösterdiği ifade edilmiştir. Yine aynı araştırmadan eğitim seviyesi düşük ve yaşlı insanların veri paylaşımında daha temkinli davrandığını öğreniyoruz.

Bütün bu yaşanan skandallara rağmen bu araştırmaya katılan insanlar kişisel sağlık verilerinin

güvenli olarak sağlık sistemlerinde korunduğunu düşünmektedirler. Duyulan güven düzeyi sıralamasında 1. sırada sağlık sistemleri yer almakta, sonrasında işverenler ve ardından bankaların veritabanları gelmektedir. İnsanlar verilerini paylaşmak konusunda en çok sigorta şirketlerine şüpheyle yaklaşmaktadırlar.

Yukarıda tartışılan büyük verinin giderek sağlık sisteminde daha fazla yer almasıyla ortaya çıkan etik sorunların çözümü için neler yapılmalıdır? Acaba bu yeni sorun alanları insan davranışlarını ve bazı durumları değerlendirirken kullandığımız klasik etik temel ilkelerin pratik hayata geleneksel bir şekilde uygulanması mümkün müdür? Uygulamalı Etik alanındaki bu soruların ötesinde acaba epistemolojik ve felsefi anlamda hangi sorunlar oluşmaktadır? Bu bağlamda bu konularla ilgili insanın özgürlüğü, özerkliği/otonomisi, mahremiyeti, sorumluluğu, adalet ve hakimiyet gibi temel normatif kavramları yeniden nasıl düşünmeliyiz ve yeni problemler bağlamında nasıl yeniden yorumlamalıyız?

İşte bu sorular çerçevesinde acaba hangi yeni normatif kavramların oluşması gerekir veya kadim kavramların nasıl yorumlanması gerekir diye sorulabilir. Alman Etik Konseyi 2017 yılında bu alanda yazılmış en şümüllü raporunda büyük verinin sağlık alanında kullanılmasında normatif bir kavram geliştirmiştir. Bu raporda 'Veri-Hakimiyeti' (Alm. *Datensouveränität*) kavramı büyük verinin getirmiş olduğu imkan ve risklerle sorumlu ve bilgili bir şekilde informatik bireysel özgürlüğün şekillenmesi (*informationelle Freiheitsgestaltung*) olarak tanımlanmakta ve bunu büyük veri dünyasında hem etik açıdan, hem de hukuki açıdan ulaştırılması gereken bir hedef olarak görmektedir. Burada Informatik Bireysel Özgürlük büyük veri dünyasında kişinin verilerine sahip çıkabilmesi ve istediği gibi bu konuda bu verilere kimlerin ulaşabileceğine ve kullanabileceğine karar verebilme hakkını içerir.⁴⁰ Bu hakkın doğru kullanılmasıyla kişi değerler sistemi çerçevesinde mahremiyetini de koruma imkanına kavuşacaktır. Veri Hakimiyeti bu bağlamda sadece bahsedilen hakkı değil, aynı zamanda verilerin ürettiği ve sağlıklıyla ilgili yararlı bilgilerin üretildiği

alandaki kişinin sorumluluk ve görevlerini de kapsamaktadır. Yani üretmiş olduğu veriler gerçekten toplumun yararına kullanılması durumunda kişi bu verileri bilimin istifadesine sunma yükümlülüğünü de hissetmesi gerekir.^{41,42,43}

SONUÇ

Makalemizde büyük verinin sağlık alanında kullanılması sırasındaki temel özelliklerinin yanısıra bu alandaki şu anda ve ileride muhtemel kullanma şekillerine değindik. Yine bu uygulamalar sırasında farklı karakterlerde ortaya çıkacak olan epistemolojik ve etik alandaki sorunları tartıştık. Konuya genel olarak yaklaştığımızda bir taraftan yeni ve insanlık tarihinde görülmemiş imkanların ortaya çıktığını görüyoruz. Diğer taraftan ise yine sağlık alanında kısmen yeni, kısmen ise daha önceden mevcut olan epistemolojik ve etik sorunların daha karmaşıklaşmış şekliyle karşılaşıyoruz. Bu süreç içerisinde bilgi oluşum şekillerinin ve oluşan bilgilerin sağlık alanında kullanılma tarzlarının etik bağlamda temel normatif kavramları da etkilediğinin altını çizmek durumundayız. Bu temel kavramların başında insan özgürlüğü, insanın özerkliği/otonomisi, araştırma yapma özgürlüğü, hasta mahremiyeti ve sorumluluğu gelmektedir.

Yukarıda kısmen tarif ettiğimiz kadarıyla artık büyük verinin sağlık alanında kullanılıp kullanılmamasından ziyade "nasıl, hangi şartlarda ve hangi ahlaki çizgiler içerisinde kalarak kullanmalıyız" sorusunu sormak durumundayız. Bu soruların cevabı ise hem mikro düzeyde büyük verinin sağlık alanında uygulamalarını etik açıdan analiz etmek, hem de makro düzeyde bütün bu uygulamaların bizim insan anlayışı, insan özgürlüğü ve insan sorumluluğu kavramlarımız açısından ne anlama geldiğini sorgulamakla mümkündür.

Hakem Değerlendirmesi: Dış bağımsız.

Yazar Katkıları: Çalışma Konsepti/Tasarım- A.U, İ.İ.; Veri Toplama- A.U, İ.İ.; Veri Analizi/Yorumlama- A.U, İ.İ.; Yazı Taslağı- A.U, İ.İ.; İçeriğin Eleştirel İncelemesi- A.U, İ.İ.; Son Onay ve Sorumluluk- A.U, İ.İ.; Malzeme ve Teknik Destek- A.U, İ.İ.; Süpervizyon- A.U, İ.İ.

Çıkar Çatışması: Yazarlar çıkar çatışması beyan etmemişlerdir.

Finansal Destek: Yazarlar finansal destek beyan etmemişlerdir.

Peer Review: Externally peer-reviewed.

Author Contributions: Conception/Design of Study- A.U, İ.İ.; Data Acquisition- A.U, İ.İ.; Data Analysis/Interpretation- A.U, İ.İ.; Drafting Manuscript- A.U, İ.İ.; Critical Revision of Manuscript- A.U, İ.İ.; Final Approval and Accountability- A.U, İ.İ.; Technical or Material Support- A.U, İ.İ.; Supervision- A.U, İ.İ.

Conflict of Interest: Authors declared no conflict of interest.

Financial Disclosure: Authors declared no financial support.

KAYNAKLAR

- Cox M., Ellsworth D. (1997); Application Controlled Demand Paging for Out-of-Core Visualization; Proceedings. Visualization '97 (Cat. No. 97CB36155).
- Mashey R. (1998); Big Data and the Next Wave of InfraStress Problems, Solutions, Opportunities; USENIX Annual Technical Conference, <http://bit.ly/2Y0nwPZ> (30.04.2019).
- Schwardmann U. (1993); Parallelization of a Multigrid Solver On The Ksr1; J-Supercomputer 10(3) S.4-12.
- World Economic Forum (2012); Big Data, Big Impact: New Possibilities for International Development, <http://bit.ly/2vA6qMl> (30.04.2019).
- Gantz J.; Reinsel D. (2010); The Digital Universe Decade – Are You Ready?; IDC, <http://bit.ly/2VzWj9b> (30.04.2019).
- Rotella P. (2012); Is Data The New Oil? , Forbes, <http://bit.ly/2V6Z6au> (30.04.2019).
- Merriam Webster Dictionary, “Datum” maddesi, <https://www.merriam-webster.com/dictionary/datum> (30.04.2019).
- Numanoğlu N., Eynehan M.E. (2016); Türkiye'nin küresel rekabetçiliği için bir gereklilik olarak sanayi 4.0; Tüsiad, İstanbul.
- Bayrakçı S. (2015); Sosyal Bilimlerdeki Akademik Çalışmalarda Büyük Veri Kullanımı; Yüksek Lisans Tezi; Marmara Üniversitesi Sosyal Bilimler Enstitüsü, İstanbul.
- Mauro A. , Greco M., Grimaldi M. (2014): What is Big Data? A Consensual Definition and a Review of Key Research Topics; 4th International Conference on Integrated Information, Madrid.
- Ffoulkes P. (2017): inside Bigdata Guide to Use of Big Data on an Industrial Scale Report; insideBIGDATA.
- Jacobson R. (2013): 2.5 quintillion bytes of data created every day. How does CPG & Retail manage it?, IBM Inc., <https://ibm.co/2vqNHTF> (30.04.2019).
- Reinsel D, Gantz J, Rydning J; Data Age 2025 The Evolution of Data to Life-Critical; IDC White Paper; 2017.
- Omnicores Agency (2019): Facebook by the Numbers: Stats, Demographics & Fun Facts, <https://www.omnicoreagency.com/facebook-statistics/> (30.04.2019).
- Smith C. (2013): Facebook Users Are Uploading 350 Million New Photos Each Day; Business Insider, <http://www.businessinsider.com/facebook-350-million-photos-each-day-2013-9?IR=T> Erişim Tarihi: (30.04.2019).
- Soderbery R. (2013): How Many Things Are Currently Connected To The “Internet of Things” (IoT)?, Forbes, <http://bit.ly/2GPFd91> (30.04.2019).
- Kitchin R.(2016): The ethics of smart cities and urban science. Philos Trans A Math Phys Eng Sci. 28;374(2083).
- IBM Inc. (2019): What is MapReduce?, <https://www.ibm.com/analytics/hadoop/mapreduce> (30.04.2019).
- D'Avolio L. (2016); Hype and Disappointment on the Road to Healthcare's Promised Land; BigData and Healthcare Analytics Forum, <http://bit.ly/2WenjRS> (30.04.2019).
- Schönberger M., Cukier K. (2013); Büyük Veri - Yaşama, Çalışma ve Düşünme Şeklimizi Dönüştürecek Bir Devrim; Çev. Banu Erol; Paloma; İstanbul.
- Uçar A. (2018); IBM Watson ile Sağlığa Arttırılmış bir Bakış; SD Platform; 46: 6-11.
- Chen H.Y., Chuang C.H., Yang Y.J., Wu T.P. (2011); Exploring the risk factors of preterm birth using data mining; Expert Systems with Applications 38: 5384–5387.

23. Twitter Blog (2014): Twitter #DataGrants selections, https://blog.twitter.com/engineering/en_us/a/2014/twitter-datagrants-selections.html (30.04.2019).
24. Vitality Inc. (2019): Britain's Healthiest Workplace (BHW) Privacy Policy, <https://www.vitality.co.uk/business/healthiest-workplace/faqs/> (30.04.2019).
25. NIH: About the 'All of Us' Research Program, <https://allofus.nih.gov/about/about-all-us-research-program> (30.04.2019).
26. Makary A., Daniel M. (2016): Medical error—the third leading cause of death in the US; *BMJ* 3;353:i2139.
27. Deutscher Ethikrat (2017): Big Data und Gesundheit – Datensouveränität als informationelle Freiheitsgestaltung, Berlin.
28. Karakaş M. (2016): Büyük Veri, Endüstriyel İnternet ve Sağlık Alanındaki Uygulamaları, İstanbul: BETİM.
29. Lipworth W., Mason P.H., Kerridge I., Ioannidis J.P.A. (2017): Ethics and Epistemology in Big Data Research. *J Bioeth Inq.* 14(4):489-500.
30. Lipworth W., Mason P.H., Kerridge I. (2017): Ethics and Epistemology of Big Data. *J Bioeth Inq.* 14(4):485-488.
31. Hand D.J. (2018): Aspects of Data Ethics in a Changing World: Where Are We Now? ; *Big Data;* 6(3):176-190.
32. Weichert T. (2014): Big Data, Gesundheit und der Datenschutz; *Datenschutz und Datensicherheit,* 38 (12), 831-838.
33. Fisher C.B., Layman D.M. (2018): Genomics, Big Data, and Broad Consent: a New Ethics Frontier for Prevention Science; *Prev Sci.* 19(7):871-879.
34. Polonetsky J.; Tene O. (2013): Privacy and Big data: Making Ends Meet; *Stanford Law Review,* 66(25).
35. Spiegel Online (2013), Frauenarzt muss ins Gefängnis, <http://bit.ly/2ULaDHj> (30.04.2019).
36. Equifax Inc. (2017); Equifax Announces Cybersecurity Incident Involving Consumer Information, <https://investor.equifax.com/news-and-events/news/2017/09-07-2017-213000628> (30.04.2019).
37. Humer C., Finkle J. (2014): Your medical record is worth more to hackers than your credit card, *Reuters News,* <https://reut.rs/2J3Z9g0> (30.04.2019).
38. McNeal G.S. (2015): Health Insurer Anthem Struck By Massive Data Breach, *Forbes,* <http://bit.ly/2VaKTtf> (30.04.2019).
39. Vodafone Institute for Society and Communication (2016); Big Data, A European Survey on the Opportunities and Risks of Data Analytics, <http://bit.ly/2W9QRa1> (30.04.2019).
40. Buchner, B. (2006): Informationelle Selbstbestimmung im Privatrecht. ; Tübingen.
41. Chadwick R., Berg K. (2001): Solidarity and equity: new ethical frameworks for genetic databases. *Nat Rev Genet.,* 2(4):318-321.
42. Townend D. (2018): Conclusion: harmonisation in genomic and health data sharing for research: an impossible dream? *Hum Genet.* 137(8):657-664.
43. Prainsack B.; Buyx A (2017): Solidarity in Biomedicine and Beyond. Cambridge.