

Büyük Veri, Epistemoloji ve Etik Tartışmalar

Yrd. Doç. Dr. Şevki Işıklı, Marmara Üniversitesi İletişim Fakültesi, sevki.isikli@marmara.edu.tr

ÖZET: *Bilgi ekonomisinin petrolü olarak tanımlanan büyük verinin özellikle beşeri bilimlerde büyük bir dönüşüm başlattığı görüşü, son 10 yılın sık yinelenen bir söylemi haline gelmiştir. Bu dönüşüm, teknolojik yeteneklerimizin artmasıyla da ilişkilendirilir fakat buradaki asıl önemli etki, veriyi nasıl kullanacağımızı söyleyen düşünme biçimlerimizde gerçekleşmektedir. Veri tabanlı olan hesaplamalı düşünme biçimi; beşeri bilimleri, fen bilimleri ile desteklenen fizikalizmin baskısından kurtarmaya başlamıştır. Beşeri bilimler, son iki yüzyıldır “bilim” olduklarını ispatlamak için tarihselcilik, hermeneutik, fenomenoloji gibi arayışlarla yeni yöntembilimsel görüşler geliştirdiler fakat yine de büyük veri çağına gelinceye, hesaplamalı düşünme yöntemini kullanmaya başlayıncaya kadar doğa bilimcilerin hakimiyetinde bulunan “bilim camiası” tarafından sosyal ve beşeri araştırmaların bilim olarak kabul edildiklerini söylemek mümkün olmamıştır. Büyük veri, beşeri bilimlere, “bilim” olmanın temel niteliklerinden addedilen “sayısal” olmayı, bu yolla olayları önceden tahmin etme becerisini “hediyeye” etmiştir. Bu yüzden büyük veri, teknik bir yaklaşımla yalnızca Google uzmanları, istatistikçiler ve ekonomistler için değil, felsefi bir yaklaşımla genel beşeri bilimlerin yeni türden anlaşılması için yeniden ele alınmayı hak etmektedir.*

Anahtar Sözcükler: *Büyük veri, hesaplamalı düşünme, beşeri (sosyal) bilimler, bilim felsefesi, etik.*

Big Data, Epistemology and Ethics Issues

ABSTRACT: *Big data is defined as the oil of knowledge economy, especially in the beginning of a transformation in the social sciences that opinion, has become a recurring rhetoric often last 10 years. This turn is also associated with the increase in our technological capabilities, but here is important, the data is how we will use is taking place in our way of thinking said. By data-driven computational thinking, humanities, began to recover from the pressure of physicalism which supported by science. Humanities, in the last two centuries is to prove they are “science” developed historicism, hermeneutics, and phenomenology have a new methodological opinions with epistemological quest. Nevertheless, until the beginning of the era of big data using computational methods of thinking in the dominance of a naturalist “scientific community” is difficult to say by what they accepted. To be a science, the gift given by the big data to the humanities which are considered fundamental characteristic of being “formulative”, in this way has the ability to predict events. So big data, only a technical approach with Google experts, statisticians, and economists not, should be reconsidered in order to understand the new type of general humanities with a philosophical approach.*

Keywords: *Big data, computational thinking, humanities (social sciences), philosophy of science, ethics.*

1. GİRİŞ

“Kişi, kesinliklerle işe başlarsa varacağı yer kuşku olacaktır. Eğer kuşku ile başlar ise kesinliklere ulaşacaktır. Ben, bilimlerin, sanatların, bütün insan bilgisinin yeniden kurulmasını istiyorum” (Francis Bacon).

Modern anlamıyla bilim, Francis Bacon’ın tümevarımsal yöntemi üzerine kurulmuştur ve doğa araştırması anlamını içerir. Bacon, bilimsel bilginin sonu gelmez deneylerle üretilebileceğini, ulaşılan sonuçların ise gözlemlerle test edilmesi gerektiğini ileri sürmekteydi (Bkz. Bacon 2010). Bu güne kadar doğa bilimcilerin, Baconcu uyarıları dikkate aldıkları görülmektedir. Bununla birlikte Bacon, bugün özellikle sosyal bilimlerde pek de gerekli ve geçerli olmayan bir öneride bulunmuştu: İyi bir hipoteze sahip olmak için iyi gözlem sonuçlarına sahip olmak gerekir (Osborne 2006:97). Tümevarımsal yöntemi, nesnesi “fizik, kimya, biyoloji” olan alanlara tatbik eden ve orada öngörü kabiliyetine sahip yasalar arayan Brahe, Kopernik, Kepler, Galilei, Newton, Faraday, Curie gibi bilim dehaları da her zaman daha çok gözlem yapmanın, daha çok bilim anlamına geldiğine inanmaktaydılar. Gözlem ve deney, doğanın işleyişini öğrenmenin yegane yolu olarak görülmüştür. Fakat deney, gözlem ve ölçüme dayalı tümevarımsal yöntem zamanla zayıflamıştır. Mayer-Schönberger ve Cukier’e göre ise büyük veriyle gelinen bugünkü noktada insanlar, dünya hakkında bilgi edinmek için artık deney ve gözlem sonuçlarına değil, veriye bakmaktadırlar (2013:27).

Veri, bilginin temel koşulu olduğu gibi, bilgi de felsefe yapabilmenin temel koşuludur. Filozoflar ve modern anlamda bilim insanları, bugüne kadar verinin kendisine bakmak konusunda zayıf, verinin nihai evresinde erişilen bilgiye bakmak konusunda ise güçlü bir eğilime sahip oldular. Veriden bilgi ve bilgeliğe ulaşan sürecin felsefi açıdan alınması, apriori başlangıçlardan hareket eden Descartes, Kant, Husserl gibi filozofların devam ettirdiği salt felsefi bilgi kuramı geleneğinin bize sunduğundan farklı içgörüler sunabilir. 1960’lı yıllarda ortaya çıkan “bilgi devrimi” ve “dijital çağ” gibi kavramlara nesnel bir zemin sağlayan “büyük veri”, 2000’li yıllardan itibaren etkisi neredeyse bütün araştırma disiplinlerinde hissedilen bir olgu (fact) haline geldi. Bu açıdan büyük veri, genelde bilim, özelde ise beşeri bilimler için önemli bir dönüşümün başlangıcına işaret etmektedir.

Büyük veriye dair felsefemiz, sosyal bilimler, bilim felsefesi ve bilim yapma tarzını tespit etme açısından fazlasıyla önemlidir. Donah Boyd’a göre artık hiçbir araştırma disiplini tarafından göz ardı edilemez bir olgu olan büyük veri, kesinsizlik ve hızlı değişim içeren ortamlarda kendine yer bulmakta, böylece aldığımız kararları ve geleceğimizi şekillendirmektedir (Boyd 2012:664). Bunu yanında “Petabyte Çağı” teriminin mucidi

Anderson'ın haklı olarak sorduğu, “Bilim, Google'dan ne öğrenebilir?” gibi sorulara cevaplar bulmamız gerekmektedir. Aynı felsefi sorularla ilgilenen Sabina Leonelli ise dijital çağda neyin bilimsel kanıt, bilimsel kuram veya bilimsel deney olarak kabul edileceğini anlamak için anahtar kavramlara ihtiyaç olduğunu ileri sürer (Leonelli 2014).

Gerçekten de bugüne kadar, “verinin bizi götürdüğü yere bakmak için çok az girişim olduğunu” söyleyebiliriz (Michael Minelli, 2012, s. 48). Büyük veri felsefesi inşa etmenin aciliyetine inanan Leonelli'ye katılarak veriyi takip etme, arşiv araştırmaları, görüşmeler, açık bilime dair politik angajmanlar, yönetici ve kullanıcılarla işbirliği yapma gibi yöntemleri içeren, deneysel olarak kurulmuş bir bilim felsefesi önerebiliriz. Büyük veriye dair “bütüncül değer önermeleri kuramı” geliştirmeye çalışan James Golden'a göre, eğer verinin gerçekte bir platform olduğuna inanıyorsak veriyi bir insan gibi yönetmeye başlamalıyız (Michael Minelli, 2012, s. 47). Golden'ın önerisi; veriyi bir araya toplamayı, entegre etmeyi, çözümlenmeyi ve yönetme bileşenlerini içermektedir.

Bilimsel yöntemin en temel karakteristiği, deneyden türemesidir, gelenekten veya derin düşünümünden (reflection) değil. Harris, Kant'ın çok alıntılanan “Felsefesiz bilim kör, bilimsiz felsefe boştur”¹ biçimindeki ünlü özdeyişini hatırlattıktan sonra şöyle yazar: “Veri, hem bilime hem de felsefeye ihtiyaç duyar”. Felsefe ile bilim arasında, filozof ile bilim insanı arasında semantik bir ayırım bulunur fakat bilgi söz konusu olduğunda ikisi arasında ilişki ve etkileşim zorunlu hale gelmektedir. Bir filozofun işi, dünya hakkındaki gerçekleri gözleyerek değil, düşünerek bulmaktır. Bilim insanın işi ise filozofun düşündüğü “aynı gerçekliği”, gözlem, deney ve ölçüm yaparak bulmaktır. Olgu ve düşünce arasındaki ilişki konusunda Firestein, Harris'in aktardığına göre şöyle yazmaktadır: “Bilim insanın çalışması, olgusal bataklıkta bir çıkmaza saplanmaz çünkü onlar, olguların sayıca çok olmasına aldırış etmezler. Bu, onları hesaba katmama veya görmezden gelme değildir, daha ziyade onları kendinde sınırlar olarak görmemektir” (Harris 2013).

Büyük veri felsefesinin içinde bilişsel, epistemolojik ve etik meseleler iç içe geçmiş bir halde bulunmaktadır. Bilginin değeri, sınırları, kaynağı ve ölçütleri açısından, etiğin ölçütleri ve sosyoloji açısından büyük veriye hem delicesine âşık olma hem de onu lanetleme, yaşadığımız büyük veri çağını yanlış anlamamıza sebebiyet verebilir. Tıpkı teleskobun evreni kavramamıza ve mikroskobun hücreleri anlamamıza imkân tanıdığı gibi, çok büyük bir veri yığınına toplamak ve analiz etmek için kullanılan yeni teknikler de henüz takdir etmeye başladığımız biçimlerde dünyamızın anlamını kavramamıza yardım edebilir (Mayer-Schönberger ve Cukier 2013:15).

¹ Newton mekaniği ışığında kavram ile duyumun, akıl ile deneyimin etkili bir sentezini geliştiren Immanuel Kant'ın felsefesini özetleyici meşhur sözü şöyledir: “Kavramsız algılar kör, algısız kavramlar boştur” (Kant 2014: &B17).

2. BÜYÜK VERİNİN FARKLI TANIMLARI VE KAYNAKLARI

a) **Büyük verideki büyüklük:** Latince tekil “datum” sözcüğünün çoğulu² olan veri (data); bir uslamlama, tartışma veya hesaplama dayanağı olarak kullanılan olgudan elde edilmiş bilgi anlamına gelir (www.merriam-webster.com 2015). Veri; kaydedilen, yeniden düzenlenebilen ve analiz edilebilen bilgi (information) birimidir.

Bugüne kadar yüzlerce bilim dalı, araştırma disiplini, edebi ve felsefi eser ortaya konuldu. Bütün literatür, tek kelimeyle “devasa” olarak nitelendirilebilir. Büyük veri söz konusu olduğunda ise geçmiş literatürün tümü, devede kulak denilebilecek kadar küçüktür. Sahip olduğumuz ve yakın gelecekte sahip olabileceğimiz veri miktarı hakkında, aşağıdaki örnekler bize bir fikir verebilir:

- i. 2012 rakamları ile dünyada günlük 2.5 kentirilyon byte veri üretilmektedir. Bazı istatistiklere göre 2020’de ulaşılacak veri miktarı, 2009’un 44 katı olacaktır (Wikipedia 2013).
- ii. 2010’lu yıllarda dünyadaki toplam bilişim harcamaları yılda %5, üretilen veri miktarı %40 artmaktadır (Wikipedia 2013).
- iii. Sloan Dijital Gökyüzü araştırması 2000 yılında başladığında, New Mexico’daki teleskop, ilk birkaç haftasında tüm astronomi tarihi boyunca toplanandan daha fazla veri topladı. Topladığı bilgi miktarı 10 yıl içinde 140 terabayta ulaştı. 2016 yılında Şili’de hizmete başlayacak olan Geniş Snoopik Araştırma Teleskobu ise aynı miktar veriyi, 5 günde toplayacak (Mayer-Schönberger ve Cukier 2013:15).
- iv. 2000 yılına kadar dünyada saklanan bilginin yaklaşık dörtte biri dijital idi. Geri kalan kâğıt, film, vinil uzun çalar plaklar, manyetik kasetler vb. üzerindedir. Hilbert’in belirttiğine göre, şimdilerde dijital veriler, her 3.5 yılda 2 katına çıkmaktadır. Analog bilgi ise tam tersine neredeyse hiç büyümektedir. 2007 yılında % 7 olan analog bilgi miktarı, 2013 yılında %2’ye gerilemiş durumdadır (Mayer-Schönberger ve Cukier 2013: 17).

Bir veri setinin “büyük” addedilmesi için onun terabayt ölçeğinde olması gerekmez. “Büyük veriyi büyük yapan, tüm veri setinin mümkün olan en büyük bölümünün kullanılmasıdır” (Mayer-Schönberger ve Cukier 2013: 36). Edd Dumbill’e göre büyüklük, “konvensiyonel yöntemlerin gücünü aşmak” ile ilgilidir. McKinsey ise saklaması, yönetmesi ve analizi “tipik veri bankası yazılım araçlarının yeteneklerini aşan ölçekteki veri setlerini” büyük olarak kabul etmektedir (Bkz. Minelli 2012:5).

² Bir galat-ı meşhur olarak data, yapı bakımından çoğul olmasına rağmen tekil yapım eki alır. Örneğin İngilizce yazımında da “data are...” olması gerekirken çoğullukla “data is...” olarak kullanılmaktadır, tıpkı “agenda” ve “agendum” isimlerinde olduğu gibi. Bu artık yerleşiktir.

Son birkaç yılda popüler olmuş³ kavramlardan biri olan büyük veri, klasik veri tabanlarının kaldıramayacağı büyüklükteki veriyi ifade eder. Dijital çağda verinin büyüme hızı, klasik depolama ünitelerini aşmaktadır. Konvansiyonel yöntem ve tekniklerle saklanamaz, akışı takip edilemez, analiz edilemez. İlişkisel veri tabanları gigabaytlar, büyük veri tabanları ise petabaytlar seviyesinde veri saklayabilir.

b) Büyük veri kaynakları: Veri oluşumundaki hızlı artışın başlıca nedeni, sosyal ağ etkileşimlerinin büyüyen hacmi, lokasyona duyarlı cihazların yaygınlaşması, fiziksel dünya hakkında bilgi yakalayan ve ileten 'akıllı sensörlerin sayısındaki artış olduğu gözlenmektedir (Wikipedia, 2013). Tabiki bunlara bir kısmı aşağıda verilen multi-medya kaynakları da eklenebilir:

- i. Fotoğraf, ses ve video belgeleri
- ii. Web sunucularının log dosyaları
- iii. Sosyal medya yayınları
- iv. İnternet istatistikleri, blog ve mikrobloglar, ağ günlükleri
- v. GSM operatörleri
- vi. Hastane kayıtları
- vii. Tanıma sistemleri
- viii. Hava durumu sensörleri
- ix. DNA dizilişlerinin analizi
- x. Mobese kayıtları
- xi. Sosyal medya paylaşımları
- xii. İklim algılayıcıları vb. sensörler

Büyük veri analizi, bütün veri kaynaklarından toplanan, belli bir yapısı olmayan (yapısal-olmayan) tüm verinin anlamlı ve işlenebilir bir biçime dönüştürülmesidir. İlişkisel veri tabanları yapısaldir. Büyük veri tabanları, yapısal olmayan değersiz sayılan bilgi yığınlarıyla doludur. Fakat büyük veri analiz sistemleriyle neredeyse "çöplükten hazine çıkarma" işlemi başlatılmış (Wikipedia, 2013), böylece önemli, kullanılabilir ve yararlı yönleri ortaya çıkarılmıştır. Gelineen noktada, büyük veri çağında, konvansiyonel veri bankaları ve veri madenciliği yöntemleriyle karar almaya devam etmek pratik ve makul değildir.

Bir veri setinin ne zaman "büyük" kabul edileceğiyle ilgili tartışmalar var olsa da "büyük veri" ile ilgili, beş nitelik üzerinde bir tartışma yoktur. Bunlar: 1. Çeşitlilik (variety), 2. Hız (velocity), 3. Büyüklük (volume), 4. Doğrulama (verification), 5. Değer (value). Bu özellikler, İngilizce karşılıklarına uygun olarak "5V" özellikleri diye kodlanmaktadır. Çeşitlilik; her teknolojinin kendine özgü veri üretebilmesinden ileri gelir. Büyük veri tabanındaki veriler

³ Wikipedia, 2013 yılında büyük veri (big data) maddesini açmıştır (Bkz. www.wikipedia.org 2013).

aynı kodla yazılmış değillerdir (non-unicode). Hız, büyük verinin üretilme hızının gittikçe artmasını ifade eder. Büyüklük, toplanan veri miktarının her an arttığını; doğrulama, enformasyon yoğunluğu içindeki gelen verinin “güvenli” olmasını, başka bir ifadeyle verinin olması gereken güvenlik seviyesinde izlenmesini, doğru kişiler tarafından görünebilir olmasını veya gizli kalmasını ifade eder. Ve son olarak değer, veri üretim ve işleme katmanlarınızdan sonra kurum için bir artı değer yaratmasına gönderme yapar.

c) Teknolojik ilerleme, dijital çağ, bilgi toplumu: Büyük veri olgusunun yeter-sebebi, yalnızca bilgisayarların veya sensörlü akıllı cihazların yaygınlaşması değildir. Büyük veri dönüşümünde bilgi-işlem teknolojilerine yapılan vurgu aşırıdır ve paradigma dönüşümünü kavramak konusunda yanlış yönlendirmektedir. Vurgu, insanlığın teknoloji üretme yeteneğine değil, gerçeği arama, deney, gözlem ve ölçüm yapma, verileri kaydetme ve analiz etme isteğinin kesintisizliğine ve bu konudaki başarısına yapılmalıdır. Burada dikkat çekmesi ve çözümlenmesi gereken şey, bir araç olan teknoloji olamaz, bir amaç olan veya teknolojiyle erişilen bilgi olabilir. Dijital çağın bilgi toplumu, ansiklopedik kütüphane bilgileriyle (information) değil, geniş veri depolama alanlarındaki devasa verileri (data) işleyen ileri hesaplama teknolojileri, yaşam biçimleri, düşünme pratikleri ve bilim yapma tarzıyla desteklenen bir toplumdur. “Elektronik iletişim ve bilişim teknolojileri; üretim, transfer ve tüketimleri esnasında ‘kendilerine has’ yeni bir yaşam pratiği oluşturmuş, bir yere kadar mevcut düzende değişikliklere yol açmıştır” (Işıklı 2014a). Bu yüzden internet destekli dijital evrenin yapısı ve işleyişi, kendine özgü ve biriciktir; daha öncekilerden kökensel biçimde farklılaşır.

İnternet sayesinde, neredeyse “küresel bilinç” düzeyine erişen dünya, tüm geçmiş zamanların toplamından daha fazla bilgi üretmekte ve işlemektedir. Bilgi, ilk kez olay ufkunun genişleme hızını takip edebilme umudunu yakalamıştır. Fiziki evrenin büyük patlamayla başlayan genişleme serüvenine paralel olarak dijital veri evrenimiz sürekli genişlemektedir. Bu genişlemeyi devam ettiren enerjinin kaynağı veridir. Veriyi sosyo-ekonomik değeriyle ele alan Cukier’in pragmatik tespitleriyle söyleyelim, “Veri, bilgi ekonomisinin petrolüdür” (2013:23). Bu durumda büyük veri için şunu söylemek uygundur: Zengin ve yaygın iletişim, bilişim ve diğer yüksek teknolojiyle donatılı bir bilgi toplumunun bir çıktısıdır. Bilgi toplumunda, simülasyonlar sayesinde milyarlarca veri hızlı ve kolayca işlenebilmektedir.

3. BÜYÜK VERİNİN EPİSTEMOLOJİK DEĞERİ

Hesaplamalı düşünmeye dayalı yeni bilim yapma tarzı, temel dayanağı olan büyük verinin nitelikleri yüzünden modern bilimin pozitivist yorumuyla örtüşmeyen yapısal farklılıklar gösterir. Pozitivizmin ve onun çağdaş bir yorumu olan analitikçi neo-pozitivizmin, bilimsel araştırmalar için sunduğu ontolojik ve epistemolojik varsayımları, hesaplamalı düşünme

biçimini kuşatamamaktadır. Pozitivizm; felsefede Descartes, bilimde ise Galilei ile başlayan “modern” paradigmanın önkabullerini (sayıltılar) sunmaktadır. 20 ve 21. yüzyılda ortaya çıkan bazı bilimsel gelişmeler Bu önkabüllerden bazıları şöyledir:

- i. Gözlemciden bağımsız dışsal gerçeklik (nesne anlayışı)
- ii. Ölçüm veya gözlem esnasında nesneyi tedirgin etmeyen yalıtık gözlemci
- iii. Araştırma süreçlerine asla karıştırılmayan ve araştırma sonuçlarının evrensel olmasını temin eden tarafsızlık ve objektiflik
- iv. Çıkarımların deney, gözlem ve ölçüm verilerine dayandırılması
- v. Genel yasalara ulaşmak için tümevarımsal akıl yürütmeye (uslamlama), genel önermelerden hareketle tekil olayları açıklamak için tümdengelim (çıkarm) başvurulması
- vi. Matematikselleştirilme
- vii. Önce etkinin gerçekleşmesi, daha sonra tepkinin ortaya çıkması olarak belirlenimcilik (nedensellik): Belirlenimciliğe göre meydana gelen her tekil olayın en az bir gözlenebilir bir nedeni olmalıdır; ve neden, zamandizinsel bir çizgide, olaydan önce gerçekleşmelidir.
- viii. Öngörü (prediction), olayların önceden yer ve zamanın kesin olarak bilinebilmesidir.

Büyük veriden önce, modern pozitivist bilim paradigmasının belirlenimcilik (nedensellik), dışsal gerçeklik ve yalıtık gözlemci gibi mütemmim cüzleri (olmazsa olmazlar) kuantum mekaniği tarafından ihlâl edilmiştir (Işıklı 2012:236). Bağlamsallık, araştırmacının yönelmişliği, bilginin temel dayanağı olarak fiziksel olarak sert ve gözlemciden bağımsız nesne tasarımı, etki-tepki zincirinin takip edilebilirliği, kesin öngörü (prediction), veri toplama tekniği olarak deney, gözlem ve ölçüme başvurma, tekil olayları açıklamak için geliştirilmiş olan temelde etkileri (nedenler) eş zamanlı veya ardışık olarak sıralamaktan öteye geçmeyen bilimsel açıklama gibi unsurlar, büyük veride ya zayıflatılmış ya terk edilmiş ya da ihlâl edilmiş durumdadır. Bu durum, büyük veri paradigmasıyla pozitivist paradigma arasına bir ayırım koymamıza izin verir. Böylece veriye bağlı hesaplamalı düşünme biçimi ve büyük veri temelli bilim yapma tarzının nitelikleri, modern tarzda olmayan bilgi edinme sürecini ortaya koyabilir.

a) Kullanılabilirlik: Bazı uzmanlara göre büyük veriyi anlamak için gerekli olan anahtar kavram, “kullanılabilirlik”tir (Minelli vd. 2012:5). Bu perspektife göre tüm mesele, terabayt ve eksabayt ölçeğindeki veri yığınlarından işimize yarayan, bir sorunumuzu çözen, bize pratik yararlar sağlayan “değerli” bilgiler elde edebilmektir. Kullanılabilirlik, kullanılmayan verilere de sahip olduğumuz anlamına gelir. Kullanılmayan veri kötüdür ve John Locke’ın “ıvır zıvır (triviality) dediği, bilgiye ulaşma sürecinde, hiçbir olumlu katkısı olmadığından ayıklanması gereken, aralarında anlamlı ilişki kurulamayan algı unsurlarına benzer (Bkz. Locke 1992:221). Ne var ki büyük veri paradigmasında, veri yığınlarından veya ıvır zıvır verilerden, kayda değer bilgiler elde edebileceği fark edilmiştir. Veri madenciliği

çalışmalarında, yapısal olmayan ilişkisel veri tabanlarında tutulan, çoğu zaman atık ve atıl halde bulunan, bilgi çöplüğü konumundaki değersiz verileri arasında içsel bağlantılar bulunabilmektedir. Yığılan (cumulative) felsefi bilginin içsel bağlantılarını, filozofların uzun zaman harcayarak ve de sezgisel biçimde keşfettikleri göz önüne alındığında, dijital veri analizlerinde atıl veri yığınları arasındaki içsel bağlantıları görece çok kısa bir zamanda ortaya çıkaran veri işleme modellerinin geliştirilmesi, bilimsel gelişme açısından şaşırtıcı bir olgu olacaktır. Üstelik dijital veri, bir kere kullanıldığında çöpe atılan kağıt peçete gibi değildir. Kullanılmış veri, tekrar kullanılabilir. Verinin sürekli yeniden kullanılabilmesinin fark edilmesiyle bit pazarında nur yağmış, ıvır zıvır bilgiler değer kazanmıştır. Bu değer, verinin entropiye karşı dayanıklılığına bağlanabilir olsa da aslında dijital verinin çoklu kullanıma elverişli doğasına atfedilmelidir.

Diğer yandan nasıl kullanılacağı bilinmeyen çok sayıda verinin varlığı, verinin kendinde tekil bir anlama indirgenmediği, bu yüzden de gerçek yaşam açısından anlamsız olduğunun kanıtıdır. Veri çoklu değer taşır. Fakat bir araştırmacı tarafından, bir model ile belli yönleri talep edilmedikçe veri, süperpoze bir durumdadır. Başka bir ifadeyle veriyi anlamlı kılan, onu enformasyon ve bilgiye dönüştüren öznedir. Öznenin geliştirdiği hipotez, hangi verilerin işleme dahil edileceğini, model ise verilerin nasıl ve niçin işleneceğini belirler. Anlam da işleme süreci de öznenin seçiminde ortaya çıkar. Kendinde şeylerde (things in itself), gözlenebilen dışsal bağlantılar yoktur. Buradaki sorun, hipotez tarafından varsayılan ve model tarafından konumlandırılan ilişkinin, şeylere yahut verilere içkin olup olmadığıdır. Ayrıca bir veri, birden fazla amaç için kullanılabilir. Çok amaçlı kullanım, verinin hammadde olmasıyla ilişkilendirilebilir fakat esasen verinin tekil mevcudiyetinin anlamsızlığını gösterir. Zira, Aristoteles'in tespit ettiği gibi, tekilin bilimi olmaz (Aristoteles 1996:435).

b) Sosyo-teknik-dijital olgular: Modern uzmanlaşma, disiplinlerarası kavramlara anlam verecek, onları işlevselleştirecek bir esneklikten yoksundur. Bu yüzden modern bilim uzmanları, yalnızca alana özgü olguları, yine alana özgü yöntemlere başvurarak araştırır; yine alanla sınırlandırılmış genel yargılara ve çıkarımlara ulaşırlar. Toplumlar, tarihsel süreçte demokrasiye evrilerek farklı birey ve kültürleri demokrasi kültürü altında, dışlayıcı olmayan bütüncül "insanlık" kavramı altında barışçıl rekabete, kuramsal düzeyde de olsa ulaşabildi. Benzer şekilde, tutarlı bilim mantığı içeren bütüncül bir bilim felsefesi oluşturma biçimindeki entelektüel yönelmişliğin doğal bir ürünü olan "bilimlerin birliği ideali", içsel türdeşlik gerektiren bir özdeşliğin sağlanmasını salık verir. Belliki modern bilimsel bakış açısıyla ele alındığı müddetçe modern "fen" ve "beşeri" bilimler, olgu ve yöntem farklılığı nedeniyle tek bir çatı altında bir araya getirilememektedir. Antidemokratik tutumla "bilim olma" niteliğini kendi tekeline alan doğa bilimleri, bilimlerdeki çeşitlilik ve farklılıkları, "olgu" tasarımına dayanarak kuvvetli bir tekbiçimleştirme eğilimi altında bastırmaktadır. Doğal olgu (naturafacte) ve yapay olgu (artefacte) ayrımı (Bkz. Özlem 2008:24), uygulanacak

yöntemi önermekten ziyade, olgu türünden hareketle bir araştırmanın “bilimsel” olup olmadığına karar vermek için kullanır. Bu ayırım, geçerliliğini artık sürdürmez. Sayısallaştırılmış, kaydedilmiş, işlenmiş sosyal olgular arasında belirlenimci ilişkiler bulunamaz fakat araştırmacılar ve analizcilerden bağımsız olarak “orada öylece duran” veri olgusundan bahsedilebilir. Bu artık ne naturafacte ne de artefacte’dir, “sosyo-teknik-dijital” bir olgudur. Dijital çağın postmodern bilim yapma tarzının konusu olan büyük veri de dijital teknolojik alt yapı ve çıktılara temellenen sosyo-teknik bir olgudur.

c) Hız ve çeviklik: Bilimsel çalışma, genelde ağır ve ciddi bir iştir; pratik çıktıları ve topluma yansımaları hemen görmek mümkün değildir. Bu yüzden bazı kültür ve teknoloji eleştircileri, bilime erişim ve nimetlerinden yararlanma konusundaki eşitsizlik bağlamında bilimi, bir ihtiyarın meyve ağacı dikmesine benzetmişlerdir. Görece kısa dönemde çıktıları elde edilen AR-GE çalışmaları dahi, bilimsel bilgiyi, tekniğe ve problem çözümüne yeterince hızlı yansıtamamaktadır. Büyük veri dönemine gelindiğinde, “zamansal kayma aralığı”nın hissedilir derecede kısaldığı görülmektedir. Büyük veri dönemi kişi ve kurumları, teknolojik kontrolün, teknolojiyi kontrol ve teknolojiyle kontrolün her şeyi değiştirdiğini fark etmeye ve hızlı değişime iyi hazırlanmaya zorlamaktadır. Bilgi ve teknolojinin toplumla etkileşimindeki tepkime süresi kısalmaktadır. Teknolojinin icadı veya bilginin keşfiyle bunların kişi ve kurumlardaki etkisinin hissedilme süresi pergelinin daralması, teknolojik tepkime süresininin azalması, her zamankinden daha hızlı bir değişim, hiperrekabet ve kesintili dünya içinde yaşadığımız anlamına gelir. Bu hız ve çevikliğin boyutunu hayal edebilmek için Minelli, bir analogik kıyas yapar:

“Geleneksel çözümleyicilerin Büyük Veri çözümleyicileriyle kıyaslamak, at arabasıyla traktörü kıyaslamaya benzer. Hız, ölçek ve karmaşıklık devasadır” (Minelli vd. 2013:6).

Yakın gelecekte kurumlar, şu üç perpektifte kendilerini çevik olacak biçimde yeniden dizayn etme ihtiyacı hissedecekler: yeni düşünme becerileriyle donanımlı personel, çevik ve esnek işlem-süreci, teknolojik inovasyon ve yaratıcılık. Bu hızlı dönüşüm süreklidir.

Hız ve çeviklik bir gelişme olduğu kadar acil bir ihtiyaçtır. Teknoloji, hız sunarak zaman kazandırma vadiyle hayatımıza girer fakat birey ve kurumları, daha hızlı ve çevik hareket edebilme kabiliyetlerini geliştirmeye zorlar. Gelineen noktada bugün hiçkimse, bir şirketin performansını veya müşteri davranışlarını anlamak için haftalarca beklemek istememektedir (Minelli vd. 2012:131). Sürekli dönüşümlerle baş edebilmek için birey ve kurumlar, çeviklik ve esneklik niteliğine sahip olmalı ki hızlı yineleme, ön yüklemeli düşünme, yeni ve sıradışı hipotezler geliştirime, görüntülemeye uyum sağlama (adaptation) yeteneklerini geliştirebilmelidirler.

Bir süredir, büyük veri yığınlarının kısa dönemli amaçsal çözümlenişi meselesi, yoğun ve acil bir işgücü talebi olarak karşımızda durmaktadır. Dijital çağda, efektif kararlar alabilmek

hayatidir. Günümüzde, çevre ürünlerine yayılan modeller, inşa edicileri tarafından periyodik olarak yeniden değerlendirilmektedir. Sürekli araştırma yaparak ve yeni ortaya çıkmış teknik, teknolojiler ve uygulamaları istimal ederek daha çok nefes alıp vermek ve problem çözmeyi derinleştirmek gerekmektedir. Bu yüzden hem sürekli gözden geçirilmekte, yeniden düzenlenmekte, yeni bir amaca uyarlanmakta hem de muazzam derecede çeşitlendirilmektedir.

d) Bütüncüllük ve hipotez: Modern bilim, parça – bütün analogisinin genel-geçer olduğunu varsayar; araştırmalarını, olgusal evreni temsil etme kabiliyetine sahip olduğunu kabul ettiği “tekil olguların gözlemi” üzerinden yürütür. Gözlem ve ölçümü ise bilim insanının özgür bir seçimle kurduğu hipotezleri doğrulamak için gerçekleştirir; gözlem ve olguları doğrulayacak gözlem ve deneyleri hesaba katar. Gözlem ve ölçüm verilerinin tümünü kullanma gibi büyük bir yükün altına girmez. Bilimselliğin modern bir ölçütü olarak böylesi bir doğrulanabilirlik, bu seçmeci tavrı yüzünden Popper gibi düşünürler tarafından haklı olarak eleştirilmiştir (Güzel 1996:49-52). Diğer yandan tümevarımcı modern bilimsel yöntem, örneklerin (aynı türden olguların) tümünü dikkate almadığı halde aynı türden olgusal evrenin tümü hakkında genellemede bulunur. Bu genelleme, genel – geçer yahut tümel yasa adını alır. Açıktır ki modern bilimin çıkarım yöntemi olarak kullandığı bu tarz bir tümevarım eksiktir. Çünkü ne olguların ne de verinin tümünü dikkate almaktadır; askine seçmeci bir tavır sergilemektedir. Burada olguyu aşan bir yön, Aristoteles ve diğer rasyonalist epistemolojiler geliştiren düşünürlerin “intellect (akıl)” dedikleri kuramdır. Rasyonalist epistemolojiye göre gözlemi mümkün, veriyi anlamlı kılan kuramdır. Kuramın yerini büyük veride model alır. Kuram veya model, paradigma içindeki alt kümelerdir.

Büyük veriyle destekli bilimsel araştırma, örneklemin dar, sınırlı ve yetersiz numelerinden evrenin tümünü analiz eden bir imkana kavuşmaktadır. Büyük veri, özellikle beşeri bilimciler için yöntem bilimsel bir “müjde” veya “sevindirici haber” gibi ilan edilmektedir. Bu sevindirici haber, şimdilik dar bir alanda şöyle yankılanır:

Verinin tümüne bakmanın bir yolu bulundu!

Son on yılda popüler olan büyük veri uzmanları, veriye bir bütün olarak bakmayı sağlayan bu yeni yöntemi “büyük veri analiz sistemleri” olarak adlandırmaktadırlar. Şurası bir gerçek ki ne olguların ne de olguya dair verilerin tümünü dikkate almak, şimdilik entelektüel bir imkansızlıktır, teknik değil. Tüm veriler arasında anlamlı korelasyonlar bulmak için verinin titiz ve sürekli bir analizine ihtiyaç vardır. Henüz o aşamada olduğumuzu söylemek imkansızdır. Yine de eldeki tüm veriye bakabiliyor olmak, ileride en azından ilişkili tüm olgulara bakabilmek için hazırdaki bir imkandır; edimselleşmesi gerekir.

Büyük veri çağında, hangi değişkenlerin inceleneceğine sadece hipoteze güvenerek karar vermek artık elverişli değildir. İncelenebilecek veri setleri çok büyük, alan çok geniş, bilgi (information) ise çok dağınıktır. Bu yüzden hipotezin yönlendirdiği yaklaşımı terk etmek,

onun yerine verinin yönlendirdiği bir yaklaşımı önemsemek gerekir (Mayer-Schönberger ve Cukier 2013:63).

e) Örneklem almanın terki: Büyük veri öncesine ait geleneksel modeller, örneklem denilen bir veri alt kümesi kullanılarak oluşturulmaktaydı. Geleneksel model bir kez oluşturulduktan sonra biraz daha büyük ikinci bir veri kümesi, yeni oluşturulan modeli test etmek veya geçerliliği bulmak için kullanılırdı.

Pozitivist bilimlerde tatbik edilen örneklem, gerçek evrenin bir prototipi olarak kabul edilir ve burada gizli bir parça ile bütün arasında bir simetri olduğunu varsayan bir analogi, eksik tümevarım olarak sunulur. Beşeri bilimlerde sık kullanılan dolaylı ölçüm araçlarının iki temel şartı vardır. Geçerlilik ve güvenilirlik. Bir modelin geçerliliği, sistematik hata içermemesiyle ilgilidir. Sistematik hata içermeyen geçerli modeller, gerçekte ölçülmesi gerekeni ölçen araçlardır. Güvenirlik, denge ve tekrarlanabilirlikle ilgilidir ve bir ölçümün tutarlı ve dengeli sonuçlar çıkarıp çıkarmayacağını söyler. Bir modelin güvenilirliği, rastlantısal hata içermemesiyle ilgilidir (Erdoğan 2007:232-236). Model, bir kez geçerlilik kazandıktan sonra bütün veriyi saymak, hesaplamak ve değerlendirmek için yetkin bir araç gibi kullanılmaktadır. Standart istatistiksel testler ve bilimsel çıkarım yapan hesaplama süreçleri, büyük popülasyondan küçük bir örneklemin analizine yönelik tasarlanmışlardır. Matematiği kullanan bu istatistiki yaklaşımın, beşeri disiplinlerin bilimsellik niteliği kazanmasında anahtar rol oynadığı görülmektedir. Bilimin fizikalist yorumuna göre, bir araştırmacının bilimsel olmasının asgari kriterlerinden biri de matematiselleştirilebilirliktir. Sosyal bilimlerde, 20. yüzyılın son çeyreğine girerken matematiği istatistik olarak kullanabilmiş, böylece “kısmen modern bilim” hüviyeti kazanabilmişlerdir.

Büyük veriyle birlikte veri uzmanları, bir model oluşturmak için, artık örnekleme olduğundan çok daha fazla veriyi ya da verinin tümünü kullanmaya başladılar. Bunu yaptıklarında, veri uzmanları, modelin doğruluğunu artırmak için modele, öngörü değişkenleri ilave edebildiklerini fark ettiler. Çünkü büyük veri, bazen popülasyonun tamamını veya büyük kısmını içeren çok büyük bir örneklem sağlar. Minelli'nin belirttiği gibi büyük veri analiz modelleri, kullanılan tarihsel veya zamansal verinin miktar ve oranına bağlı olarak mevcut trendleri tanımlayabilmektedir (Michael Minelli, 2012, s. 87).

f) Karar-destek bilimlerinin yükselişi: İş süreçlerine dahil edilmiş karar-destek bilimleri, organizasyonların yeni etkinlik öngörülerini keşfetmek ve öngörü-karar-eylemler arasındaki uçurumu azaltmak için yeteneklerini geliştirmede kritik önemdedir (Minelli vd. 2013:130). Büyük veri ve karar-destek bilimlerinde gözleme bağımlılık yoktur. Veri kümeleri, belli bir zaman periyodunda, sistemin bir çıktısı olarak kendiliğinde birikir. Hazır durumdaki devasa veriden anlamlı sonuçlar çıkarmanın görece zorluğunu kabul edilmekle birlikte, tümelleri ve trendleri tespit edebilmek, geri dönüp tekilerin davranışlarını çıkarsayabilmenin kesinliği yadsınmamakta, aksine onaylanmaktadır. Bu yüzden veri tabanlı hesaplamalı düşünme

yöntemi, büyük veri analizini olduğu gibi karar bilimlerini de (decision sciences) sonuç vermiştir.

g) Belirlenimsizlik (indeterminizm) ve bağlantı (correlation): Büyük veri analistleri, örneğin insanların neyi izledikleri ile neyi satın aldıkları arasında, hiç de nedensel görünmeyen ilişkiler (bağıntı veya korelasyonlar) kurarlar. Burada, camın kırılmasını atılan taşa bağlayan doğa bilimsel belirlenimciliğin neden ile sonuç arasındaki ardışıklık ve yeter-sebebe ilişkisi yerini, bağlantıya (correlation) bırakır. Belirlenimciliğin tutarlı ve derin felsefi biçimini, kuantum kesinsizliğine karşı hem sunan hem de savunan Einstein'ın yerellik görüşüne göre, neden ile sonuç, uzay-zamansal bir yakınlıkta, etkinin ışık hızıyla veya ışık parçacığıyla aktarılabilmesini zorunlu kılar (Einstein vd. 1935:779). Nedenselliğin bu tarz yerelci yorumu, bizi neden ile sonucu bir arada görmeye zorlar. İnsan davranışları söz konusu olduğunda, psikolojik olarak hazır bulunmuş düzeyini oluşturan unsurların satın alma eylemine neden olduğu kabaca önerilebilir. Fakat uyarıcı ile tepkinin şimdi ve burada birlikte bulunmayışı, önerimiz üzerinde daha dikkatli düşünmeye iter. Önceki bir etkinin, sonraki bir ortamda bir tepkiye doğrudan yol açması, her zaman çoklu uyarıcılara maruz kalan insan duyarlılığının sonraki bir ortamda, önceki etkiyi bünyesinde saklayarak özgürce tepkide bulunduğunu söylemek aceleci yargıda bulunmaktadır.

Bugün yaygın olarak kullanılan istatistiksel ölçülerden biri olan korelasyon kavramı, örneğin kol uzunlukları ile boy uzunluğu arasındaki ilişki gibi, eş zamanlı (simultaneously) olarak var olan, ardışık değil, iki veri arasındaki istatistiksel ilişkiyi inceler. Korelasyon, gözlem yapma biçimimizde radikal değişikliğin bir göstergesidir. Büyük veri analizi ve öngördüğü trendler, nedensel (causal) ilişkisi arama zorunluluğunu zayıflamış, korelasyonel ilişkileri bulma arayışını güçlenmiştir. Örneğin Araba sahibi olmak ile antibiyotikleri doktorun tavsiye ettiği biçimde almak arasında hiçbir nedensel ilişki yoktur; bu ilişki korelasyoneldir. Olayların nedensel açıklamasına dayanmadan, sadece korelasyonel ilişkilere odaklanarak gelecekte ne olacağını tahmin edebiliriz. Büyük veride, “neden’i değil, “ne’yi bilmenin önemli olduğunu söyleyen Mayer-Schönberger ve Cukier (2013:59), öngörü ve tahmin için “nedensel ilişkiye” niçin gereksinim duyulmadığını, 2007 yılında ABD’de grip vakalarının seyrini tespit etmeye çalışan uzmanlar hakkında söylediği şu sözlerle ortaya koymaktadır:

“Bilmiyorlardı ve bunu umursamayan bir sistem tasarladılar. Sistemlerinin yaptığı tek şey, bazı arama sorgularının sıklığı ile zaman ve mekân içinde gribin yayılması arasındaki korelasyonlara bakmaktı... Gribin nerede yayıldığını söyleyebiliyorlardı [...] Bunu olaydan bir ya da iki hafta sonra değil, gerçek zamana yakın bir şekilde söyleyebiliyorlardı” (2013:10).

Örneğin gerçek zamanlı konum bilgisi gönderen şehiriçi otobüslerdeki GPS’ler, ölçüm ile verinin kaydedilme yeri arasındaki mesafeyi ortadan kaldırmaktadır. Gerçek zamanlı bilgiler, çoklu unsurların planlanması sürecinde etkili biçimde kullanılmaktadır. Burada

sürekli ve dinamik veri akışı ile kesintili ve statik veri akışı arasındaki yöntembilimsel farkı görmekteyiz. Eş zamanlı olmasa da “olay-veri-analiz-öngörü-eylem” zincirindeki zaman aralığı gittikçe daralmaktadır. Normo fiziksel evrende daralan veya ortadan kalkan zamansal fark, önemli bir ontolojik soruna, veribilimsel bir çözüm önerisi doğurabilir. Şöyle ki parçacık fiziğinin standart yorumuna göre mikro evrendeki olgular, ancak gözlemediği müddetçe vardır. Gözlenmediği müddetçe mikro evren hakkında, klasik fiziksel anlamda konuşmak mümkün değildir. Yüzyılın en önemli fizik tartışması olan Bohr-Einstein uzlaşmazlığının temelindeki tartışma, bu konudaki şu soruya verilen iki farklı yanıtta özetlenebilmektedir:

- *Bakmadığım anda Ay, orada öylece var olmaya devam eder mi?*

Çoğumuzun sağduyusuna uygun biçimde Einstein, realist bir tutumla “Baksak da bakmasak da Ay orada öylece durur”, Bohr ise “Bakmadığın şey hakkında hiçbir şey söyleyemezsin!” diye yanıt vermişlerdi. Çünkü Bohr’un önderliğinde geliştirilen standart yoruma göre “gözlemden bağımsız nesnel gerçeklik” iddiasında bulunmak, kanıtlanabilir değildir (Işıklı 2012:159). Atom altı parçacıkların davranışlarını inceleyen ve nedensellik ilkesini ihlâl eden kuantum fiziği, Bohr’un yorumu doğrultusunda gelişmeye devam etmektedir. Nedenselliği paranteze alarak korelasyonel ilişkileri arayan hesaplamalı düşünme yöntemi, “Var olmaya devam eder mi?” gibi bir soruya, gözlemin sürekliliğini temin ederek bir adhoc çözüm sunmuş bir paradigma olarak değerlendirilmelidir.

h) Kesinsizlik (uncertainty), tedirgin etmeyen ölçüm ve gözlemin imkansızlığı: 1926 yılında Werner Heisenberg tarafından keşfedilen ve kendi adıyla anılan kesinsizlik ilkesi, bir galat-ı meşhur olarak belirlenimsizciliği önerdiği zannedilir. Ne var ki kesinsizlik ontolojik olarak belirlenimci, epistemolojik olarak bulanıklaştırıcıdır (Bkz. Heisenberg 2000:20). Ölçüm değerinin net olmayışı, öngöründe epistemolojik bir sapma veya bulanıklığa düşmemize yol açar (belirsizlik). Kesinsizlik, ölçüm süreciyle ilgilidir; olgunun kesin ve net ölçülemeyeşine gönderme yapar. Burada gözlemci ve gözlem aletlerinin bir yetkinsizliği göz ardı ediliyor değildir. Kuramsal, mantıksal ve uygulamısal sınırlar söz konusudur. Örneğin ışık hızı ve fotonun kütleli bir cisim gibi davranması, kesinsizliğin bilimsel dayanaklarından ikisidir. Gözlem veya verinin model tarafından işlenmesi, verinin veya sistemin süperpoze durumlarını çökertir. Kuantum bilişim kuramı, dinleme etkinliğini dahi, çoklu değerlerin yitirilmesine, değerlerin tekilleşmesine neden olan dışsal bir müdahale olarak tanımlar. Çünkü sistemin bileşenlerini tedirgin etmeyen bir gözlem gerçekleştirilemez. Bir kuantal sistem, gözlem yapılmadığı müddetçe süperpoze durumda kalır, bir sonraki aşamaya evrilmez. Kuantal sistemlerin devinimi ancak bilinçli gözlemcinin katılımıyla mümkündür.

İranlı Firdevsi, “Bilge kişi, güçlü kişidir” diye yazmış, Francis Bacon ve Thomas Hobbes ise “Bilgi güçtür” diye ilan etmiş, bilgi ile veriyi ilişkilendiren Lord Kelvin ise “Ölçmek bilmektir” demişti.⁴ Bilginin ölçüme indirildiği bu pozitivist yaklaşımın hakimiyetinin sarsılması ancak geçen yüzyılda, yaklaşık 90 yıl önce mümkün olabildi. Heisenberg’in kesinlik kesinsizlik ilkesinin kusursuz ve kesin ölçüm idealini ebediyen yok ettiğini söylemeliyiz. Einstein gibi dehaların bilimi, ondan arındırmaya çalıştıkları kesinsizlik ve dağınıklık, yine de bir sorun olmayabilir yahut da onunla başa çıkmanın bir yolu geliştirilebilir, böylece bilim için bir “kusur” olmaktan çıkarılabilir. Sadece az miktarda ve analog verilere sahip olduğumuz son on beş yıla girmeden önceki küçük veri çağında, kesinliğe ayrıcalık tanımak pragmatik açıdan doğru olabilirdi fakat artık elimizde büyük veri gibi bir maden vardır. Ondaki gevşekliği tercih edip olguların zaman içindeki evrim trendlerini görmek, daha akıllıcadır.

Mayer-Schönberger’e göre büyük veri çağında “netlik ve belirlilik talep ettiğimiz alanlarda bulanıklığa ve ikileme tolerans gösterebiliriz” (2013:55). Uzmanların bizzat gözlem ve ölçüm yoluyla elde ettikleri görece az sayıdaki verideki küçük bir kesinsizlik, kaos teorisinde serimlendiği gibi çıktılarda devasa bir sapmaya yol açabilir (Click 2000:XI). Böylece Çin’de kanat çırpın bir kelebek, Amerika sahillerinde fırtınaya yol açabilir, demek mümkün olur. Fakat büyük verideki kesinlik, verinin dijital olması, sayıca gerçekten çok olması, bir yetkinsizlik unsuruna dönüşen gözlem aletlerine gereksinmemesi gibi nedenlerden ötürü, kesinsiz verilerin yol açacağı sapmayı minimize etmeyi mümkün kılar.

1) Öngörü yerine tahmin, yasa yerine trend: Kesinliğe karşı kesinsizlik, öngörüye karşı tahmin, yasaya karşı trend hem büyük verinin başımıza açabileceği tehditlerle başa çıkabilme hem de pozitivist yöntemleri uygulamakta zorlanan beşeri bilimlere destek sunabilme araçlarından bazılarıdır. Gerçekten de aldığımız kararların arkasındaki nedenleri açıklayamadığımız, sabit veri-güdümlü bir tahminler evrenine doğru yol almaktayız. Büyük verinin öngörü (prediction) kabiliyetini sorgulayan Mayer-Schönberger ve Cukier şöyle yazmaktadır:

“Büyük veri, özünde kestirimlerle ilgilidir... Bir bilgisayara insanlar gibi ‘düşünmesini öğretmek’ ile ilgili değildir. Onun yerine, olasılıklar çıkarmak amacıyla çok fazla miktarda veriye matematik uygulamakla ilgilidir” (2013:19).

Burada dikkat edilmesi gereken nokta, büyük veri temelli bir öngörü modelinin iyi performans sağlamanın nedeninin belirlenimciliğe ve genel yasalara dayalı klasik öngörü sisteminden kat ve kat daha fazla sayıda veriyi hesaba katıyor. Modeller, çok sayıda veriyle

⁴ Mitolojik ve dini metinleri paranteze alırsak bilgi ile güç ve iktidar arasındaki ilişki kurma girişimini ilkin MÖ 350’li yıllarda Sokrates’in “Hiç kimse bilerek kötülük yapmaz” biçimindeki etik argümanda, daha sonra 1000’li yıllarda Firdevsi’nin Şehname’sinde, 1597’de Francis Bacon’un Kutsal Düşünceler adlı eserinde, 1651’de Hobbes’un Leviathan adlı eserinde görmekteyiz.

beslenirler, bu yüzden etki alanları daha kapsamlı, tahminleri daha isabetlidir. Daha fazla veriyle beslendikçe hangi veriye bakmaları gerektiği konusunda bir sezgiye ve zaman için kendilerini yeniden düzenlemeye daha elverişlidirler.

Geleceği, analizler yoluyla sezgisel olarak öngörmek istiyorsak büyük resme bakmalıyız. Fakat burada küçük bir sorun vardır: Resmin tamamını görmek, klasik-normo yaklaşımla mümkün değildir. Normo yaklaşım, insani duyu yeteneklerinin erişim sınırları içinde kalan büyüklüklerdir. Çok büyük (makro) veya çok küçük (mikro), normo (insani duyum eşiği) ölçeğin sınırlarını aşar. Klasik işlem ve hesaplama yeteneklerimizin normo olduğunu söyleyebiliriz. Bugün için veri, biraz büyük yani ortadır fakat yarın gerçekten de çok büyük olabilir. İşte o zaman tekrar veri içinde yüzen anlamlı sinyal adacıklarını bulacak yeni analitik motorlara ihtiyacımız olabilir. Salt verinin büyüklüğü içinde her türlü sinyal yitip gidebilir; büyüklük, sinyali gölgede bırakabilir. Böyle bir durumda sinyali bulmak, samanlıkta iğne aramak gibi olabilir, diye uyararı Minelli, bulanıklığı yeni bir öngörü imkânına dönüştürebileceğimizi söyler:

“Gelecekte eğer Büyük Veri okyanusundaki trendleri yeterince erken tanımlayabilirsek bunu, erken uyarı sinyaliyle önceden doğru eylemde bulunmak, ve yüzen çöp adaları gibi, sıkıntı içinde bulunduğumuzu gösteren sinyal birikimlerini önlemek için kullanabiliriz” (Minelli vd. 2013:107).

Sinyal, salt veri deryasında döngüsel olabilir ve uzun bir zaman dilimi boyunca gözlemlendiğinde tespit edilebilir. Örneğin birçok ekonomik döngünün, belirgin hale gelmesi, onu tespit edebilmemiz için en az on yıl gerekir. Bu yüzden, eğer verinin yalnızca, örneğin üç yılına bakarsak sinyali göremeyebilir veya yanlış yorumlayabiliriz. Bazı sinyaller, karmaşıktır ve o sinyali tanımlamak, ilişkili birkaç unsurun anlaşılmasını gerektirebilir. Bu aşamada büyük veri görüntüleme (visualation) araçları, kullanışsız veri içinde yakalamayı kolaylaştırmak için oyuncağa dahil olur. Farklı analitik araçlara ihtiyaç vardır. Minelli'ye göre analitiğin yalnızca tek bir yönüne odaklanan kurumlar, doğru öngörü ve tavsiyeler oluşturmada başarısız olacaktırlar. Analitik araçları uygun bir sentezle bir arada kullanılmalıdır.

Büyük veriyi ellerinde tutarak olanların tasvirinden olacak olanların öngörüsünü sunan, klasik olmayan yapılar gelecek tahmin sistemleri geliştirilmiştir. Bunlar kişisel, kurumsal veya toplumsal eğilimleri yakalayabilir, kişisel tercih ve davranışları öngörebilir, toplumsal olaylar henüz meydana gelmeden haber verebilir. Buradaki öngörünün, gelecekteki bir olayın yer ve zamanını kesin olarak söylemek anlamındaki doğa bilimsel öngörü (prediction) ile aynı anlamda ve işlevde olmadığı açıktır. 20 Mart 2015 yılında gerçekleşecek bir sonraki tam Güneş tutulmasının saat kaçta başlayacağını ve hangi enlem-boylamlardan gözlemlenebileceğini, kaç dakika süreceğini bilmekten farklıdır. Bu yüzden veri tabanlı öngörüler, içgörü (insight) olarak adlandırmak doğrudur. İçgörüyü de kişisel tahminden, eksik veriye dayalı yaklaşıksal akıl yürütme ürünü olan olasılıklı bilgiden ayırt etmek

gerekir. Kişisel tahminin içinde sezgi de vardır. İnsani düşünme biçimi, algoritmik olmayabilir. Bu da insani düşünmenin, hesaplanamaz veya varlığı tespit edilemez unsurlar içerdiğini söylemenin başka bir yoludur. Algoritmik tahmin sistemleri, öngörü değil, içgörü veya tahmin sunar. Öngörü, nedensellik ve yasal işleyiş; içgörü veya tahmin ise trend gerektirir.

i) Post yapısalcılık ve merkezsizlik: Büyük veri tabanları tek bir yerde (merkez) değildir, genellikle birden fazla sabit disk ve bilgisayar arasında dağıtılmıştır. Güvenirlik ve hız ihtiyacına cevap verebilmek için bir veri, farklı disklerde saklanır. Bu durum, klasik ve modern şehirlerin tek bir merkez etrafında kurulmasına mukabil, postmodern metropollerin birden fazla merkeze sahip olması, eski merkezlerinin artık tek bir merkez olmaktan çıkmasını andırır. Yapısökümcü filozof Jacques Derrida'nın da belirttiği gibi, klasik veya modern düşünce "sistemleri" de belli bir merkezi kavram etrafında örülmüştür. Derrida'ya göre, sistem düşüncesi belli merkezi kavramı önvarsayar; sistemler ise ikili (binary) karşıtlıklar mantığına göre kurulur. Karşıtlıklar mantığı yapısaldır ve merkez-çevre, iç-dış, iyi-kötü, verimli-verimsiz, büyük-küçük gibi karşıtlıklara dayanır ve karşıtlıklar yaratır (Derrida 2005:161). Yapısökümcü yaklaşımda ise ne tek bir hâkim sistem vardır ne de sistem vardır. Yalnızca pratik amaçlar için inşa edilmiş geçici modeller vardır. Ve modeller birbirini dışlamazlar çünkü bir modelin inşası için öteki modellerin dışlanması zorunlu değildir. Bir modeller çokluğundan, demokratik bir düşünce çeşitliliğinden bahsedilebilir. Büyük veriyi yapısal bir yaklaşımla kavramak zordur. "Tahminlere göre bütün dijital verinin yaklaşık %5'i 'yapısal'dır" (Schönberger ve Cukier 2013:55). Geleneksel veri tabanına tam uygun olan yapısal dijital verideki dağınıklık kabul edilmezse geri kalan %95 verinin belirsiz olduğunun kabul edilmesi gerekecektir.

j) Nesnellik ve doğruluk (accuracy) tasarımının düzeltilmesi: Nesnellik; bilginin nesneye bağımlılığı, öznenin bilgi edinme sürecinde kişisel ve sınırlı önyargılarını araştırma dışında tutulması, bilginin nesneye tam sadakatle oluşturulması, nesneyi aşkın bir bilginin nesneye atfedilmemesi gibi anlamlar içerir. Nesnelliğin, uzay – zamandaki realite anlamına gelen bir yönü de vardır ki bu neo-pozitivist bilim felsefecileri tarafından bilimselliğin olmazsa olmaz bir koşulu olarak önerilmiştir. Duyulabilir (sensible) nesneyle ilgili olmayan önermeler, duyu sınırlarını aşan, belki de duyulurüstü (transendantal) olan, bu yüzden de bilim alanından kapı-dışarı edilmesi gereken bir metafizik olarak tukaka edilmiştir. Bu bilim felsefesine göre bilimsel araştırmanın, elle tutulur, gözle görülür nesnelere dayanması zorunludur.

Türkçe'de kimi zaman hakikat (truht) yerine de kullanılan doğruluk (accuracy), bağlamsız genel yargıların, kapsadığı olgular kümesindeki tekil bir nesne için geçerli olduğunun gösterilmesiyle elde edilir. Doğruluk, bilgide dile getirilen yargının, yargıda içerilen nesnelere tam uygunluğudur ki kısaca "bilginin nesnesine uygunluğu" olarak tanımlanır. Tümevarımsal yöntemin nihai evresinde tescillenen önermeler, genel geçerlilik kazanırlar ve genel yasa ya da bilimsel yasa adını alırlar. Neo-pozitivizme göre deney, gözlem ve

matematikte doğrulanabilir olmayan hiçbir hipotez, “bilimsel” diye kabul edilemez. Bu görüşte açıktır ki bilimi tekeline alma, onu sahiplenme ve sınır çizme eğilimi güçlüdür.

Latour’a (2009) göre sosyoloji, “nicel bilim olma amacı tarafından baskı altında tutulmuş” (Akt. Boyd 2012:666), fakat bu amaca, sosyal alanda niceleştirilebilir ile nicelleştirilemez bilgi arasında bir çizgi olduğundan, asla ulaşamamıştır. Boyd ve Crawford’a göre ise nesnellik ve doğruluk beklentisi yanlış yönlendirmektedir. Sosyal bilimlerin araştırma yöntemleri hakkında yürütülen uzun soluklu tartışmada ortaya konan ayrımlar üzerinde, şöyle yeniden düşünülmelidir:

“Büyük Veri, beşeri disiplinlere niceliksel bilim konumu ve nesnel metod iddia edecekleri yeni bir yol önerir. Bu, çok daha fazla sosyal alanı nicelleştirir. Gerçekte, Büyük Veri çalışmasıyla birlikte öznel, varlığı hala sürdürür [...] özellikle de sosyal medya sitelerinden gelen mesajlar düşünüldüğünde. Fakat hâlâ nitel araştırmaların hikâye yorumlama, nicel araştırmaların ise olgu üretme işi olduğu hatalı inancı devam eder” (Boyd 2012:667).

Nesnellik kavramı, bilim felsefesi ve bilimsel yöntem hakkındaki erken tartışmalar için merkezi bir soru olmuştur. Bilim, nesnelliğin elde edilmesine çalışır. Ne var ki nesnellik bir iddia olarak, “özne tarafından yapılmak ve öznenin gözlem ve tercihlerine dayanmak zorundadır” (Boyd 2012:667). Bu görüşe göre bütün araştırmacılar veriyi yorumlarlar. Veri, enformasyon ve bilgiye dönüştürülmeden önce, ilk başta veri olarak tasarlanmaya ihtiyaç duyar. Bu yüzden verinin bu tasarımı süreci, öznel olan yorumlayıcı bir zemin içerir. Herbir uzmanlık alanı, veriyi nasıl tasarlayacaklarına dair belli norm veya standartlara sahiptir. Leonelli yalnızca verinin değil, enformasyonun bile hala yorum ve senteze karşı fazlaca verimli olduğu görüşünü paylaşır (Leonelli 2014).

Boyd ve Crawford’ın tespit ettiği gibi, büyük veri uzmanları da yaptıkları işin olgusal olduğunu yani yorumlama işi olmadığını iddia etme eğilimindedirler. Bir model, matematiksel olarak anlamlı gelebilir, bir deney geçerli görünebilir fakat bir araştırmacı olarak ne anlama geldiğini anlamaya giriştiğimizde, modeli veya deneyi yorumlamaya başlamış oluruz. Bu, her şeyi istediğimiz gibi yorumlayabiliriz, yorumlamanın geçerliliğin nesnel ölçütleri yoktur anlamına gelmez. “[Y]alnızca hiçbir sayı yansız değildir, anlamına gelir” (Boyd 2012:667). Bollier’e göre diğer tüm veriler gibi “büyük veri kendi kendini açıklayamaz [...] Veriyi yorumlayacak özel yöntemler, her türlü felsefi tartışmaya çoktan açıktır” (2010:13). Sosyal medya verilerinin analizindeki veri temizleme süreci, hangi nitelik ve değişkenlerin hesaba katılacağı, hangilerinin göz ardı edileceği hakkında karar verme sürecidir. Bu süreç, doğası gereği belirlenimci olamaz, araştırmacının aradığı bağıntıları bulmaya yönelik öznel ve seçmeci karakterli bir süreçtir.

4. BÜYÜK VERİNİN ELEŞTİREL ÇÖZÜMLEMESİ

Bütün bunların yanında büyük veri, her şeye çare bir ilaç, bir hokus pokus değildir. Onun da dezavantajları, başa çıkılması gereken sorunları, zayıflıkları ve en önemlisi yan etkileri vardır. Siegfried'in tespit ettiği gibi, devasa veri kümelerinin baştan çıkarıcı olduğu kabul edilmelidir (Siegfried 2013b). Bu yüzden eleştirel yaklaşımla incelenmesi gereken birtakım temel varsayımlar üzerine inşa edildiğini fark etmemiz gerekir (Boyd ve Crawford 2012:665). Bu yüzden genetik kodlar, sosyal medya etkileşimleri, hastane kayıtları, telefon logları, hükümet kayıtları ve insanlar tarafından bırakılan diğer dijital izler gibi büyük veri kaynaklarının analizi yapılmalı; analizlerin potansiyel fayda ve tehditleri tartışılmalıdır. Elbette bu süreçteki etkinlik en genel anlamıyla felsefi, özelde ise kuramsal, etik, epistemolojik ve politik olacaktır. Kesinlikle bilimsel olmayacaktır. Tam da bu noktada, bazıları aşağıda sıralanan önemli sorunlar ortaya çıkmaktadır:

- i. Geniş çaplı veri araştırmaları, daha iyi araçlar, hizmetler ve mallar üretmeye yardım edecek mi?
- ii. Büyük veriye mahremiyetin ihlâl ve istila eden pazarlama dalgasının eşlik etmesi önlenebilir mi?
- iii. Büyük veri analizleri, çevrim içi iletişimi veya siyasi hareketleri anlamaya, analiz etmeye ve öngörmeye yardımcı olabilir mi?
- iv. Büyük veri, örneğin protestocuların izini takip etmek ve ifade özgürlüğünü baskı altında tutmak için kullanılır mı?
- v. Geniş nitel veriler, insani iletişim ve kültür araştırmalarını nasıl dönüştürecek veya dar araştırma paletleri seçeneklerinde araştırmanın anlamı ne olacak? (Danah Boyd, 2012, s. 663).
- vi. Sosyal tabakalar ve kuşaklar arasında yeni türden bir ayrımı yol açabilir mi?
- vii. Bilgi ve gündelik yaşama dair kritik değişimlere yol açar mı?

İnsan hakları, demokrasi ve eleştirel kavramlara duyarlı birçok araştırmacıya göre, diğer sosyo-teknik olgular gibi büyük veri de ütöpik ve disütöpik söylemi harekete geçirmektedir. Bir yandan büyük veri, terörizm, güvenlik, kayıt dışı ekonomi, ekonomik krizler, kanser araştırmaları gibi çok çeşitli sorunların çözümüne yönelik yeni içgörüler sunan bir teknik olarak görülmektedir. Diğer yandan Büyük Birader'in popüler kıldığı bir baş belası olarak büyük veri, kişisel olanı (mahremiyeti) görünür kılan, sivil özgürlükleri kısıtlayan, resmi ve kurumsal kontrolü artıran bir manifesto gibi görülmektedir. Sonuçta büyük verinin kötü kullanımı, büyük tehdit ve tehlikeleri içermektedir ve bu yüzden büyük verinin altını oyan etik içerimler hakkında tartışmamız gereken birçok sorun bizi beklemektedir.

a) Dijital uzmanların etik yoksunluğu: Beşeri meselelere dair veriler, kaçınılmaz olarak “özne[l]lik (su[b]je[ctivity])” ve “özel-lik (priva[te/cy])” taşır. Özellik, özneyi refere eder; “belli bir özneye has olan” anlamına gelir. Sorun şu ki büyük veri araştırmalarında bu özel olma durumundan türetilen mahremiyeti koruyan normlar henüz geliştirilmiş ya da

etkinleştirilmiş değildir. Araştırmalarını dijital kaynaklar üzerinden yürüten uzmanlar için henüz dijital telif hakları yasaları bilinir değildir veya üzerinde uzlaşmış dijital etik ilkelerden henüz bahsetmek güçtür. Etik ve tüze bağlamında sık tartışılan şu sorulara bakalım:

- i. Uzmanlar, herkese açık bile olsa kişisel paylaşımları, kişisel bilgilendirme ve[ya] izin olmaksızın analiz etme hakkına sahip midir?
- ii. Gerçekleştirilen araştırma projesi, başkalarına zarar verebilir mi? Bunun sorumluluğu kime aittir?

Büyük veri çalışmalarında hakikat, kontrol ve güçle ilgili kayda değer sorular olduğu görülmektedir. Çevrimiçi bir paylaşımın herkese açık oluşu yahut da özel bir verinin erişilebilir oluşu, onu kullanmayı etik kılar mı? Boyd ve Crawford'a göre "yalnızca erişilebilir oluşu, onu etik kılmaz." Bu konuda kişisel haklara duyarlı bir tutumla araştırmacıların kendilerine ve diğer meslektaşlarına büyük veri analizi ve yayınları konusunda etik sorular sormaları gerekir fakat araştırmacılar "çok nadiren" kendilerini kullanıcı yerine koyarlar. Çünkü:

"Büyük Veri araştırmacıları, kamusalda olmakla (being in public) ile kamusal olmak (being public) arasındaki dikkate değer fark olduğunu çok nadiren bilirler" (Boyd ve Crawford 2012:673).

Birçok araştırmacı için kamudaki veriler, herkesin kullanımına açık kamusal veridir. Halbuki aracınızı herkesin görebileceği bir alana bırakmanız, herkesin o aracı kullanabileceği anlamına gelmez. Öte yandan birçok çevrimiçi kullanıcı, kişisel verilerinin gelecekte kullanılıp kullanılmayacağı konusunda, bir fikre sahip değildir. Belki de birçok kullanıcı, kendi ürettiği ya da kendine ait verilerin başka amaçlar için kullanılmasına izin vermeyecektir. Elbette kullanıcıların, kişisel verilerinin veya paylaşımlarının çoklu kullanımlarının, kâr ve diğer kazançlarının baştan farkında olması gerekmez. Fakat onu kullananların bu konuda bir norm veya etik engellerin olup olmadığını sorgulamaya yönelik zorunlu bir eğilimlerinin olmasını beklemeliyiz. Çünkü veriyi bilgi ve kapitale dönüştüren bu uzmanlar, verinin ekonomik değer taşıdığı, bu değer ne karşılığı oluştuğunu bilmektedirler. Bilinçli ve girişken bazı büyük veri sensörlerini kullanan tüketicilerin, izinsiz veri kullanımını mahkeme gündemine taşımalarını bekleyebiliriz.

b) Veri tabanlı dolandırıcılık: Leonelli'ye göre başka sorunlarla birlikte veri-ağırlıklı bilim, "kötü veriyi, delillerin dijital manipülasyonunu, intihal ve yeni dolandırıcılık fırsatlarını" bünyesinde barındırır. Bu yüzden verinin kanıt değeri kazandığı bağlamı ve etki alanını (domain) anlamak hayattır (Leonelli 2014).

Dolandırıcılık, kişisel kazanç sağlamak veya başkasına zarar vermek için kasten başkalarını aldatmak olarak tanımlanmaktadır (Minelli vd. 2013:34). Kişisel bilgiye dönüşen dijital veri deposu da denilen sosyal medya hesapları ve cep telefonları, dolandırıcılığın sınırlarını genişletmektedir. Çoğu zaman uzmanların, iletişim araçlarında kişisel bilgilerin

paylaşılmaması gerektiği uyarılarının, özellikle sosyal medyada, çevrimiçi işlemlerde kullanıcıların kimliklerinin gerçekliğini tespit etmede kullanılan çok sayıda kişisel veriyi paylaşmaya devam ettiklerinden, işe yaramadığı gözlenmektedir. Herkese görünür olan açık profillerle de kişisel bilgiler muhtemelen daha çok teşhir edildiğini düşünmek için şimdiden önemli araştırmalar yapılmış durumdadır.⁵

İnsan hak ve özgürlükleri konusunda bilgilenmek, bir insan hakkı olarak tanınmıştır ve buna bağlı olarak da Türkiye, Avrupa ve ABD gibi demokratik ülkelerde insan hakları ve vatandaşlık eğitimi, yerine getirilmesi gereken zorunlu görevlerden birinin ifası olarak kabul edilmektedir. Peki, kişisel verileri saklayan bilişim-iletişim-teknoloji sistemleri konusunda eğitim almak, kişisel bir hak meselesi olarak kabul edilemez mi? Ne de olsa hakları bilmeden aktif vatandaş olunamayacağı gibi, teknolojiyi kullanmadan toplumsal katılım sağlanamamaktadır. Bu yüzden mesele tartışmaya değerdir.

c) Mahremiyetin ihlâli, hukuk, olasılık ve cezalandırma: İlk otomatik işlemci 1890'da, ilk ilişki veri tabanları ise ondan 70 yıl sonra, yani 1960'larda geliştirildi. Kişiselleştirme gibi özellikler, daha ilgili bilgilere hızlı erişime izin vermekte fakat birçok uzman, büyük verinin cevaplandırılması zor etik sorularına kafa yormayı tercih etmektedir. Üzerinde tartışılan konulardan bazıları aşağıda sıralanmaktadır.

- i. **Geleneksel etik kavramlarının yetersiz kalması:** Büyük veri olgusunun yol açtığı tehditleri, geleneksel etik meseleler ve yaklaşımlar çerçevesinde tartışmak verimsizdir. Etiğin kullanılagelen kavramları, "modern" dönemin meselelerini çözmek için üretilmiş araçlardır. Dijital çağın yeni türden etik meseleleri ve onları kuşatacak yeni kavramlara ihtiyaç vardır. Örneğin mobese kameraları gibi kamusal veri toplayıcıları tartışmak, kişinin görüntüsünün şahsiliği ve mahremiyeti bağlamında tartışmak, yeni teknik olgu, çıkar ve zararlar bağlamında mümkün olabilir.
- ii. **Sürekli izle[n]mek ve panoptikon:** İnternet, mahremiyeti tehlikeye attıysa büyük veri onu iki kere tehlikeye atmıştır. Büyük veri başlangıçta kişisel bilgi içermez. Çünkü bilginin değeri, birincil amaca yönelik değildir; ikincil kullanımlara bağlıdır. Kişisel verilerin toplanması, saklanması, başka amaçlarla kullanılması (dijital işleme) izne bağlı olmalı mıdır? Güvenlik gibi toplumsal faydalar açısından nasıl karar verilmelidir? Kamusal alanda ve kişisel alanda her yanı sarmış durumdaki teknolojik cihazlar, tüm hareketlerimizin izini kaydetmektedir. Dijital cihazları kullanmak veya kamusal alana çıkmak, buralarda sürekli izlendiğimiz anlamına gelmekten nasıl kurtarılabilir?

⁵ Örneğin Bkz. Javelins'in 2012 Kimlik Dolandırıcılığı Raporu: Sosyal Medya ve Yeni Dolandırıcılığın Sınırlarının Mobil Cihazlar Tarafından Çizilmesi Araştırması Raporu.

- iii. **Masumiyet karinesinin ihlâli, suç ve ceza, özgür irade:** Toplumsal tehdit ve tehlikleri bertaraf edebilmek amacıyla gerçekleştirilen öngörücü ya da önleyici polisiyenin veriyi kullanırken mahremiyeti ve kişisel hakları ihlal edip etmediği konusu, hakkaniyet, adalet ve özgür irade fikirleri bağlamında tartışılmalıdır. Henüz gerçekleşmemiş eylemler için vatandaşların hak ve özgürlüklerinin önleyici polisiye etkinliği gerekçesiyle kısıtlanması başka tahmin edilemez sorunlara yol açabilir. Kişisel mahremiyetin korunması, özgür irade, seçme özgürlüğü... Bir kişinin suç işleme olasılığının çok yüksek olması ne anlama gelecek? Potansiyel suçlu veya muhtemel zanlı gibi kavramlar hukuk literatüründe nasıl tanımlanacak? Bunların yasal veya toplumsal yaptırım olarak karşılığın nedir? Suç işlemeye yatkın olanlar tutuklanacak mı? Henüz suç, fiilen gerçekleşmediği halde suçlu veya zanlı muamelesi mi yapılacak? Zan, suçla birlikte ortaya çıkan bir durum iken büyük veri, zannı, suçtan önce mi ya da suçla eş zamanlı olarak mı ortaya çıkaracak? Potansiyel suçluyu, henüz suç işlemeyen tutukladığımızda, tıpkı istatistiklerin nedensiz biçimde kendilerini yenilemeleri gibi, suç da ceza da nedensiz (gerekçesiz) mi gerçekleşecek? İstatistiksel olarak suç işlemeye eğilimi kuvvetli biri, bir suç işlediğinde “özgür iradesiyle suç işlemiş” mi sayılacak yoksa “istatistiksel zorlamayla suç işlemiş” mi sayılacaktır? Bir sonraki kurban bayramında kaç kişinin trafik kazası kurbanı olacağını tahmin eden büyük veri analiz sistemi, o bayramda ölecekler için felaket tellallığı mı, felakete davetiye mi yoksa uyarı mı yapmış sayılacak? Peki, her şeye rağmen istatistiki eğilim veya trendlerin gerçekleşmesi, yaşam ve ölüm açısından ne anlama gelecek?
- iv. **Büyük veri diktatörlüğü:** Büyük veriye sahip büyük şirketler, veriyi nasıl kullanacaklar? Güçlünün elinde bir enstrümana dönüşüp yanlış kullanılırsa ne yapılabilir? Burada verinin çirkin suçlara ortaklık etmesi söz konusudur. Büyük veri baronlarının idare edilmesi için ulusal ve uluslar arası yasa ve müeyyidelerin oluşturulmaya başlanması ivedi bir sorundur.
- v. **Rızanın yokluğu veya büyük verinin karanlık yüzü:** İnternet logları veya cep telefonu görüşmelerinin kayıtlarına dayanarak kişiye özel (kişiselleştirilmiş) reklamlar, öneriler, bildirimler, simge ve mesajların gönderilmesi, hem kişilerin ıvır zıvırla meşgul edilmesi hem de izni olmadan kişisel zamanın ve bilginin istismarı olarak değerlendirilebilir. Bu, büyük veride rızanın yokluğu sorununu gündeme getirir. Şirket ve kurumların, ticaret ve alışveriş gibi gündelik basit işlerde dahi, kimlik no, cep telefonu, adres gibi kişisel bilgileri talep etmesi kuşku uyandırmaktadır. İster devlet isterse tüzel kişilerin olsun, kaydettikleri kişisel verileri nasıl muhafaza ettiği tartışılmalıdır.

Belki de bazılarının önerdiği gibi gizliliği muhafaza etmek için şu üç tekniği etkin kılmak gerekecektir: 1. Kişisel uyarı ve rıza, 2. Vazgeçme, 3. Anonimleştirme... Mahremiyetin ihlalinden kurtulmak için veriyi oluşturma ve veriyi kullanma sorumluluğunun

yaygınlaştırılması, gelenek haline getirilmesi ve verinin yasal bir teminata kavuşturulmasını beklemeliyiz.

d) Veride büyüklük sorunu: Jiang Fan'a göre "Yüklü miktarda veri, hem fırsatlar hem de veri analizi güçlüklerini beraberinde getirmektedir" (Siegfried, 2013). Büyük verinin bilim için kötücül yönlerini sorgulayan Tom Siegfried'a göre, büyük veri analiz modelleri, araştırmacıları, kötü hatta berbat denilebilecek istatistiksel özellikler içeren ad hoc (geçici olarak eklenmiş) prosedürler oluşturmaya zorlayabilir. Bilim için büyük verinin niçin kötücül olduğunu tartışabiliriz. Bilim insanları artık sonsuz - süper bilgi okyanusunda yüzyorlar; veri nirvanasına dair vahşice rüyalarını gerçekleştirmeye başladılar.

Bazı uzmanlara göre, büyük veriyi kullanarak araştırma yapmak, denize olta atmaya benzer. Ne arandığı ve ne bulunacağını belirsizdir (Schönberger ve Cukier 2013:37). Boyd ve Crawford ise büyük veriyle birlikte beşeri bilimlerdeki mevcut yöntem sorunlarının, artık sorun olmaktan çıkmış olduğunu söylemek mümkün değildir. "Daha büyük veri", koşulsuzca "daha iyi veri" anlamına gelmez (2012:668). Crawford'un, büyük veri üreticilerinden biri olan Twitter'a dair şu değerlendirmesi, büyük veri temelli araştırmalardaki bazı esaslı sorunları sezmeye yardımcı olabilir:

Twitter herkesi temsil etmez; üstelik halk ile twitter kullanıcılarının eş anlamlı olduğunu varsaymak bir hatadır: Twitter kullanıcıları çok özel bir alt kümedir. Ne Twitter kullanan nüfus, küresel nüfusun temsilidir ne de hesap ile kullanıcıyı eşdeğer sayılabilir. Bazı hesapların çoklu kullanıcıları olduğu gibi bazı kullanıcıların çoklu hesabı vardır. Bazı kişiler asla hesap açmazlar, Twitter'a web aracılığıyla erişirler. Bazı hesaplar, doğrudan biriyle ilgili olmaksızın otomatik içerik üreten 'bot' hesaplardır. Üstelik 'aktif' hesap kavramı da sorundur. Bazı kullanıcılar Twitter'dan sıkça içerik paylaşırken diğerleri 'dinleyici olarak katılır' (Boyd ve Crawford 2009:532).

Haklı olarak büyük veri setlerinin, vatandaşların tümünü örneklediğini, büyük veriye dayalı tahminlerde bir şeyleri gözden kaçırmış olabileceğimizi önceden kabul etmemiz gerekir. Çünkü "büyük" veri ile "bütün" veri aynı değildir. Bir veri kümesinin örneklemi dikkate almadığımızda veri kümesinin ne denli büyük olduğunun bir anlamı kalmaz. Üstelik, verinin kaynağına dair bilgi kesin olmadıkça örneği veya örnekleme anlamlandırmak zordur. Örneğin Twitter Inc.'teki büyük veri, korumalı hesapları dışta bıraktığı için bütün tweetleri temsil etmez. Dahası açık hesapları temsil etmesi de zorunlu değildir. Bu yüzden Boyd ve Crawford şu uyarıyı yaparlar:

Twitter, veri madenciliğinin çok popüler bir kaynağı haline gelmiştir fakat twitter ile çalışma, ciddi yöntemsel tehlikelere sahiptir" (2012, s.669).

Bu yüzden dijital çağın hesaplamalı düşünme yöntemiyle gerçekleştirilen veri analizi, modelleme, bilgi – işlem süreci, kişiselleştirme, tahmin veya trendlerin değerini tayin ederken büyük verinin berbat yönlerini hesaba katmak gerekmektedir.

5. BÜYÜK VERİYE DAYALI YENİ BİLİM YAPMA TARZI

internet ve büyük veri, Savage ve Burrows tarafından deneysel sosyolojide bir kriz olarak tasvir edilmiştir. Açıktır ki büyük veri, hem doğa bilimlerini hem de sosyal bilimleri yeni tarzda bilim yapmaya hazırlamaktadır. Geleneksel laboratuvar veya literatür çalışmalarını önemseyen modern bilim insanları, yeni paradigmanın baskı veya etkisini hissetmeye çoktan başladılar. Öte yandan Siegfried'e göre "büyük verinin yükselişi, kuram ihtiyacını belirginleştirmektedir" (Siegfried, 2013).

a) Paradigma değişimi: Verinin nasıl kullanılacağına dair bir zihniyet değişimine ihtiyaç duyulduğu açıktır. Geneleksen veri işleme paradigmasına sahip veri analiz uzmanları, çoğu zaman binbir güçlükle toplanmış analog veri setlerini, tek bir amaç için kullanmaktadırlar. Verinin değeri de buna bağlı olarak "kağıt peçete" gibi, bir kez kullanılınca çöpe atılması gereken "statik" bir şey olur. Yeni bir hipotez için, hipotezi doğrulayıcı yeni veriler toplanmak gerekir. Fakat bilgi toplumunun dijital verisi, farklı amaçlara ışık tutan farklı matematiksel modeller tarafından çok kez işlenen, dinamik ve kıvrak oyun hamurları gibi çok katmanlı bir hammadDEDİR. Bir veri, farklı amaçlar için, farklı modeller tarafından tekrar tekrar kullanılabilir. Buradaki kritik ayırım şudur: Büyük veri, hipoteze yönelik üretilmez; logaritmik sistemin bilgi – işlem sürecinde doğal olarak üretilen veridir.

Modern uzmanlar ile büyük veri uzmanlarını, çalışma prensiplerini göz önünde bulundurarak "analog" ve "dijital" paradigma iki farklı gruba ayrılmaktadır. Geleneksel araştırma ve veri analiz yöntemlerini kullanan uzmanları, "analog paradigmaya saplanmak" biçiminde olumsuzlayıcı biçimde tasvir eder (2013:13). Ne var ki bu ayırım, tarihsel açıdan "bilimsel araştırma" denilen olguyu kuşatmakta yetersizdir. Olguların gözleminini sayılarla raporlama ve bilimsel araştırmanın neticelerini sayısal formlarda ilan etme, Galilei ile başlamış modern tarzda bilim yapma paradigmasıdır ve ondan önce de insanlar, bilim üretmişlerdir. Beşeri bilimler için yeni bir sayısallaştırmadan bahsedebiliriz. Böylece bir usamlama yöntemi olarak sosyal bilimler, bir düşünme becerisi olarak "dijital doğanlar"a özgü olan hesaplamalı düşünme, dijital paradigmayı çözümlmek ve genişletmek için önemlidir.

b) Hesaplamalı düşünme ve dijital beşeri bilimler: 16. yüzyılda başlayan doğayı niceleme yoluyla anlama çabası, 19. yüzyıla gelindiğinde bilimi tanımlayan en önemli nitelik olarak ön plana çıkmıştır. Bugün biraz yol aldık ve gerçekliğin daha fazla yönünü veri formatında ifade etmeye başladık. Bir olayı verileştirmek, çizelgelenebilir ve analiz edilebilir niceliksel formata dönüştürmek demektir. Verileştirme, bilgisayarlarla yapılan ilk işlem değildir. Bilgisayar, bir hesaplama teknolojisi olarak geliştirilmiştir, veri hazırlama, dijitalleştirme veya verileştirme aracı olarak değil. Dijitalleşme, analog bilgiyi bilgisayarların okuyabileceği hale getirmektir. Sayısallaştırma (digitalizing) denilmesinin nedeni, analog bilginin bilgisayarların kullanabileceği şekilde 0 ve 1 sayılarından oluşan ikili (binary) koda

dönüştürülüyor olmasıdır. Verileştirme; güneş altındaki her şeye dair bilgi toplama, bunu nicel hale getirmek için bir veri formatına dönüştürme girişimini ifade eder. Bir kişinin konumu, bir köprüdeki vibrasyon ve gerilim, bir motorun sesi gibi asla bilgi olarak düşünmediğimiz şeyleri bile verileştirmek... Bir seyyahın not tutması, dokunduğumuz nesnelere üzerindeki parmak izlerimiz, Tyco Brahe'nin astronomi gözlem kayıtları, Galilei'nin Pisa Kulesinde tuttuğu notlar, denizcilerin seyir defterleri verileştirme olarak kabul edilir. İlişki ve ayırım, Mayer-Schönberger'in tespit ettiği gibidir:

Sayısallaştırma, verileştirmeye büyük güç kazandırır ama onun yerini alamaz. Sayısallaştırma edimi [...] kendi başına verileştirme değildir (2013, s.90).

Analog dönemde yıllar alan bilgi toplama ve analiz etme işleri artık günler içinde ya da daha kısa bir zaman diliminde yapılabilmektedir. Dijitalize edilen bilgi, daha kolay ve daha ucuz saklanabilir, işlenebilir. İstatistikçi Jiangın Fan, bilimsel ilerlemenin her geçen gün veriye daha fazla bağlı ve daha fazla veri odaklı hale geldiğini ileri sürmektedir (Siegfried 2013b). Bu iddianın, diğer tüm dijital-veri analiz uzmanları tarafından onaylandığını söylemek mümkündür.

Daha önce tanımlanmamış bir yetenek veya bilim yapma tarzı olarak hesaplamalı düşünmeden bahsetmek ancak dijitalleşmeyle birlikte mümkün olabilmiştir. Hesaplamalı düşünme, maruz kaldıkları yoğun dijital uyarıcı ortamında yetişen bireylere özgü bir düşünme becerisi ve dijital uzmanların geliştirdiği bir bilim yapma tarzıdır. Bu, birden bire ortaya çıkmış bir hokus pokus değildir. Beşeri bilimlerde hesaplamalı düşünmeye yol açan hesaplama teknolojileri, ilk başlarda çoğu zaman, proje yürüten gerçek uzmanlara, araştırmalarında teknik bir destek olarak görülürdü. David Berry'nin yerinde ifade ettiği gibi, zaten beşeri bilimlerde, projelerin büyüklük ve karmaşıklık düzeylerinin artması, hesaplamalı teknikleri, usta araştırmacıların nazarında araştırma sürecinin ayrılmaz bir parçası haline gelmişti (Berry 2011:2). Bugün, hesaplama teknolojisinin gelişimini takip eden hesaplamalı düşünme, farklı evrelerden geçirmiş yaygın ve etkili bir yöntemdir. Hesaplamalı teknoloji ve düşünme biçimi, bugün beşeri bilimlerde farklı evrelerden geçerek ortaya çıkan sorunlar hakkında düşünmek için gerekli bir koşul haline gelmiştir.

Bazı uzmanlar tarafından dijitalleşmenin farklı evreleri tespit edilebilen bir süreç olduğunu belirtilmektedir. Örneğin Schnapp ve Presner, Digital Humanities Manifesto 2.0'de, dijital dönüşümün ikincisi evresine çoktan girdiğimizi ileri sürmektedir:

“Birinci dalga dijital beşeri bilim çalışmaları nicelikseldi, araştırmaya hareketlilik kazandırıyor, veri bankası gücünü tekrar ele geçiriyor, yapı dilbilimi otomatikleştiriyor, kritik dizilerdeki hiper kartları istifliyordu. İkinci dalga ise karakter olarak niteliksel, yorumlayıcı, deneyimsel ve duygusal idi. Beşeri bilimlerin hizmetinde olan dijital araç takımlarını saf yöntembilimsel güçle donattı:

karmaşıklığa dikkat etme, ortam özelleştirme, tarihsel bağlam, analitik kuşku, eleştiri ve yorum” (Schnapp ve Presner 2009).

Esasında birinci dijital dalga, 1990’ların sonu ile 2000’lerin başında geniş ölçekli dijitalleştirme projeleriyle ve teknik ekipmanların kurulmasıyla başlamıştır. Birinci dalga, dar anlamıyla sistem sınıflandırmaları, işaretleme, metin çözümleme, akademik editörlük vb. alanlarda yoğunlaşıyordu. Tümüyle yeni, disiplinler arası bir paradigma, yondeşik alanlar, hibrid metodolojiler ve hatta baskı kültürüyle sınırlandırılmayan ya da onlardan türetilmeyen yeni yayıncılık modellerini sunan ikincil dijital dalga, Presner ve diğerlerinin Dijital Beşeri Bilimler 2.0 diye adlandırdıkları yönde devam etmektedir. Birinci dalgada, dijital raporlama yoluyla beşeri bilim metin çalışmalarında alt yapı inşa etme ve metin işaretlemeyle sürecini içerirken ikinci dalga, arşivlerin kavramsal sınırlarını, örneğin dijital eser ve materyalleri, elektronik literatürü (e-lit), interaktif işlevleri, web tabanlı ürünleri vb. kapsayacak biçimde genişletmektedir (Mayer-Schönberger ve Cukier 2013:3).

Presner’e göre aynı anda iyi ve kötü niteliklere sahip olan dijital beşeri bilimler; son derece üreticidir, yeni üretim araçları ve çevreleri yaratır. Çeşitli dijital bağlamlarda canlı bilgiyle etkileşen yeni bir nesile sebebiyet vermişlerdir: “Dijital doğanlar” (Presner 2010). Şimdiye kadar bu çağ, “Z kuşağı” veya “dijital kuşak” olarak adlandırılmıştı fakat “dijital doğanlar”, esaslı bir ayırım ve vurgu taşımaktadır. Burada artık “dijitale ve dijital ile” doğanlardan bahsetmekteyiz.

Berry’e göre ise dijital çağdaki “hesaplamalı dönüşüm (computational turn)” aslında bu noktada artık üçüncü dalgayı ifade etmektedir. Dijital üçüncü dalga bir bilgisayar ortamında tutulan formaların hesaplamasallığını (computationality) vurgulamaya yoğunlaşır (Berry 2011). Burada dijital içeriklere, dijital verinin spesifik ortamı ışığında, ortam değişiminin bir epistemik değişim üreteceğini öngören bir yaklaşımla bakmayı salık verilmektedir. Lakatos’un sorunsallaştırdığı bazı kavramları, örneğin “salt beşeri bilim çalışması” kavramı, “normal” araştırmaları destekleyen fakat üzerine konuşulmayan varsayımları veya örtük sayıltıları, ontolojik temelleri açısından yeniden düşünmeye davet eder. Aslında Berry, üçüncü dalganın; beşeri araştırmalarda içselleştirilmiş yakın okuma, ölçütlerin oluşturulması, periyodikleştirme, liberal beşeri bilim vb. anomalileri görünür kılan bir yolu işaret ettiğini söylemek ister. Öyleyse bu noktada Berry’nin Mathew Fuller’dan aktardığı şu tespite katılmamız mümkündür:

“Bütün entelektüel çalışmalar, bir anlamda ‘yazılım araştırması’dır, yazılım tarafından sağlanan ortamda ve bağlamda yapılır” (Berry 2011:4).

Hesaplamalı düşünmeyi, kültürel bir olgu olarak analiz eden Wing, durumu anlamak için çocukların nasıl bilgisayar uzmanı gibi düşündüklerini gözümüzde canlandırmayı önerir (2006, s. 34). Bu öneriye, sosyolojik gözlemler eşlik eder: Fiziki çevre ile özel veya resmi tüm

kurum ve kuruluşların istatistik departmanlarında artık neredeyse tümüyle bilgisayar uzmanları çalıştırmaktadırlar. İstatistik bölümleri, üniversitelerin bilgisayar bilimleri fakültelerinde olmazsa olmazlar arasında sayılmaya çoktan başlandı. Özellikle makine öğrenme istatistiğini yadsınamaz biçimde dönüştüren Hesaplamalı Düşünme (HD)'nin, Wing'e göre iki temel teknik niteliği vardır: 1. Soyutlama (Abstraction), 2. Otomasyon (Automation) (2006:34). Otomasyondan ötürü, her ne kadar yapay zekâyı çağrıştırırsa da HD, logaritmik değildir; özünde bir kavramsallaştırma sürecidir, programlama değil; bir fikir edinme sürecidir, yaratılmış bir nesne veya teknoloji değildir. Belli bir uzman türüne has olmayan, herkes için olan tümüyle insani bir düşünme biçimidir. Modern toplumlarda iş yapabilmek için herkesin ihtiyaç duyduğu temel bir yetenektir, mekanik ezber değildir. Wing'in tespitleri ışığında bir beceri olarak hesaplamalı düşünmenin nitelikleri şöyle sıralayabiliriz (Bkz. Wing 2006:34):

- i. HD, rekürsif (kendini tekrarlayan, özyinelemeli) bir düşünmedir.
- ii. HD, nasıl çözebileceğimizi bildiğimiz zahiren zor görünen soruların, indirgeme, imbit etme, dönüştürme veya simülasyon yoluyla yeniden formüle edilmesidir.
- iii. HD, kolay izlenir kılmak için bir problemin farklı yönleriyle ilgili uygun bir tasarım veya model seçmektir.
- iv. HD, büyük karmaşık görevlerle mücadele etmek için bir soyutlama ve analiz kullanmadır.
- v. "Bu problem ne kadar zordur ve bunu en iyi nasıl çözebilirim?" türünden teorik bilgisayar bilimlerinin sorularına cevaplar verir.
- vi. HD, bir sistemin basitlik ve zerafet tasarımını yargılama, değerlendirmedir.
- vii. HD, boyutlu analizlerin bir genelleştirmesi olarak tip kontrolü yapmaktır.
- viii. HD, çoklu kullanıcıların katılımındaki şeyleri, modülize etmek ve sonraki kullanımlara katılımı tespit etmektir.
- ix. HD, problem çözme, sistem tasarımına ve bilgisayar bilimlerinin kavramlarıyla açıklanabilen insani davranışları anlama yaklaşımına sahip olmaktır.

Dijital teknolojiler, araştırma yapma biçimimizi etkilemekte ve değiştirmektedir. Her türlü araştırmanın, artan biçimde dijital teknoloji aracılığıyla yapıldığı gün yüzüne çıkmaktadır. David Berry gibi düşünürler, bu aracılığın, bir araştırma yürütmenin anlamını usulca değiştirmeye başladığını, bir araştırma programının önemini vurgulayan epistemolojiyi ve ontolojiyi etkilediğini iddia etmektedirler. Bugün, araştırma etkinliğinin bir parçası olarak dijital teknoloji erişimine sahip olmayan akademisyen çok enderdir. Bir akademisyenin kütüphanelere erişimi için kütüphane kataloglarına başvurması, bilgisayara kullanmadan araştırma yapması daha az rastlanan bir durumdur. Baker'ın (1996 ve 2011) araştırmalarına atıf yaparak "Kart indeksleri yavaş ve kesin biçimde ölüyor" diyen Berry'e göre, bir yandan dünya kütüphaneleri taranıp çevrim içine aktarılmakta, diğer yandan elektronik posta,

Google araştırmaları ve bibliyografik veri bankaları gittikçe hayati hale gelmektedir (Berry 2011:1).

Berry, beşeri bilimlerin bilgisayar teknolojileri ve hesaplamalı yaklaşıma doğru gidişatının yeni bir adlandırmayı gerektirdiğini düşünür, böylece “dijital beşeri bilimler” ifadesini önerir. “Dijital beşeri bilimler, dijital formların esnekliğini ve gerçekliğin dijital katmanları denilebilecek yeni bir tasarımlama ve dolayım ile çalışma yolunu hesaba katmaya çalışır” (Berry 2011:1). Buradaki anahtar olgu, somut kodlama imkanı bulunmayan hesaplama araçları için işlem yapacak hiçbir nesnenin olmamasıdır. Başka bir ifadeyle bir bilgisayar, gündelik yaşamımızın sürekli akışındaki her şeyi, bir algoritma kullanarak manipüle edilebilen gerçekliğin bir tasarımı olarak matematiksel gridlere dönüştürmeyi gerektirmektedir. Berry’nin de kabul edebileceği gibi, sayılar ve sayısal modeller, “kendinde gerçekliğin tasviri” olarak görülemez:

“Bu, gerçekliği eksiltici anlama yöntemi (episteme), yeni bilgiler ve gerçekliği kontrol yöntemleri (teknik) üretir. Onlar, dijital medya vasıtasıyla dijital beşeri bilimleri kendi sorunsalları olarak ciddiye almaya başlarlar” (Berry 2011:2).

Veri ağırlıklı bilimin ne tür bir felsefeye sahip olduğunu irdeleyen Sabina Loenelli, mutemel epistemik sakıncaları olsa da böylesi bir bilimin felsefi soruşturmalar için verimli bir alan olduğunu ileri sürer (Leonelli, 2014). Hesaplamalı düşünme biçimini, analitik veya kritik düşünme biçimlerini tasvir eder gibi belirtilerini ve kullanım alanlarını tespit eden Jenette M. Wing, önceki zihinsel yeteneklere bir yenisinin eklendiği şöyle ifade eder:

“Hesaplamalı düşünme, 21. yüzyılın ortalarına doğru dünyadaki herkes tarafından kullanılan temel bir yetenek olacaktır” (Wing 2006:34).

Hesaplamalı düşünmeyi zihinsel bir dönüşüm gibi kabul etmek, hızlı verilmiş bir karar olsa da kesinlikle mümkündür. Boyd ve Crawford’un titiz bir ilgiyle inceledikleri “büyük veri hesaplama kültürü” (2012:665), öyle görünüyor ki bir olgu artık araştırmacıların ilgisine cevap verir düzeyde içerik ve etki zengiliğine sahiptir. Bu noktada Wing, bir analogi yapar ve şöyle bir tahminde bulunur: “Genel hesaplama, dünün rüyası idi, bugünün ise gerçeğidir. HD de bugünün rüyası fakat yarının gerçeğidir” (Wing 2006:35). Teknik olduğu kadar entelektüel bir dönüşüm olarak da kabul edilen, halihazırda üçüncü dalgasının yaşandığı ileri sürülen dijital büyük veriye dayanan “hesaplamalı düşünme” yaklaşımı, birçok çalışmada başat yöntem olarak kullanılmaktadır. Özellikle sosyal bilimlerin uygulamalı alan araştırmalarında tatbik edilmektedir.

b) Multidisipliner yetenekler ve çalışmalar: Büyük veri hesaplamaları, belli paralel programlama türlerine gereksinen paralel hesaplama çevrelerinin tümüdür. Paralel programlama, standart programlamadan daha zor farklı bir genlik (magnitude) düzenidir.

Burada belli bir zorluk vardır ve o Minelli'ye göre, paralel programlamanın, programcının deneyim ve yeteneklerine bağımlı olmasıyla ilgilidir (Michael Minelli, 2012, s. 94).

Donah ve Boyd'a göre büyük veri, araştırma üzerine nasıl düşüneceğimiz konusunda "radikal bir sıçrama" yaratmıştır (2012:665). Ayrıca hesaplamalı düşünme yöntemi gerektiren büyük veri paradigması, ilişkisi zor kurulan alanların dahi etkileşimini mümkün kılan hatta gerektiren çalışmaları teşvik etmektedir. Örneğin veri uzmanları, iş hedeflerini karşılamak için analitik model yaratırken çoklu teknik kullanırlar (Michael Minelli, 2012, s. 122). En iyi sonucu elde etmek için teknikleri bir araya getirip birbirine zincirlerler; en iyi sonucu öngörebilmek için bazı modellerin en etkili özellikleri seçilip çoklu-eklektik bir model içinde birleştirirler. Dahası logaritmik teknikler vasıtasıyla ilk önce bir tahmin modeli yaratıp ardından simülasyon tekniklerini kullanarak milyonlarca senaryoyu değerlendirebilirler. Sonra da optimizasyon teknikleri tatbik ederek modelin çıktılarını maksimize edebilirler. Tüm bunlar için çoklu- disiplinler arası- yeteneklere gereksinim duyulur.

Minelli'ye göre multidisipliner yetenekler, büyük veri tabanlı karar-destek bilimlerini sürdürülebilmek için birleştirilmesi zorunlu bir unsurdur ve en azından "iş + matematik + teknoloji + davranış bilimlerini" bir arada uygulayabilmeyi içermektedir. Analitik tabanlı öngörüler, birçok geleneksel çalışma biçimini tehdit etmektedir. Uzman desteğindeki veri tabanlı kararlar, "iş+teknoloji+matematik+davranış ekonomisi+sosyal antropoloji"yi kullanan disiplinler arası bir yaklaşımı geliştirmeye ihtiyaç duymaktadır (Minelli vd. 2013: 130,132).

c) Bilgi tanımının değişmesi: Büyük veri yalnızca geniş veri kümeleri ve veri toplama araçlarına veya onları manipüle ve analiz etmek için kullanılan süreçlere gönderme yapmaz. Aynı zamanda düşünme biçimlerinde ve bilimsel araştırmalarda hesaplama devrimini gündeme getirir. Büyük veri, gerçek iletişim ağlarını nasıl anlayacağımızı söyleyen, geleneksel bilginin nesnesini değiştiren yeni bir bilgi sistemi ortaya çıkarır.

Danah ve Boyd'a göre, "Büyük veri, bilginin tanımını değiştirmektedir" (2012: 665). Berry'ye göre, büyük veri, düzenleyici felsefe gücünün eksikliğinde bilişim ve bilginin yeniden kuruluşunu sağlar. Felsefenin yerini, Kant'ın tüm sezgilerin ussal temeli olarak gördüğü hesaplamasallık (computationality) almaya başlar. Bu yüzden, yeni bir ontolojik yaklaşım yaratan bir onto-teoloji olarak büyük veri yaklaşımı, anlığın (zekânın) yeni bir tarihsel tesellisi olarak anlaşılabilir (Berry 2011:12). Moretti'ye (2007) göre, büyük veri ne salt bir ölçek meselesidir, ne onu yaklaşıksallık terimlerinde düşünmek yeterlidir. Büyük veri, metnin uzak veya yakın analizi olarak gönderme yaptığı şey de değildir. O daha çok, epistemolojik ve etik seviyede derinlemesine bir değişimdir. Boyd'a göre, bilginin yapısı ve araştırma süreçleri, bilgiyle nasıl meşgul olacağımız, doğa ve gerçekliğin kategorileştirilmesi gibi anahtar soruları, yeni bir çerçeveye oturtmaktadır (2012:665).

d) Erişilebilirlik ve dijital akademik uçurum: Büyük veri kümeleri, ilk başlarda anlamsızdı ve yönetmesi zordu. Bu yüzden onlarla yalnızca sosyal bilimciler ilgilendiler. Bugün ise aldıkları eğitime ilgisiz olarak herkes için, kısmen bile olsa kolayca erişilebilir hale gelmiştir (Danah ve Boyd 2012:664). Leonelli'ye göre veriyi herkes için açık kılmanın akıllıca yollarına ihtiyacımız vardır (Leonelli 2014).

Yapısökümcü filozof Jaques Derrida, demokratik katılımı, “bir arşive erişim ve onun yapılarına ve yorumlarına erişim” olarak tanımlamaktadır (Derrida 1996:4). Erişim, akademi için her zaman temel sorunlardan biri olmuştur. Ekonomi, kişisel bağlantı, saygınlık ve dil gibi erişim anahtarlarına sahip olmak, akademik araştırmalarda avantaj sağlar. Fakat dijital çağdaki hesaplamalı düşünme devrimiyle birlikte veriye erişim de bir anahtar statüsü kazanmıştır. Google veya sosyal medya şirketlerinde çalışan uzmanların elinin altındaki devasa veriye dayalı popüler ve de prestijli araştırma raporları, gerçekte büyük veriye sınırlı erişimin doğurduğu yeni bir dijital uçumu da göstermektedir. “Google, her gün 3 milyardan fazla arama sorgusu alıp kaydettiğinden çalışacağı çok miktarda veri vardı[r]” (Mayer-Schönberger, 2013, s. 10). Facebook için çalışan bir antropolog veya Google için çalışan bir sosyologun, bilim topluluğunun geri kalanının ulaşamayacağı verilere erişim imtiyazına sahip olduğunu belirten Boyd ve Crawford, şöyle yazmaktadır:

“Bazı şirketler bütün veriye ulaşmayı engellemekte; diğerleri özel erişimi bir ücretle satmakta, diğerleri ise üniversite destekli araştırmacılara küçük miktarda veri önermektedir” (Boyd ve Crawford 2012:673).

Hiç kuşku yok ki erişim sınırlılığı olgusu, dijital sistemde düzeltilmesi gereken bir “yamukluk” olduğunu gösterir. Bununla birlikte büyük veri şirketlerinin ergin olmayan tutumu dile getirilmelidir: Büyük veri sınıfı, akademik ilginin kendilerine yönelmesiyle birlikte tanınıp güçlendiklerini göz ardı ederek ellerindeki [tüm] veriye erişimini kısıtlamışlardır. Böylece muhafazakar ve dışa açık olmayan bir baron tavrı sergileyerek akademiye karşı diyet borçlarını unutmuşlardır. Veri baronlarının adil rekabeti engelleyen bu tekelci tavrı, akademik açıdan sorunun yalnızca tek bir yönünü ifade eder. Akademinin kendi içinde ivedi çözüm talep eden teknoloji okuryazarlığı gibi bir yeterlilik sorunu vardır. Dijital veriye veya büyük veriye erişim, veriyi saklama ve analiz işlemleri, belli düzeyde bir bilgisayar müktesebatı gerektirmektedir. Bizce bu, rakamları okuyabilen ve teknolojik cihazları etkin kullanabilen akademisyenler lehine yeni bir hiyerarşiye yol açar.

Vatandaşların, başta ekonomik parametreler olmak üzere, diğer sosyal kurumlara katılma ve onlardan pay alma oranlarına göre oluşturulan sosyal tabakalaşma pramidinde dijital uçuruma maruz kalmalarının anlamlı gerekçeleri bulgulanabilir. Tabakalaşma pramidindeki katmanları birbirinden ayıran sosyal uçuruma paralel gerçekleşen dijital uçurumu, meslekler ve uzmanlıklar arasında gözleyebilmek mümkün iken dijital akademik uçurumu, aynı meslek grubu içinde bölünmeye yol açan bir farklılaşma olarak gözlenebilmektedir. Yani

dijital uçurum, sosyal tabakalaşma pramidine paralel bir dağılım gösterirken akademik dijital uçurum ise bir meslek grubunun kendi içindeki bölünmesine gönderme yapar.

Özellikle elektronik ve akıllı teknolojilerin etki alanlarının genişlemesiyle artan büyük veri bankalarına erişim, veri toplayan büyük yazılım ve donanım şirketlerinin tekeline terk edilmesinin etkileri, demokratik hak ve özgürlüklük, baskı unsuru, âdil gelir dağılımı gibi kavramlar bağlamında analiz edilmeye ihtiyaç duyulmaktadır. Feminist açıdan ele alındığında akademik dijital uçurumun, yalnızca kuşaklar arasında değil, cinsiyetler arasında da bir eşitsizliğe yol açtığını söylemek mümkündür. Örneğin erkek araştırmacıların, kadın araştırmacılara göre bilgi-işlem teknolojilerini kullanma konusunda daha eğilimli ve eğitilmiş, becerili ve istekli oldukları belirtilmektedir. Feministler, bilim ve teknolojinin eril tarzda kurumsallaşmasını bu açıdan zaten eleştirmektedirler (Bkz. Işıklı 2014a:13). Dijital toplum veya büyük veri çağı, dışıl usun saf dışı bırakılma, pasivize edilme, eril usa bağımlılığının artırılmasının yeni bir tekniği olarak okunmaktadır.

e) Google uzmanları, akademiye karşı: Leonelli'nin "muhafazakarlık tehlikesi" diye adlandırdığı "büyük veri durumu", verinin herkes için erişilebilir olmayışını ifade eder. Eğer veriye erişimi, bilgiye erişimin veya sosyo-ekonomi-politik eşitlik hakkının bir genişletilmesi olarak görme aşamasına ulaşırsak veriyi herkes için erişilebilir kılmanın yollarını bulmamız gerekir. Ne de olsa veri, kapitale dönüşen bir hammadDEDİR; sık söylendiği gibi, bilgi toplumunun petrolüdür. Üstelik toplumsal veri, toplumdaki tüm bireylerin katkısıyla oluşmaktadır. Veriyi, bir tür gayri safi millî hasılı gibi değerlendirmenin zamanı gelmiştir. Bir sosyal sermaye olarak veri, tüm vatandaşlara eşit biçimde sunulabilmelidir. Fakat halihazırda veri tekelleşmesi yaşanmaktadır. Minelli'ye göre her trend, karşı trendini yaratır. Veri analiz alanı da bu fenomene karşı bağışık değildir. Birkaç yıldır bu alan, büyük veri dağlarının artışı yoluyla çatırdadığı için heyecan verici yeni teknolojiler ve araçlarının utanılacak derecede hızlı ortaya çıkışı sayesinde çabucak demokratikleştirilmelidir (Michael Minelli, 2012, s. 128).

Geçmişten günümüze kadar sosyal bilimciler; alan araştırması, görüşme, gözlem ve deneyler yoluyla kişiler arası ilişkiler hakkında veri topladılar. Bu veriyi kullanarak insanların kişisel ilişki ağlarını ve kişilerin sürdürdüğü ilişki gruplarını tasvir etmeye odaklandılar. Bugün, "sosyal grafiklere kafayı takmış bir uzmanlar grubu" nun ortaya çıkışı, sosyal ağ sitelerinin artışıyla ilişkili bir olgudur. Binlerce araştırmacı, mesajlar ile hesaplar arasındaki bağlantıyı analiz etmek, sosyal ağlar hakkında bir öneride bulunmak için sosyal medyaya koşar durumdadır (Boyd ve Crawford 2012:670). Açık ki Google, Facebook, Twitter gibi sosyal medya şirketleri, insanlar arasındaki ilişkileri, trendleri ve kişisel tercihleri incelerken büyük verinin tamamını ellerinde bulundurmak gibi bir güç ve ayrıcalığa sahiptirler. Çevrim içi uygulamalarla çalışan bu şirketler, yarattıkları yeni uzman tipi ve uzmanlık alanı sayesinde özellikle beşeri bilimlerde akademiye bir alternatif ve alternatif bir sosyal araştırma tarzı trendi oluşturdular. Dijital veri tabanlarına erişim imtiyazına sahip uzamanların, akademik

uzmanlık ve akademik araştırmayı, en hafif ifadesiyle egale etme, fütürist bir sezgiyle de ekarte etme gücüne sahip oldukları görülmektedir. Öyle ki Google uzmanları, bugün akademi uzmanlarından daha fazla popüler, daha fazla etkilidiler. Bu yüzden şimdiden başkalarıyla birlikte şunu söylememiz mümkündür: Bazı mesleklerin eskide kalması kadar, halkın çoğunun bilimden anladıklarını dille getiren akademi uzmanlarının itibarsızlaşması, alan uzmanlığının önemsizleştirilmesi söz konusudur.

Büyük veri tabanlı araştırmalarda alan uzmanlığı, olasılık ve korelasyonun çok büyük ve çok önemli olduğu bir dünyada daha az veriyle çalışan önemsiz ve demode bir alan haline gelir. İstatistiksel hesaplama, alan uzmanlığının derin tecrübesine ortak olmaya başlamıştır. İçgüdü, sezgi ve feraset yerini, yoğun istatistik, karmaşık matematiksel modellere ve veri analizine bırakmaktadır. Sosyal medya uzmanlarının, üniversitelerin istatistik veya sosyoloji bölümlerine paralel yeni bir etki alanı yarattıklarını kabul etmemiz gerekir. Yeni alanın yükselişi, akademinin itibarsızlaşmasına yol açıp açmayacağını zaman gösterecek. Schönberger ve Cukier'a göre akademi uzmanları yok olmayacaklar fakat büyük vericilerin söyledikleriyle başa çıkmaları gerekecektir (Mayer-Schönberger 2013:24).

5. SONUÇ

Büyük veri; yeni bir olgu olmasına rağmen eski ve köklü kavramlar üzerine esaslı bir tartışma başlatmış, bazılarında şimdiden bir dönüşüm sağlamış durumdadır. Sağlık, işletme, ekonomi, spor gibi alanlarda önemli farklar yaratmış, bilgi ve bilimle ilgili tasarımlara yol açmıştır. Beşeri bilimlerin fizikalizm baskısından kurtulmasına yönelik matematiksel bir model sunmakla kalmamış, klasik mekaniğe dayalı bilim yapma tarzında sorunlu görünen bazı kavramlara gerek duymadan bilim yapma imkanı sunmuştur. Büyük veri, beşeri bilimler için yeni bir başarı alanı haline gelmiş, veri ve bilgiden maksimum yararlanma imkanını yaratmış; iklim değişikliğini ele almak, hastalıkların kökünü kurutmak, iyi yönetim ve ekonomik kalkınmayı teşvik etmek gibi ivedi global problemlerin çözümünün bir parçası haline gelmiştir. Aslında büyük veri, en nihayetinde doğanın kendisinin büyük veri tabanı olduğu gerçeğiyle bilim insanlarını başa çıkmaya zorlamaktadır. Büyük veriyle birlikte bilim insanları, doğanın ayrıntıları, tüm karmaşıklığı ve içsel zenginliği ile yüzleşmeye motive edilmektedir. Büyük veride, eski usül fen deneyleriyle elde edilen az miktardaki verinin içindeki doğanın karmaşıklığını aramaktan farklı bir durum söz konusudur. Yaneer Bar-Yam'ın tespit ettiği gibi, ne kadar büyük miktarda veri elde ettiğinizin önemi yoktur, çünkü ilgili olgunun tümünü kapsayacak büyüklükteki yeterli veri elde edilemez. Fakat şimdilik elde ettiğimiz kadarı bile etik, akademik, epistemik tartışmaları ateşlemeye yetmiştir.

Sonuçta, yeni bir dönüşüm çağın arefesinde olduğumuzu söyleyebiliriz. Ne getireceğini az çok öngörebildiğimiz fakat tehdit ve imkanlarına hazırlanmakta yeterince başarılı olamadığımız hatta elde ettiğimiz veriyi tümüyle işleyip bilgi ve bilgeliğe dönüştürmekte başarılı olamayacağımız dijital ve sibernetik bir çağa doğru yol almaktayız. Artık bu dijital veriye temellenen hesaplamalı düşünme, dijital kültür, dijital bilgi içeren büyük veri paradigmasının egemenliğine kendimizi hazırlamamız gerekir.

Kaynakça

Aristoteles. (1996). *Metafizik*. (A. Arslan, Çev.) İstanbul: Sosyal Yayınları.

Bacon, F. (2010). *New Organon or: True Directions Concerning the Interpretation of Nature*. Jonathan Bennett: Copyright ©2010–2015 A.

Berry, D. M. (2011). The Computational Turn: Thinking About The Dijital Humanities. *Culture Machine*(12), ss.1-22.

Click, J. (2000). *Kaos*. (F. Üçcan, Çev.) İstanbul: TUBİTAK.

- Danah Boyd ve K. Cukier. (2012). Critical questions for big data. *Information, Communication and Society*, 15(5), ss. 662-679.
- Derrida, J. (1996). *Archive Fever: A Freudian Impression*. (T. E. Prenowitz, Çev.) Chicago: Chicago University Press.
- Derrida, J. (2005). *Writing and Difference*. (A. Bass, Çev.) London: Routledge.
- Einstein, A. vd. (1935). Can quantum mechanical description of physical reality be considered complete? *Physical Review*(47), ss.777-782.
- Erdoğan, İ. (2007). *Pozitivist Metodoloji*. Ankara: Erk Yayınları.
- Güzel, H. C. (1996). *Sağduyu Filozofu: Popper*. (H. C. Güzel, Derl. ve Çev.) Ankara: Bilim ve Sanat Yayınları.
- Harris, J. (2013, Mart 14). *On Philosophy, Science, and Data*. Ocak 3, 2015 tarihinde “<http://www.ocdqblog.com/home/on-philosophy-science-and-data.html>” adresinden alındı.
- Heisenberg, W. (2000). *Fizik ve Felsefe (İdealizm: Determinizmden Olasılığa Doğru)*. (Y. Öner, Çev.) İstanbul: Belge Uluslararası Yayıncılık.
- Işıkli, Ş. (2012). *Kuantum Felsefesi: Postmodern Bilimin Doğuşu*. Ankara: Birleşik Yayınevi.
- Işıkli, Ş. (2014a, Mayıs-Haziran). Kadına Tahakküm ya da Eril Usun Tavırları. *Uluslararası Online Akademik Bakış Dergisi*(43), 1-31.
- Işıkli, Ş. (2014b, Eylül 18). *Dijital Kültür ve Türkiye’de Çevrimiçi Kullanıcı Profili*. Şubat 13, 2015 tarihinde www.dijitalhayat.tv: “<http://dijitalhayat.tv/yazarlar/dijital-kultur-ve-turkiyede-cevrimici-kullanici-profil>” adresinden alındı.
- Kant, I. (2014). *Critique of Pure Reason (1781)*. (J. M. Meiklejohn, Çev.) web edition: eBooks@Adelaide.
- Leonelli, S. (2014). *Philosophy of data – intensive science*,. Aralık 27, 2014 tarihinde “<http://mith.umd.edu/sharedhorizons/wp-content/uploads/Shared-Horizons.Sabina-Leonelli.pdf>” adresinden alındı.
- Locke, J. (1992). *İnsan Anlığı Üzerine Bir Deneme*. İstanbul: Kabalcı Yayınevi.
- Mayer-Schönberger, K. Cukier. (2013). *Büyük Veri: Yaşama, Çalışma ve Düşünme Şeklimizi Dönüştürecek Bir Devrim*. (B. Erol, Çev.) İstanbul: Paloma Yayıncılık.
- Michael Minelli, M. C. (2012). *Big Data, Big Analytics: Emerging Intelligence and Analytics Trends for Today's Business* (Cilt 578). USA: John Wiley and Sons, Inc.
- Osborne, R. (2006). *Yeni Başlayanlar İçin Felsefe*. (İ. Sezer, Çev.) İstanbul: Nokta Kitap.
- Özlem, D. (2008). *Felsefe ve Doğa Bilimleri* (4. baskı). Ankara: Doğu Batı Yayınları.

Presner, J. S. (2009, October 14). Digital Humanities Manifestos 2.0.

Presner, T. (2010). *Digital Humanities 2.0: A Report on Knowledge*. October 15, 2010 tarihinde <http://cnx.org>: <http://cnx.org/content/m34246/1.6/?format=pdf> adresinden alındı

Siegfried, T. (2013a, November 26). *Why big data bad for science*. Aralık 23, 2014 tarihinde Science News Magazine of the Society for Science and The Public: ["https://www.sciencenews.org/blog/context/why-big-data-bad-science"](https://www.sciencenews.org/blog/context/why-big-data-bad-science) adresinden alındı

Siegfried, T. (2013b, December 3). *Rise of Big Data underscores need for theory*. Ocak 01, 2015 tarihinde Science News Magazine of the Society for Science and The Public: <https://www.sciencenews.org/blog/context/rise-big-data-underscores-need-theory> adresinden alındı.

Wikipedia. (2013, Şubat). *Büyük Veri*. Ocak 22, 2015 tarihinde www.wikipedia.org: ["http://tr.wikipedia.org/wiki/B%C3%BCy%C3%BCk_veri"](http://tr.wikipedia.org/wiki/B%C3%BCy%C3%BCk_veri) adresinden alındı

Wing, J. M. (2006, Mart). Computational Thingking. *CACM Viewpoint*, ss. 33-35.

www.wikipedia.org. (2013, Şubat). Aralık 23, 2014 tarihinde www.wikipedia.org: ["http://tr.wikipedia.org/wiki/B%C3%BCy%C3%BCk_veri"](http://tr.wikipedia.org/wiki/B%C3%BCy%C3%BCk_veri) adresinden alındı.

www.merriam-webster.com. (2015). Ocak 15, 2015 tarihinde ["http://www.merriam-webster.com/dictionary/data"](http://www.merriam-webster.com/dictionary/data) adresinden alındı.