

Büyük Veri, Duygular, Dil ve Etik Üzerine Bir Değerlendirme

Şeyma Kuru¹, Fırat Başbuğ², Sezin Andıç³

Kuru, Ş., Başbuğ, F. ve Andıç, S. (2023). Büyük veri, duygular, dil ve etik üzerine bir değerlendirme. *nesne*, 11(28), 336-345. DOI: 10.7816/nesne-11-28-10

Anahtar kelimeler

Büyük veri, duygular, dil, etik

Keywords

Big data, emotion, language, ethics

Öz

İnsan-makine etkileşimi sonucunda devasa boyutta, yüksek hız ve çeşitlilikte olan büyük veri, bilgileri toplama, düzenleme ve analiz edebilme potansiyeli nedeniyle psikoloji araştırmaları için yeni bir disiplinin gelişmesine katkıda bulunmaktadır. Psiko-informatik adı verilen bu disiplin insan duygularını, düşünce ve davranışlarını tahmin etmede bilgisayar bilimlerinden yararlanmaktadır. Teknolojinin hayatın her alanına temas ettiği içinde bulunduğumuz yüzyılda internetin kullanımı, insanların sınırsız bilgiye çok hızlı bir şekilde ulaşmasının yanı sıra kişilerin dijital dünyaya izler bırakarak sınırsız bilgi oluşturmasını da sağlamaktadır. Bugün psiko-informatik çalışmalarına bakıldığında uluslararası alanyazında daha çok bireylerin kişilik özelliklerini; akıllı telefon, sosyal medya ve internet kullanım davranışlarını; psikolojik sağlıklarını ve problemleri tahmin etme üzerine çalışmalar olduğu görülmektedir. Mevcut çalışma psiko-informatik konusuna dikkat çekmeyi amaçlarken, psiko-informatiğe dair kısa bir değerlendirme sunulmuş ve ortaya çıkabilecek etik problemler tartışılmıştır. Psikoloji, dilbilim ve bilgisayar bilimleri gibi alanlarda çalışan araştırmacıların, ortak çalışmalar yürütmesi gerekliliği ve çalışmaların sağlıklı yürütülebilmesi için temel etik standart konusunda bilgilendirilmesi öncelikli görünmektedir.

An Evaluation of Big Data, Emotions, Language and Ethics

Abstract

As a result of human-machine interaction, big data of gigantic size, high speed and variety contributes to the development of many new disciplines for social science research due to its potential to collect, organise and analyse information. This discipline, called psycho-informatics, uses computer science to predict human emotions, thoughts and behaviours. In the current century, where technology touches every aspect of life, the use of the internet enables people to access unlimited information very quickly, as well as creating unlimited information by leaving traces in the digital world. Today, when we look at the studies in this field, it is seen that in the international literature, there are mostly studies on predicting individuals' personality traits; smartphone, social media and internet usage behaviours; psychological health and problematic behaviours. While the present study aims to draw attention to psycho-informatics, a brief overview of psycho-informatics is presented and ethical problems that may arise are discussed. It seems to be a priority for researchers working in fields such as psychology, linguistics and computer sciences to be informed about the basic ethical standards in order to carry out joint studies and to carry out healthy studies.

Makale Bilgisi

Geliş tarihi: 27 Eylül 2022

Düzeltilme tarihi: 30 Nisan 2023

Kabul tarihi: 3 Haziran 2023

DOI: 10.7816/nesne-11-28-10

¹ Araş. Gör., Antalya Bilim Üniversitesi, Psikoloji Bölümü, seyma.kuru(at)Antalya.edu.tr, ORCID: 0000-0001-7785-9903

² Dr. Öğr. Üyesi, İstanbul Medeniyet Üniversitesi, Dilbilimi Bölümü, firatbasbug(at)gmail.com, ORCID: 0000-0001-9824-123X

³ Dr. Öğr. Üyesi, Çukurova Üniversitesi, Psikoloji Bölümü, sezinandic(at)gmail.com, ORCID: 0000-0002-5144-0685

2016 yılında Dünya Ekonomik Forumu, Endüstri 4.0 konusunu gündemine alırken, WEF'in Başkanı Klaus Schwab hepimiz için yeni ama biraz da ürkütücü olan bu devrimi şu sözlerle anlatıyordu: “Yaşamımızı, çalışma ve birbirimizle ilişki kurma şekillerimizi temelden değiştirecek teknolojik bir devrimin eşiğindeyiz. Ölçeği, kapsamı ve karmaşıklığı ile bu dönüşüm insanlığın daha önce tecrübe ettiği herhangi bir şeye benzemiyor. Henüz nasıl ortaya çıkacağını bilmiyoruz, ancak bir şey açık: Buna verilecek yanıt, kamu ve özel sektörden akademiye ve sivil topluma küresel politikanın tüm paydaşlarının bütünleşmiş ve geniş katılımlı olmasını mecbur kılıyor (Schwab, 2016). Böylece Klaus Schwab Endüstri 4.0'ın ilk tanımlarından birini yaparken, aslında devrimin anahtar kelimelerinden birinin de altını çizmiş oluyordu. Birbirine mecbur kurumlar, birbirine bütünleşik reaksiyonlar. Sonradan anladığımız ve tecrübe ettiğimiz kavramla biraz değiştirirsek aslında, yolun başındaki Endüstri 4.0 her şeyin birbirine bağlı olduğu akıllı bir toplumu işaret etmektedir. Akıllı toplum ve birbirine bağlı olmak Endüstri 4.0'ın temeldeki işlevini ve devrimsel karakterini ortaya koysa da bu iki temel kavramın henüz yolun çok başında olan E4 için ne kadar güncel kalacağı belirsiz. İşte, E4 için yeni bir kavram daha: Her şey her şeyin yerine kolayca geçebilir. Her şeyin her şeyin ikamesi olabilecek bir sistemde başlangıçtaki şeyi yani henüz ikamesi olmayanı kimin üreteceği tartışmalı görünse de E4'ün bize sunduğu imkanlar, Andre Kudelski'nin ifadesi ile henüz yeterince anlaşılmadı. “Yetenekli bir mühendis birbiriyle bağlantılı olan herhangi bir şeyin kontrolünü uzaktan da alabilir. Toplum henüz bu yeteneğin nasıl inanılmaz senaryolar gerçekleştirebileceğini anlamadı (World Economic Forum, 2016). 2016'da “önümüzdeki yıllarda iş yapış ve yaşama şekillerimizi temelden değiştireceği” söylenen, Yale Üniversitesi Ekonomi Profesörü, 2013 Nobel Ödüllü Robert J. Shiller'ın “Yangın sigortası yaptırmak için evinizin yanmasını bekleyemezsiniz. Biz de Dördüncü Endüstri Devrimi'ne hazırlanmak için toplumumuzun kökten altüst olmasını bekleyemeyiz” diye nitelediği bu görece “yüksek sesli” devrim, kendisinden önceki bütün devrimlerin kazanımlarıyla beslenip, kendisinden sonra gelecek olan -ki gelirse- bütün devrimlerin de temelini oluşturmaya hazır (Endüstri 4.0 Platformu, 2016). Bu büyük devrim bizi yapay zekâ, büyük veri, doğal dil işleme, yapay zekâ tabanlı makineler gibi yarı ütopyik yarı distopik yeni kavramlarla tanıştırdı. Bu tanışmaya ilk vurgu yapanlardan Keynes 1929'da ekonomik büyümenin işsizliği artıracakını öngörürken aslında bunda endüstriyel otomasyonun rolüne de dikkat çekmişti. Keynes'e göre otomasyon- işçileri makinelerle değiştirme- açısından çok büyük fırsatlar olduğunu fark etmiş ve istihdamdaki azalmanın ve teknolojik gelişmelerin bağlantılı olduğunu dile getirmişti. Bu teknolojik işsizlik bizim de yakın zamanda sıkça duyacağımız yapay zekanın her mesleğin ya da bireyin yerine geçebileceği anlamına gelebilir (Acemoğlu, 2022).

İnsan-makine etkileşimi sonucunda devasa boyutta, yüksek hız ve çeşitlilikte olan büyük veri, bilgileri toplama, düzenleme ve analiz edebilme potansiyeli olarak ifade edilebilir (Zadrozny ve Kodali, 2013, s.2.). Yapay zekâ bizi duymadığımız, görmediğimiz, etkileşim içinde olmadığımız olaylar ya da kişiler hakkında bilgi sahibi yapabilir mi? Yanıt biraz ironik olmakla birlikte kesinlikle hayır. Yapay zekânın, bizimle etkileşim halinde olmadığı, bizi görmediği ya da duymadığı koşullarda ne bizi ne de duygularımızı yönlendirmesi beklenemez. Ancak böyle bir koşulda yaşamak ancak dağ başında bir kulübede herkesten ve birbirine bağlı olan her şeyden uzak yaşamakla mümkün olabilir. Endüstri 4.0 konseptiyle “Birbirine mecbur kurumlar, birbirine bütünleşik reaksiyonlar”, bizi birbirimizin duygularına, karşıtlıklarına, tercihlerine bağlı bir bireye dönüştürüyor. Yapay zekâ ve büyük veri işleme, bu mecburiyetlerden beslenip büyüyerek bize bu mecburiyetleri daha da zorunlu kılıyor. Yapay zekânın bizi görebilmesi ya da yapay zekâyla etkileşim içinde olmak bir yan anlam mı? Yanıt yine hayır. Dünyadan bilgi akışına -ki bu, sabah kahvaltıda yediğiniz bir simitin tazeliği hakkındaki bir arkadaş yorumu da olabilir- maruz kaldığımız her an aslında büyük verinin bir parçası yapay zekânın da iletişime geçtiği bir “nesne” haline gelmiş oluyoruz. Bu noktada düşüncemiz ve

duygularımız ve kullandığımız dil, başka bir araçtan beslenen bir aparata dönüşüyor. Son 30 yılda demokrasiler bazı önemli sınavlardan kaldı. Acemoğlu'nun (2022) Amerika özelinde söylediği aslında bütün dünya demokrasilerinin ortak problemlerinden üç anahtar kelime seçelim: Kutuplaşma, Güvenilir Medya, Paranın Siyasetteki Rolü. Yapay zekâ devleri, bu sorunlu bölgeleri çok iyi bildikleri için, bu fay kırıkları yapay zekâ ile devasa boyutlara ulaşmaya başlamaktadır (Acemoğlu, 2022, s.16).

Büyük verinin yakın tarihte en çok konuşulduğu olaylardan birini hatırlayalım. İngiltere merkezli Cambridge Analytica olayında, Facebook kurucusu M. Zuckerberg'in Kongre'de verdiği ifadeye göre 87 milyon Facebook kullanıcısının (NY Times, 2018) verileri "This is Your Digital Life" isimli uygulamayı kullanan yaklaşık 300 bin kullanıcı (The Guardian, 2018) sayesinde elde edilmişti. Her şeyin birbirine bağlı olduğu ideal sistem, 300 bin kişinin arkadaş ağını kullanarak 87 milyonun acılarına, heyecanlarına, sevinçlerine, onları mutlu/mutsuz eden şeylere doğrudan ortak oldular, günü geldiğinde kullanmak üzere bu verileri işlediler. Sonrasındaki spekülasyonlar da Amerikan seçimlerinde, önceden zaten dünya görüşlerini, sevdiği ve öfkelendiği şeyleri iyi bildikleri son kullanıcıların önüne, yine onları öfkeli edilecek ya da onların hoşuna gidecek içerikleri çıkardıkları yönünde. Bu içeriklerle kaç kişinin oy tercihlerini değiştirdikleri bilinmemektedir. Burada ürkütücü olanın, 87 milyon kişinin önüne, kutuplaştırmayı artırıcı, demokratik kurumları yıpratıcı dahası bireysel özgürlükleri ihlal edici içerikler koyabilme ihtimallerinin olabileceğidir. Algoritmaları kullanarak, kullanıcıların karşılaştıkları reklamlara ve haber içeriklerine müdahale gücü sadece seçim dönemi kampanyalarında değil, internete bağlandığımız her an bizim özgürlük alanımızda bizi bekliyor görünmektedir.

Teknolojinin hayatın her alanına nüfuz ettiği 21. yüzyılda internetin kullanımı, insanların sınırsız bilgiye çok hızlı bir şekilde ulaşmasının yanı sıra kişilerin dijital dünyaya izler bırakarak sınırsız bilgi oluşturmasını da sağlamaktadır. Akıllı cihazlar, sosyal medya platformları, çeşitli uygulamalar ve web sayfaları gibi dijital araçlar üzerinden kaydedilen bu bilgiler, saf veriler halinde internet ortamında toplanmaktadır (Ergen, 2018). Söz konusu veriler, hacimlerinin oldukça büyük olması, aynı anda birçok farklı platformdan veri girişinin gerçekleşmesine olanak sağlayan hızı ve çeşitliliği nedeniyle daha önce kullanılan yöntemlerle analiz edilememektedir. Bu durum "büyük veri" kavramını ortaya çıkarmıştır (Chen ve arkadaşları, 2014, s.4; Doğan ve Arslantekin, 2016). Büyük veriye dair üzerinde uzlaşmış bir tanım olmamakla birlikte; Zadrozny ve Kodali (2013, s.2.), büyük veriyi "gelişmiş karar verme, içgörü keşfi ve süreç optimizasyonu sağlamak için yeni işleme biçimleri gerektiren yüksek hacimli, yüksek hızlı ve çok çeşitli bilgi varlıkları" olarak tanımlamanın kapsayıcı olacağını düşünmüştür. Bir diğer tanımla büyük veri, büyük veri kümeleri ve hesaplama yöntemleri kullanarak metinlerin, videoların, seslerin kısacası insan ya da makineler tarafından üretilen her türlü verinin depolanmasını, incelenmesini, işlenmesini ve yeni çıktılar üretilebilme kabiliyetini ifade eder. Bu tanımdan yola çıkarak büyük verinin dijital nesnelere alınan bilgileri depolayabildiği, gerektiğinde analiz edebildiği, çeşitli algoritmalar kullanarak kişilerin tercih ve davranışlarını tahmin edebildiği söylenebilir. Bu özellikleri nedeniyle de ekonomi, politika, endüstri ve akademi alanlarında büyük veriye olan dikkat de ilgi de giderek artmaktadır (Chen ve arkadaşları, 2014, s.1).

Büyük Veri ve Sosyal Bilimler

Alan Turing (1950), yapay zekâ çalışmalarının öncüsü sayılan makalesine, "Makineler düşünebilir mi? sorusu ile başlamıştı. Turing'in (1950), "makine" ve "düşünmek" kavramları üzerinden, kısa ama güçlü bir tartışma açtığı makalesinin hayatımızı bu kadar değiştirebileceğini kim tahmin edebilirdi? Endüstriden sanata, tıptan turizme, psikolojiden spora, dilbilimden uzay çalışmalarına kadar büyük veri hayatın her alanına hızla yerleşmeye başladı. İnsan makine etkileşiminin bu denli yüksek olması ve bu etkileşim

sonucunda bireylerin duygu, düşünce ve davranışlarının büyük veri olarak dijital dünyada kaydedilmesi psikoloji, sosyoloji, dilbilimi, iktisat, finans, pazarlama gibi sosyal bilim alanlarında oldukça önemli bir rol oynuyor. Sosyal bilimlerde, insanlarla yapılan çalışmaların önemli bir bölümünde anketler kullanılmaktadır. Örneğin psikolojiyi araştırmalarında anketlerin kullanımı, anket sorularının sıklıkla kapalı uçlu olması ve buna bağlı olarak anketlerden kişiler hakkında edinilen bilgilerin görece yüzeysel olması; anketlerin öz bildirime dayanması sonucu kişilerin verdiği yanıtların yanlılık içerebileceği gibi sınırlılıklara sahiptir (Katschnig, 2006; Seçer, 2018:32). Bununla birlikte hem psikoloji araştırmacıları hem de klinisyenler için bireyleri yirmi dört saat gözlemlemek ya da bir psikolojik testi gün içinde birden fazla kez uygulamak mümkün değildir (Markowitz ve ark., 2014). Benzer şekilde, dilbilimde ya da sosyoloji alanında bir saha araştırmasını ele alalım. Yarı yapılandırılmış görüşmelerdeki yanıtların deşifresi ve incelenmesi de saatler hatta günler sürebilmektedir. Diğer taraftan, bireyler hakkındaki verilerin; bireylerin, araştırmacıların ya da klinisyenlerin çabası olmaksızın, yanlılıklardan uzak bir şekilde, çeşitli akıllı cihazlar aracılığıyla otomatik olarak toplanabilmesi ve kaydedilebilmesi (Markowitz ve ark., 2014) bize çok daha farklı sonuçlar verebilmekte ve sonuçlarımızı çok daha hızlı işleme imkanı sunmaktadır.

Odağına insanı alan psikolojiden sosyolojiye, dilbilime kadar farklı bilimlerde işlenebilen ve derlenebilen verinin önemi büyüktür. Bu bağlamda psikoloji ya da diğer sosyal bilim alanlarında büyük veri kullanımı, davranışın olduğu gibi kaydedilmesini sağlaması, geleneksel yöntemlerle elde edilemeyecek kadar büyük miktarda veriye ulaşması üstelik bunları yaparken hızlı, ekonomik ve görece yansız olması nedeniyle alternatif yöntem olarak görülmektedir. Ancak çok fazla kullanıcı bilgisine filtresiz bir şekilde sahip olan büyük veri, bu bilgilerin temizlenmesi, analiz edilmesi ve modellenmesi için bilgisayar bilimine ihtiyaç duymaktadır (Montag ve ark., 2016). Söz konusu ihtiyaçlar, sosyal bilim araştırmalarında bilgisayar bilimlerine ait yöntem ve teknikleri kullanan yeni bir enformatik alt alanı gündeme getirmiştir. Örneğin, psikoloji ve enformatiğin iş birliği yaptığı Psiko-enformatik alanı, geleneksel veri toplama ve analiz sürecinde teknoloji kullanımından çok daha ötede, insan-makine etkileşimi sonucu ortaya çıkan büyük veriyi analiz etmeye yönelik algoritmaların kullanımına odaklanmaktadır (Montag ve arkadaşları, 2016). Bu bağlamda akıllı cihazlar, sosyal medya platformları, arama motorları, dijital oyunlar, web siteleri, bloglar ve çok daha fazlası psiko-informatik için bireylerin tutum ve davranışlarını gözlemlemek açısından oldukça önemli malzemelerdir. İlgili alanyazında psiko-enformatiğe dair yapılan çalışmalara bakıldığında sosyal medya platformlarından elde edilen verilerden yola çıkarak Büyük Beşli kişilik özelliklerinin tahmin edilip edilemeyeceğini araştıran çalışmaların olduğu görülmektedir (Azucar ve arkadaşları, 2018; Marengo ve Montag, 2020).

Azucar ve arkadaşlarının (2018) meta-analiz çalışmasında yer alan ve Büyük Beşli kişilik tahmini üzerine çalışılan yirmi dört farklı araştırma, sosyal medyadan elde edilen demografik bilgilerin; kullanıcıların sosyal medya platformlarındaki gönderi ve takip sayılarını içeren kullanıcı etkinliğinin; sosyal medya beğenilerinin, paylaşılan metinler ve resimlerin analizine odaklanmaktadır. Bu meta-analiz çalışması sonucunda, analiz edilen verilerin, çoklu olarak ele alındığında kişilik özelliklerini tahmin etmede başarılı olduğu söylenebilmektedir (Azucar ve arkadaşları, 2018). Bu araştırmaya benzer şekilde Facebook özelinde yapılan bir diğer meta-analiz çalışması da çoklu özelliklere sahip Facebook verilerinin bireylerin kişilik özelliklerini tahmin etmede orta düzey bir doğruluğa sahip olduğunu ve bu nedenle bireysel değerlendirmelerden, gruplara özgü değerlendirmelerin yapılmasının daha doğru olacağını ifade etmektedir (Marengo ve Montag, 2020). Bu doğrultuda sosyal medya verilerinden yola çıkarak bireylerin kişilik özelliklerinin tahmin edilmesinde birden çok sosyal medya içeriğinin kullanılmasının ve yapılan tahminlerin bireysel değerlendirmelerde kullanılmamasının önemli olduğu düşünülmektedir. Dilbilim ve psikoloji araştırmalarının yanı sıra farklı uygulama alanlarında da psiko-enformatiğin kullanılmasının önemli

olduğu düşünölen bir diđer alıřma alanı internet, akıllı telefon ve sosyal medyanın problemlı kullanımınıdır. Yapılan arařtırmalar problemlı internet kullanımını ile zaman algısında bozulma arasında anlamlı iliřkiler olduđuna iřaret etmektedir (Greenfield, 1999; Lin ve ark., 2015). Buradan yola ıkarak hem klinisyenler hem de arařtırmacılar iin bireylerin interneti ve bununla bađlantılı olarak sosyal medyayı veya akıllı telefonlarını ne kadar kullandıkları konusundaki öz bildirimlerindenense, psiko-enformatik yoluyla elde edilen nesnel bilgileri kullanmanın daha dođru sonulara ulařılmasına yardımcı olacađı düşünölmektedir. Nitekim kiřilerin akıllı telefon kullanımına dair soruları ieren öz bildirim ölekleri ile bir uygulama yoluyla telefonlarına kaydedilen nesnel kullanımları karřılařtırıldıđında kiřilerin akıllı telefon kullanımları konusunda yanıldıkları görölmüřtür (Andrews ve ark., 2015; Lin ve ark., 2015; Montag ve ark., 2015). Buna ek olarak, uygulamadan ekilen nesnel veriler akıllı telefon bađımlılıđını kiřilerin öz bildirimlerinden daha iyi tahmin etmektedir (Montag ve arkadaşları, 2015).

Psiko-informatik alanının; kiřilerin internet, sosyal medya veya akıllı telefon kullanımlarına dair nesnel verilere ulařılmasını sađlamasının yanı sıra söz konusu yapıların problemlı kullanımlarının anlaşılmasına da yardımcı olacađı düşünölmektedir. Bu dođrultuda yapılan arařtırmalar incelendiđinde özellikle sosyal medyaya iliřkin hangi bileřenlerin problemlı kullanımla iliřkili olduđuna dair alıřmaların olduđu görölmektedir (Marengo ve ark., 2020; Marengo ve ark., 2022; Marino ve ark., 2017; Peterka-Bonetta ve ark., 2021). Marengo ve arkadaşları (2020), genç yetiřkin Facebook kullanıcılarına ait nesnel durum güncellemeleri ve beđeni sayılarını ieren evrimii etkinliklerin, kullanıcıların kiřilik özellikleri ve problemlı sosyal medya kullanımları arasındaki aracı rolünü incelerken; Peterka- Bonetta ve arkadaşları (2021), Twitter ve Instagram kullanımına dair nesnel verilerin kiřilik, problemlı sosyal medya ve problemlı akıllı telefon kullanımı ile olan iliřkisini incelemiřlerdir. Facebook özelinde yapılan bařka bir arařtırmada ise kullanıcıların son bir yıldaki gönderi etkinliđi, dil kullanımı ve beđeni kategorilerinden oluřan “dijital izler” in problemlı sosyal medya kullanımını tahmin edebildiđi sonucuna varılmıřtır (Marengo ve ark., 2022). Tehlikedeki dillerin incelenmesi, deđiřimin deđerlendirilmesi ve belgelenmesi aısından akıllı cihazlar hem topluluk üyelerinin hem de arařtırmacıların eř zamanlı alıřma, veri depolama ve veri iřleme kabiliyetlerini artırıyor. Dahası artık kaybolan diller üzerine geniř alıřmalar yapılabilmesine olanak veren veri modelleme araları gittike daha fazla kiři tarafından kullanılıyor. Sosyal medya üzerinde alıřan dilbilimciler, psikologlar ve toplumbilimciler iin veri ve iřleme zenginliđi ise üst düzeyde. Günümüzde sosyal medya, bireylerin birbirleriyle iletiřim kurabildiđi, eřitli olay ve durumlar hakkında kendi duygu ve düşünöcelerini ifade edebildiđi bir alan olarak karřımıza ıkmaktadır (Schwartz ve ark., 2013; Settanni ve ark., 2018). Kullanıcıların sosyal medya platformlarındaki ifadeleri, duygu ve düşünöcelerini yansıtmaları nedeniyle kiřilerin psikolojik sađlıkları hakkında ip ucu sađlayabilmektedir (Eichstaedt ve ark., 2018; Marengo ve ark., 2021; Marengo ve ark., 2019; Settanni ve Marengo, 2015). Nitekim Eichstaedt ve arkadaşları (2018), Facebook gönderilerinin metinsel ieriđi, gönderi uzunluđu, gönderi sıklıđı, gönderi modelleri ve demografiden oluřan Facebook dili modelinin depresyonu tahmin edebildiđini ifade etmiřlerdir. Söz konusu alıřmada onamı alınmıř olan katılımcıların depresyon tanısı almadan üç ay önceki nesnel Facebook hareketleri incelenmiř ve her katılımcı iin depresyonun varlıđı veya yokluđu, modeli sınavan algoritma tarafından tahmin edilen depresyon olasılıđı ile karřılařtırılmıřtır (Eichstaedt ve ark., 2018). Buna ek olarak Facebook dilinin bireylerin psikolojik iyi oluřuyla iliřkili olduđu; daha sık olumsuz duygu ifadesi kullanan kiřilerin depresyon ve kayđı düzeyleri yüksekken, daha ok olumlu duygu ifade edenlerin stres düzeylerinin görece daha düşük olduđu bilinmektedir (Settanni ve Marengo, 2015). Facebook özelinde yapılan bir diđer alıřmada da benzer řekilde Facebook’ta kullanılan olumsuz dilin ve uyku, rüya, yorgunluk, bedensel yakınmalar ieren ifadelerin genel yařam kalitesi ile negatif iliřkiler gösterdiđi; kullanılan Facebook dilinin bireylerin yařam kalitesini tahmin etmede bařarılı olduđu ifade edilmiřtir

(Marengo ve ark., 2021). Facebook dili aynı zamanda bireylerin problemleri alkol kullanımlarını da tahmin edebilmektedir (Marengo ve ark., 2019). Tüm bu çalışmalar hakkında bilinmesi gereken önemli bir nokta, büyük verinin herhangi bir davranışı tahmin ederken davranışı açıklamakla ilgilenmediği; davranış ve analiz edilen veriler arasındaki ilişkiye odaklanarak tahminde bulunduğudır. Psiko-informatik bu tahminin doğruluğu konusunda, elde edilen verilerin çokluğu nedeniyle istatistiksel açıklamalara güvenmektedir (Hinds ve Joinson, 2019).

Büyük verinin hayatımıza kazandırdığı kavramlardan biri de nesnelerin internetidir. Akıllı telefonlar, sosyal medya platformları, internet siteleri insanların günlük yaşamının bir parçası haline gelmiş ve insanlar buna çoktan alışmışken hızla gelişen teknoloji insan hayatına yeni akıllı cihazlar kazandırmaktadır. Akıllı bileklikler, akıllı kahve makineleri, akıllı buzdolapları ve içinde yaşanan akıllı evler interneti barındırmasıyla “Nesnelerin İnterneti” kavramını yaşamımıza dahil etmiştir (Miraz ve ark., 2015). Nesnelerin İnterneti, çeşitli bağlamlarda kullanılan nesnelerin sensörlerle donatılarak ortamdaki algılayabilmesine, internet ve çeşitli dijital platformlar aracılığıyla diğer nesnelerle ve insanlarla iletişim kurabilmesine olanak sağlayan çevrimiçi ulaşılabilir ve kontrol edilebilir çok büyük ölçekli bir cihaz ağıdır (Davies ve ark., 2019; Qazi, 2019). Bu düşünebildiğimiz ya da öngörebildiklerimizin çok ötesinde olan yeni bir dil, yeni bir iletişim kanalı gibi görünmektedir. Evans (2012), Nesnelerin İnterneti kavramıyla asıl vurgulanması gereken nesnelere daha çok insanların, süreçlerin, verilerin şeylerle oluşturduğu bağlantılar olduğunu ifade etmiştir. Bu bağlamda aslında “Her Şeyin İnterneti” olarak Nesnelerin İnterneti, kurduğu bağlantılar sayesinde gerçek zamanlı verilere ulaşması, bu verileri analiz edebilmesi ve aktarabilmesiyle kamu politikalarından çevresel sürdürülebilirliğe, ekonomi politikalarından sosyal bilimlere insana dair süreçlerde kullanılabilir (Bradley ve ark., 2013). Nesnelerin interneti ile bireyler, beslenme şekilleri, uyku düzenleri, spor alışkanlıkları, iletişim şekilleri, kullanılan dil ve yaşam biçimleriyle her geçen gün büyük verinin bir parçası haline gelmektedir. Bu doğrultuda insan duygu, düşünce ve davranışlarını anlama konusunda psiko-informatik giderek önem kazanmakta ama aynı zamanda beraberinde birçok etik problemi de gündeme getirmeye başlamaktadır.

Sosyal Bilimlerde Büyük Veri Tabanlı Enformatiğin Gelişiminde Etik Problemler ve Zorluklar

Yaşamın günden güne dijitalleşmesi ile bireylerin interneti, akıllı cihazları, sosyal medya platformlarını, arama motorlarını, e-posta hesaplarını kullanmasının neredeyse zorunlu olduğu düşünülmektedir. Bu durum kişileri isteyerek ya da istemeyerek birer veri haline getirmektedir. Çoğu zaman bunun için konuşmaya ya da yazmaya da gerek kalmamaktadır. Nitekim kişilerin söz konusu uygulamalara erişimi için, kişisel verilerin gizli tutulacağı taahhüdü ile verilerin mevcut uygulamaların geliştirilmesi ve kişiselleştirilmesi amacıyla paylaşılmasını kabul etmeleri gerekmektedir. Bu bağlamda en önemli etik problem verilerin gizliliğidir. Ancak insanların büyük bir çoğunluğu bu durum hakkında aşırı endişeli değildir.

Psiko-informatiğin kullanımı ile ortaya çıkan ve veri gizliliğini kapsayan bir diğer problem ise kişilerin anonim olarak dahi araştırmaların bir parçası olmak isteyip istemediğidir. Kullanıcılara buna dair bir bilgilendirilmiş onam sunulmamaktadır. Öte yandan insanların bir kısmının bu durumu kabul etmesi, kabul etmeyen kişilerin verilerinin de kullanılmasına yol açacak gibi görünmektedir. Nitekim özellikle sosyal medya platformlarından çekilen veriler yalnızca veri kullanımını kabul edenlerin değil onların etkileşimde oldukları kişilerin verilerini de içermektedir.

Gizlilikle bağlantılı olarak etik problem olarak değinilebilecek bir diğer konu, kaydedilen verilere kimlerin; ne amaçla erişebildiğidir. Psiko-informatikle birlikte sanal dünyada oluşturulan dijital izlerin insanların kişilik özelliklerini, davranışlarını tahmin edebileceği görülmektedir (Azucar ve ark., 2018; Hinds ve Joinson, 2019). Bu durum kişilerin bu özelliklerinden yola çıkarak, reklam şirketlerinin, politikaların ya da iktidarların kişileri manipüle etme potansiyelini düşündürmektedir (Montag ve arkadaşları, 2016). Bu manipülasyonun en büyük etki alanlarından biri de kullanılan dildir. Cambridge Analytica örneğinde olduğu gibi, size ait veriler günü geldiğinde sizin fikirlerinizi değiştirmek ve sizi çevrenizdeki diğer insanlardan uzaklaştırmak için kullanılabilir.

Psiko-informatiğin gelişiminde ortaya çıkan zorluklardan birisi de dijital dünyada özellikle sosyal medya platformlarında insanların kendilerini, duygu, düşünce ve davranışlarını ne ölçüde doğru yansıttığının bilinmemesidir. Bu doğrultuda özellikle sosyal medya gönderilerinin analizi sonucu toplanan veriler ile yapılan araştırma sonuçlarının yanlışlıklar taşıyabileceği düşünülmektedir. Ancak veri büyüklüğünün devasa boyutlarda olmasının bu yanlışlığı hafifleteceği şeklindeki görüşler ile de tartışma genişletilerek farklı noktalara taşınabilir. Ancak bunun için daha fazla çalışmaya, araştırma ve sonuca gereksinim olduğu anlaşılmaktadır.

Sonuç

İnternetin insan yaşamında edindiği yer ile bireylerin duygu, düşünce ve davranışları çeşitli platformlar aracılığıyla veri olarak kaydedilebilmektedir. İnsan- makine etkileşiminin hayatın her alanında olması, her an milyarlarca verinin kaydedilmesi anlamına gelmektedir. Bu devasa boyutta, yüksek hız ve çeşitlilikte ortaya çıkan büyük veri, bilgileri toplama, düzenleme ve analiz edebilme potansiyeli nedeniyle psikoloji araştırmaları için yeni bir disiplinin gelişmesine katkıda bulunmaktadır. Psiko-informatik denilen bu yeni disiplin psikoloji araştırmaları için platformlardan çekilen büyük verilerin temizlenmesi, analiz edilmesi ve kullanılması için bilgisayar bilimleri ve psikoloji biliminin iş birliği içinde çalışmasını gerektirmektedir (Yarkoni, 2012). Bu ihtiyacın karşılanması amacıyla sosyal bilimler lisans programlarına bilgisayar ve veri bilimlerinden dersler eklenmesinin faydalı olacağı düşünülmekte, bu şekilde öğrencilerin söz konusu bilimler ile sosyal bilimlerin iş birliğine dair bilgilendirilmesi önerilmektedir.

Yeni bir disiplinin ortaya çıkması, özellikle bu disiplinin internetle bu denli bağlantılı olması beraberinde etik problemleri de getirmektedir. Bu bağlamda psiko-informatik ve lingo-informatik çalışmalarında en önemli etik problem verilerin gizliliği, kişilere ait verilerin kullanımı konusunda kişilerin rızasının alınıp alınmayacağı ve kimlerin hangi amaçlarla verilere erişim sağlayabileceği konularıdır. Olası etik sorunların önüne geçmesi açısından bu konulara yönelik atılabilecek bir adım verilerin gizliliğini ve mahremiyeti sağlayan, verilere erişim ve verileri kullanıma dair amaçları yansıtan net kuralların ya da sınırların belirlenmesi olabilir. Bu amaç doğrultusunda psikoloji, dilbilim ve bilgisayar biliminden uzmanların bir araya gelmesi ile söz konusu problemlere yönelik ortak bir etik çerçeve çizilmesinin mümkün olabileceği düşünülmektedir. Öte yandan problemlerin tartışılıp, bu problemlere yönelik önlem alınması, o konunun bilinmesini ve gündeme gelmesini gerektirmektedir. Bu bağlamda mevcut makalenin psiko-informatiğe dair bilgilendirme ve ortaya çıkabilecek problemlere çözüm üretme amacına hizmet ettiği söylenebilir.

Bugün psiko-informatik çalışmalarına bakıldığında uluslararası alanyazında daha çok bireylerin kişilik özelliklerini; akıllı telefon, sosyal medya ve internet kullanım davranışlarını; psikolojik sağlıklarını ve problemleri davranışlarını tahmin etme üzerine çalışmalar olduğu görülmektedir. Ulusal alanyazın

tarandığında ise psiko-informatiğe dair herhangi bir çalışmaya rastlanmamıştır. Mevcut çalışma psiko-informatik konusunda kısa bir bilgilendirme sağlamaktadır. Psiko-informatiğin, özellikle insan davranışlarını tahmin etme konusunda araştırmacılara ve alanda çalışan psikoloji sağlık uzmanlarına farklı bir bakış açısı kazandıracağı düşünülmekte ve bu konuya çekilecek dikkat ile ileri çalışmaların planlanması ve gerçekleştirilmesi önerilmektedir.

Dilsel fenomenleri, dilin yapısını, kullanım kalıplarını devasa dilbilimsel verilere dayalı olarak inceleyen bir alan olarak büyük veri dilbilimi, veri toplama, kişisel verileri işleme, gizlilik, rıza beyanı, teknik zorlukların getirdiği potansiyel önyargılar gibi konularda hala eksiklikleri olan yeni bir alandır. Dil, düşünmeden bağımsız olarak işlenemez. Büyük verinin sadece dile ait somut araçları kullanarak düşünmeden bağımsız bir işleme modeliyle hareket etmesi bazı riskleri de beraberinde getirebilir. Büyük veri kümeleri işlenirken, konuşurların demografik özelliklerini yansıtan önyargılar, dil araştırmalarında çarpık veya hatalı sonuçlar alınmasına yol açabilir. Bu ve benzer sorunlar bütün sosyal bilim alanları için geçerlidir. Dilbilim, psikoloji, sosyoloji gibi alanlarda çalışan araştırmacıların, çalışmaların sağlıklı yürütülebilmesi için temel etik standart konusunda bilgilenebilmesi öncelikli görünmektedir.

Kaynaklar

- Acemoğlu, D. (2022). *Yapay Zekâyı Yeniden Tasarlamak*. (Çev. H. Dölkeleş). Efil Yayınevi.
- Andrews, S., Ellis, D.A., Shaw, H. ve Piwek, L. (2015). Beyond self-report: Tools to compare estimated and real-world smartphone use. *PLoS ONE*, 10(10). doi: <https://doi.org/10.1371/journal.pone.0139004>
- Azucar, D., Marengo, D. ve Settanni, M. (2018). Predicting the Big 5 personality traits from digital footprints on social media: A meta-analysis. *Personality and Individual Differences*, (124), 150-159. <https://doi.org/10.1016/j.paid.2017.12.018>
- Bradley, J., Reberger, C., Dixit, A. ve Gupta, V. (2013). Internet of Everything (IoE) top 10 insights from cisco's IoE value at stake analysis for the public sector. Cisco.
- Chen, M., Mao, S., Zhang, Y. ve Leung, V. (2014). Big data: Related technologies, challenges and future prospects (1.baskı). Springer.
- Davies, D., Beauchamp, G., Davies, J. ve Price, R. (2019). The potential of the 'Internet of Things' to enhance inquiry in Singapore schools. *Research in Science & Technological Education*, 38(4), 484-506.
- Doğan, K. ve Arslantekin, S. (2016). Büyük veri: Önemi, yapısı ve günümüzdeki durum. *DTCF Dergisi*, 56(1), 15-36.
- Eichstaedt, J.C., Smith, R.J., Merchant, R. M., Ungar, L.H., Crutchley, P., Preotiuc-Pietro, D., Asch, D.A. ve Schwartz, H.A. (2018). Facebook language predicts depression in medical records. *PNAS*, (115), 11203-11208. <https://doi.org/10.1073/pnas.1802331115>
- Endüstri 4.0 Platformu (2016, 1 Şubat). Davos'ta gündem: Endüstri 4.0. <https://www.endustri40.com/davosungundemi-endustri-4-0/>
- Ergen, Y. (2018). Büyük veri, sosyal medya ve etik: Facebook örneğinde bir değerlendirme. *Yeni Düşünceler*, (10), 53-64.
- Evans, D. (2012). The internet of everything: How more relevant and valuable connections will change the world ?. Cisco. https://www.cisco.com/c/dam/global/en_my/assets/ciscoinnovate/pdfs/IoE.pdf
- Greenfield, D.N. (1999). Psychological characteristics of compulsive internet use: a preliminary analysis. *CyberPsychology & Behavior*, (2), 403-412.

- Hinds, J. ve Joinson, A. (2019). Human and computer personality prediction from digital footprints. *Current Directions in Psychological Science*, 28(2), 1-8. <https://doi.org/10.1177/0963721419827849>.
- Katschnig, H. (2006). Quality of life in mental disorders: Challenges for research and clinical practice. *World Psychiatry*, 5(3), 139–145.
- Lin, Y.H., Lin, Y.C., Lee Y.H., Lin, P.H., Lin, S.H., Chang, L.R., Tseng, H.W., Yen, L.Y., Yang, C.C.H. ve Kuo, T.B.J. (2015). Time distortion associated with smartphone addiction: Identifying smartphone addiction via a mobile application (App). *Journal of Psychiatric Research* (65), 139-145. <https://doi.org/10.1016/j.jpsychires.2015.04.003>
- Marengo D., Azucar D., Giannotta F., Basile V. ve Settanni M. (2019). Exploring the association between problem drinking and language use on Facebook in young adults. *Heliyon*, (5), e02523. <https://doi.org/10.1016/j.heliyon.2019.e02523>
- Marengo, D. ve Montag, C. (2020). Digital phenotyping of big five personality traits via Facebook data mining: A meta-analysis. *Digital Psychology*, (1), 52-64. doi: <https://doi.org/10.24989/dp.v1i1.1823>
- Marengo, D., Azucar, D., Longobardi, C. ve Settanni, M. (2021). Mining Facebook data for quality of life assessment. *Behaviour & Information Technology*, (40), 597-607. <https://doi.org/10.1080/0144929X.2019.1711454>
- Marengo, D., Montag, C., Mignogna, A. ve Settanni M. (2022). Mining digital traces of Facebook activity for the prediction of individual differences in tendencies toward social networks use disorder: a machine learning approach. *Frontiers in Psychology*, (13). doi: <https://doi.org/10.3389%2Ffpsyg.2022.830120>
- Marengo, D., Poletti I., Settanni M. (2020). The interplay between neuroticism, extraversion, and social media addiction in young adult Facebook users: testing the mediating role of online activity using objective data. *Addictive Behaviors*, (102). doi: <https://doi.org/10.1016/j.addbeh.2019.106150>
- Marino C., Finos L., Vieno A., Lenzi M. ve Spada M. M. (2017). Objective Facebook behaviour: differences between problematic and non-problematic users. *Computers in Human Behavior*, (73), 541–546. doi: <https://doi.org/10.1016/j.chb.2017.04.015>
- Markowetz, A., Błaszczewicz, K., Montag, C., Switala, C. ve Schlaepfer, T. E. (2014). Psycho-informatics: Big data shaping modern psychometrics. *Medical Hypotheses*, (82), 405-411.
- Miraz, M.H., Ali, M. Excell, P.S. ve Picking, R. (2015). A review on Internet of Things (IoT), Internet of Everything (IoE) and Internet of Nano Things (IoNT). *2015 Internet Technologies and Applications (ITA)*, 219-224. doi: 10.1109/ITechA.2015.7317398
- Montag, C., Błaszczewicz, K., Lachman, B., Sariyska, R., Andone, I., Trendafilov, B. Ve Markowetz, A. (2015). Recorded behavior as a valuable resource for diagnostics in mobile phone addiction: Evidence from psychoinformatics. *Behavioral Sciences*, 5(4), 434-442. <https://doi.org/10.3390/bs5040434>.
- Montag, C., Duke, E. ve Markowetz, A. (2016). Toward psychoinformatics: Computer science meets psychology. *Computational and Mathematical Methods in Medicine*, Makale 2983685. <https://doi.org/10.1155/2016/2983685>.
- New York Times (2018, 4 Nisan). Facebook says Cambridge Analytica harvested data of up to 87 million users. <https://www.nytimes.com/2018/04/04/technology/mark-zuckerberg-testify-congress.html>
- Peterka-Bonetta J., Sindermann C., Elhai J. D. ve Montag C. (2021). How objectively measured Twitter and Instagram use relate to self-reported personality and tendencies toward Internet/Smartphone Use Disorder. *Human Behavior and Emerging Technologies*, (3), 898–911. doi: <https://doi.org/10.1002/hbe2.243>
- Qazi, A. (2019, 30 Mayıs). How will IoT change the education sphere?. *EmergingEdTech*. <https://www.emergingedtech.com/2019/05/how-will-iot-change-impact-education/>

- Schwab, K. (2016.) The fourth industrial revolution. <https://www.weforum.org/about/the-fourth-industrial-revolution-by-klaus-schwab>
- Schwartz, H.A., Eichstaedt, J.C., Kern, M.L., Dziurzynski, L., Ramones, S.M., Agrawal, M., Shah, A., Kosinski, M., Stillwell, D., Seligman, M.E.P. ve Ungar, L.H. (2013). Personality, gender, and age in the language of social media: the open-vocabulary approach. *PLoS ONE*, 8(9). doi: <https://doi.org/10.1371/journal.pone.0073791>
- Seçer, İ. (2018). Psikolojide araştırma yöntemler. Işık, Ş. (Ed.). *Psikolojiye giriş* (25-46). doi:10.14527/9786052414132.02
- Settanni M. ve Marengo D. (2015). Sharing feelings online: studying emotional well-being via automated text analysis of Facebook posts. *Frontiers in Psychology*, (6). doi: <https://doi.org/10.3389/fpsyg.2015.01045>
- Settanni, M., Azucar, D. ve Marengo, D. (2018). Predicting individual characteristics from digital traces on social media: a meta-analysis. *Cyberpsychology, Behavior, and Social Networking*, (21), 217-228. doi: <http://doi.org/10.1089/cyber.2017.0384>
- Simon, H. A. ve Newell, A. (1958). Heuristic problem solving: The next advance in operations research. *Operations research*, 6(1), 1-10. doi: <https://doi.org/10.1287/opre.6.1.1>
- The Guardian (2018, 10 Nisan). How to check whether Facebook shared your data with Cambridge Analytica. <https://www.theguardian.com/technology/2018/apr/10/facebook-notify-users-data-harvested-cambridge-analytica>
- World Economic Forum (2016). *Annual Meeting 2016 Mastering the Fourth Industrial Revolution*, 14. https://www3.weforum.org/docs/WEF_AM16_Report.pdf?_gl=1*ciz32p*_up*MQ..&gclid=CjwKCAjw5_GmBhBIEiwA5QSMxE4GMaVfnrptrmr9UkPRoSX-IXV77_d1krV4_3lqVarzQ9xeJU-G6RoCu2EQAvD_BwE
- Yarkoni, T. (2012). Psychoinformatics: New horizons at the interface of the psychological and computing sciences. *Current Directions in Psychological Science*, 21(6), 391–397.
- Zadrozny, P. ve Kodali, R. (2013). Big data analytics using splunk. Apress.