

TÜRKİYE'DE BÜYÜK VERİ VE VERİ MADENCİLİĞİNE İLİŞKİN POLİTİKA VE STRATEJİLER: ULUSAL POLİTİKA BELGELERİNİN İÇERİK ANALİZİ

POLICIES AND STRATEGIES OF BIG DATA AND DATA MINING IN TURKEY: A CONTENT ANALYSIS OF NATIONAL POLICY DOCUMENTS

Özer KÖSEOĞLU*, Yılmaz DEMİRCİ**

* Doç. Dr., Sakarya Üniversitesi Siyasal Bilgiler Fakültesi Siyaset Bilimi ve Kamu Yönetimi Bölümü,
ozerk@sakarya.edu.tr

** Doktora Öğrencisi, Sakarya Üniversitesi Sosyal Bilimler Enstitüsü, yilmazdemirci@gmail.com

ÖZ

Bilişim teknolojilerinde yaşanan son değişim birbiriyle ilişkili olan büyük veri, veri madenciliği, makine öğrenmesi, Endüstri 4.0 ve açık veri gibi kavramlarla ifade edilmektedir. Özellikle büyük veri ve veri madenciliği uygulamaları daha şimdiden pazarlamadan ürün tedarikine, planlamadan mal ve hizmet üretimine kadar birçok alanda önemli değişimleri ortaya çıkarmıştır. Büyük veri uygulamaları ve veri madenciliği projeleri, son yıllarda kamu sektörü kuruluşlarında da yaygınlaşmaktadır. Bu uygulamalar kamu hizmetlerinin dağıtımını, kamu kurumlarının yönetimini ve kamu politikalarının oluşturulması ve uygulanması süreçlerini geleneksel yönetim anlayışından olduğu kadar e-devletten de farklılaştırmaktadır. Bu makalede, Türkiye'de büyük veri ve veri madenciliğinin kamu hizmetleri ve politikalarına olan etkilerini analiz etmek amacıyla temel üst politika belgeleri ile bakanlıkların stratejik planları içerik analizi yöntemiyle incelenmiştir. Sonuç olarak, üst politika belgelerinde ve bakanlıkların nezdinde büyük veri ve veri madenciliğine yönelik bir farkındalığın olduğu, ancak bu konuda daha kapsamlı ve bütüncül bir strateji ve eylem planına ihtiyaç bulunduğu tespit edilmiştir.

Anahtar Sözcükler: Büyük Veri, Veri Madenciliği, Makine Öğrenmesi, Açık Veri, Endüstri 4.0

Jel Kodları: O3, O33, M15, L63.

ABSTRACT

Recent developments in the field of knowledge technologies are expressed with concepts such as big data, data mining, machine learning, open data, and Industry 4.0. Among these big data and data mining practices have already started a substantive change in the fields of marketing, planning, and provision of goods and services. Big data practices and data mining projects, in recent years, have begun to be implemented in public sector organisations as well. These projects and implementations have been differentiated provision of public services, management of public organisations, developing and implementing public policies not only from traditional public administration but also from e-government approach. In this study, national policy documents and ministry strategic plans have been analysed through content analysis method to comprehend the impact of big data and data mining on public services and policies in Turkey. As a conclusion, national policy documents and ministry policies have revealed that there is an awareness of big data and data mining practices, however it is necessary to develop a comprehensive and integrated strategy and action plan for effective and sustainable implementation.

Keywords: Big Data, Data Mining, Machine Learning, Open Data, Industry 4.0

Jel Codes: O3, O33, M15, L63.

1. GİRİŞ

Sanayi toplumundan bilgi toplumuna geçiş, teknolojik anlamda yaşanan gelişmelerin de katkısıyla verinin üretilmesi, depolanması, işlenmesi ve paylaşılmasını hızlandırarak bilginin öneminin artmasını sağlamıştır. Bilgi teknolojilerinde yaşanan hızlı ilerleme insanların gündelik hayatından kamu ve özel sektör örgütlerinin iş yapma süreçlerine, sağlık ve eğitim gibi kamusal hizmetlerin sunumundan yeni uzmanlık alanları ve mesleklerin ortaya çıkmasına kadar birçok değişime kapı aralamıştır.

Bilginin yönetsel süreçlerde kullanılması, bilgi ve iletişim teknolojilerine adaptasyon örgütler arasında olduğu kadar ülkeler arasında da karşılaştırmalı rekabet üstünlüğü sağlayan bir unsur olarak görülmektedir. Bu nedenle gerek örgütler için gerekse ülkeler için bilgi ve iletişim teknolojilerine yapılacak olan yatırımlar stratejik bir öneme sahip hale gelmiştir.

Günümüzde geleneksel donanım ve yazılım çözümleriyle saklanması ve analiz edilmesi maliyetli olan büyük hacimli verileri gelişen bilgi teknolojileri, donanım ve yazılım çözümleri sayesinde saklamak, işlemek, paylaşmak ve analiz etmek mümkün hale gelmiştir. Devletler ve kamu örgütleri 1990'lı yıllarda geleneksel yolların yanında işletme tekniklerini ve e-devletin sağladığı imkanları kullanarak işlem ve prosedürlerini değiştirmiş, kamusal mal ve hizmetleri sunmuş, kamu politikaları geliştirmiş ve uygulamıştır. E-devlet ile başlayan devletin bilgi teknolojilerine adaptasyon süreci günümüzde sosyal medya ve Web 2.0, açık kaynak yazılımı, büyük veri, makine öğrenme ve açık veri gibi yeni meydan okumalarla karşı karşıyadır. Bu makalede, yukarıda bahsedilen gelişmelerle çok yakın ilişkisi bulunan, ancak dijital dönüşümün başka bir boyutunu teşkil eden “büyük veri” ve büyük verinin işlenmesine ve analizine imkân sağlayan “veri madenciliği” konu edilmektedir.

Veri madenciliği bankacılık, pazarlama, bilişim, telekomünikasyon ve borsacılık gibi sektörlerde gelişmiş bir uygulama

ağına sahip olmakla birlikte devletler ve kamu sektörü için nispeten daha yeni bir olgudur. Genellikle risk analizi ve usulsüzlük tespiti, müşteri kazanma, kredi kartı dolandırıcılıklarının tespiti, müşteri kaybı sebeplerinin belirlenmesi, hile tespiti, hatların yoğunluk tahminleri, tıbbi teşhis ve uygun tedavi süreçlerinin belirlenmesi gibi birçok alanda veri madenciliği uygulamalarına rastlanmaktadır.

Türkiye’de de büyük veri ve veri madenciliğine hükümet programları, kalkınma planları ve ulusal eylem planları gibi üst politika belgelerinde yer verilmektedir. Ayrıca kurumsal politika ve planlarda da bu kapsamda hedef, strateji ve faaliyetler ortaya konulmaktadır. Alanyazın açısından değerlendirildiğinde kavramsal ve kuramsal çalışmalar (Erbay ve Kör, 2016; Altunışık, 2015; Işıklı, 2014) olduğu kadar işletmelerde büyük veri uygulamalarına odaklanan araştırmalar (Çakırel, 2016; Atan, 2016; Özbilgin, 2015) da bulunmaktadır. Ayrıca verinin gizliliği ve kişi mahremiyeti gibi büyük veri ve veri madenciliği uygulamalarının ortaya çıkardığı hukuki ve etik sorunları inceleyen çalışmalara (Eyüpoğlu vd., 2017) da rastlanmaktadır. İlgili alanyazın sorgulandığında bu kapsamdaki çalışmaların özel sektör işletmeleri üzerinde yoğunlaştığı, veri madenciliğinden ziyade büyük veri odaklı yürütüldüğü ve son yıllarda bu kapsamdaki çalışmaların sayı olarak artış eğiliminde olduğu söylenebilir. Bu alanın oldukça dinamik ve gelişime açık olduğu anlaşılmaktadır. Ancak yine de kamu hizmetleri ve politikalarına ilişkin çalışmaların sağlık gibi belirli alanlarla sınırlı kaldığı görülmektedir. Bu nedenle elinizdeki makale, kamu kurum ve kuruluşlarına veri madenciliği açısından odaklanarak yazında var olan eksikliğe katkı yapmayı hedeflemektedir.

Bu bağlamda, makalenin amacı Türkiye’de hükümet politikaları, kalkınma planları ve eylem planları gibi üst politika belgeleri ile bakanlıkların stratejik planlarının “içerik analizi” yöntemiyle analiz edilerek büyük

veri ve veri madenciliğinin kamu politikaları ve hizmetlerine olan etkilerini değerlendirmektedir. Böylece özellikle bakanlıklar nezdinde geliştirilen iyi uygulama örnekleri de tespit edilmiş olmaktadır. Bu amaçla makalenin birinci bölümünde kamu yönetiminde yaşanan dijital dönüşümün izi sürülmektedir. İkinci bölümde veri madenciliği ve ilgili kavramların açıklanması yapıldıktan sonra üçüncü ve dördüncü bölümde veri madenciliği türleri, uygulama süreci, kamu ve özel sektörde başlıca kullanım alanları değerlendirilmektedir. Son bölümde ise Türkiye'de veri madenciliği ve büyük veriye ilişkin politika, hedef ve stratejiler çeşitli politika belgeleri kullanılarak analiz edilmektedir.

2. KAMU HİZMETLERİNDE DİJİTAL DÖNÜŞÜMÜN GELİŞİMİ: E-DEVLETEN BÜYÜK VERİ VE VERİ MADENCİLİĞİNE

Kamu hizmetlerinin sunumu ve uygulanacak kamu politikalarına karar verilmesi 1980'li yıllardan itibaren idari reform dalgasıyla birlikte yeniden şekillenmiştir. Örgütsel yapı ve süreçlerde özel sektörün esas alınması neticesinde kamu örgütleriyle vatandaş arasında piyasa mekanizmalarına benzer bir ilişki kurularak yenilikçi hizmet yöntemleri uygulanmaya başlanmıştır. Yaşanan değişim kamu yöneticilerini hesap verebilir olmaya, kamu örgütlerini ise verimlilik artışına yöneltmiştir. Aynı zamanda örgütsel desantralizasyonun sağlanması, kamu yöneticilerinin performans hedeflerinden sorumlu tutulması ve kamu örgütlerinin yeniden yapılandırılması gündeme gelmiştir (Denhardt ve Denhardt, 2000: 550).

Kamu sektöründe profesyonel yönetim, açık performans standartları ve ölçütleri, çıktıların kontrol edilmesine odaklanma, rekabete geçiş, özel sektör yönetim tekniklerinin uygulaması, kaynak kullanımında verimlilik ve disiplin 1980'li yıllarda başlayan ve "yeni kamu işletmeciliği" (YKİ) olarak adlandırılan reform dalgasının temel ilkelerini

oluşturmuştur (Hood, 1991: 4-5). Hükümetler ve kamu örgütleri dünyanın dört bir yanında söz konusu ilkeleri kamu politikalarına ve reform uygulamalarına dönüştürmüştür. Dünya Bankası ve Uluslararası Para Fonu gibi uluslararası kuruluşlar ise işletmecî reformları tavsiye ve teşvik etmiş, böylece yeni kamu işletmeciliği küresel ölçekte yayılım göstermiştir.

1990'lı yıllarda ise bilgi ve iletişim teknolojilerindeki gelişmeler kamu sektöründe yeni bir reform dalgasının habercisi olmuştur. Aslında teknolojideki hızlı ilerlemeler hem özel sektör hem kamu sektörü yönetiminde değişimin sürükleyici motoru olmuştur. Bilgisayar sistemleri gibi bilgi ve iletişim sistemlerinin kullanımı, internet ve yeni veritabanları gibi yenilikler kamu sektöründe "e-devlet" olarak adlandırılan eğilime önyak olmuştur (Hughes, 2003: 13-14). YKİ'nin ortaya çıktığı ülkelerden biri olan ABD'de 2000'li yılların başında Başkan Bush verimlilik artışı yoluyla maliyetleri azaltmaya yöneldiğinden bilgi teknolojilerinden daha fazla yararlanma politikası takip edilmiştir. İngiltere'de ise bilgi ve iletişim teknolojilerinin kullanımı, 1999 yılında iktidara gelen İşçi Partisi hükümetinin temel amaçlarından biri olmuştur. Bu dönemde Avrupa Birliği kapsamında da tartışmalar, yeni bilgi ve iletişim teknolojilerinin demokratik potansiyelinden çok etkinlik, verimlilik ve hizmet sunumuna odaklanmıştır. Diğer gelişmiş ve gelişmekte olan ülkelerde de yaygınlaşmasıyla birlikte 2000'lerin başına gelindiğinde e-devlet uygulamaları uluslararası bir nitelik kazanarak tüm dünyaya yayılmıştır (Chadwick ve May, 2003: 272). Bu dönemde kamu örgütleri e-posta, web sitesi yönetimi, online işlem yapma gibi uygulamaları adapte etmiş ve artık web tabanlı hizmetler devletin ayrılmaz bir parçası haline gelmiştir (Ho, 2002: 434). E-devlet, 21. yüzyılın ilk çeyreğinde birey-devlet ilişkilerinde yaşanacak olan dönüşüme önyak olmuştur. E-devlet kapsamında verilen hizmetler, paylaşılan veriler ve gerçekleştirilen

işlemler dolaşımdaki verinin miktarını ciddi ölçüde artırmıştır. Bu dönemde akıllı telefonlar, tablet bilgisayarlar gibi teknolojilerin gelişmesi ve internetin kullanımının yaygınlaşması devletten beklentileri de değiştirmiştir.

Özelleştirilmiş forum sayfaları, sohbet odaları, gruba e-posta gönderme gibi teknolojiler toplu ve özel iletişim imkanlarını artırmıştır (Miller vd., 2016: 2). İnternete bağlı bir şekilde ortalama bir kullanıcının katılım ve bilgisini paylaşabileceği yeni nesil internet uygulamaları olarak adlandırılan Web 2.0 teknolojileriyle birlikte kullanıcılar herhangi bir içeriğe katkıda bulunma ve mevcut içeriği güncelleyerek yeniden yayımlama imkanına kavuşmuş, internet statik yapısından dönüşerek sosyal ağlar ve kullanıcılar arası işbirliğini içeren dinamik bir yapıya kavuşmuştur (Thompson: 2008: 711). Web 2.0 teknolojilerine dayanan sosyal medya araçlarının, 2000'li yılların başından itibaren giderek kullanıcı sayısını artırmasıyla birlikte özel sektör ve kamu örgütleri bu alandaki gelişmelere kayıtsız kalamamıştır. Özel sektör kuruluşları pazarlama, reklam ve satış gibi iş süreçleri için sosyal medyayı kullanarak rekabet avantajı elde etmeye başlamıştır. Kamu örgütleri ise kurumsal sosyal medya hesaplarını ağırlıklı olarak bilgi paylaşma, haber verme, etkileşime geçme ve kısmen de katılım ve işbirliği kurma gibi amaçlar doğrultusunda kullanmaya başlamıştır. Böylece kamu hizmetlerinin daha etkili ve hızlı sunulması, yurttaşların ise kamusal kararların alınması ve politikaların oluşturulması süreçlerine daha etkin katılımı imkânı gelişmiştir.

2009 yılında ABD'de Başkan Obama tarafından yayımlanan bir memorandumun ardından şeffaflık, katılım ve işbirliğine dayanan ve dijital dönüşümün getirdiği olanakların yönetim süreçlerine eklemlenmesini savunan açık yönetim (*open government*) ve açık veri (*open data*) politikaları gelişmiştir. Yönetimde açıklık yeni bir kavram olmamakla birlikte, günümüz açık yönetim politikalarının öncekilerden farkı şeffaflık, katılım ve

işbirliğini içerecek biçimde daha bütüncül olması ve ayrıca bilişim teknolojileri, internet ve sosyal medya gibi araçların entegre edilmesiyle birlikte daha etkili bir biçimde kullanılmasıdır. Açık yönetimin bu hedefi, onu online işlem ve hizmet sunumunu önceleyen klasik e-devlet uygulamalarından da farklılaştırmaktadır (Harrison vd., 2012: 84). 2011 yılında kurulan Açık Yönetim Ortaklığı (*Open Government Partnership*) gibi uluslararası oluşumların da etkisiyle açık yönetim kısa süre içinde çok sayıda ülke arasında yaygın bir politika haline gelmiştir. Açık yönetimin gelişimini destekleyen yenilikler arasında açık veri ve açık kaynak yazılımlar bulunmaktadır. Açık veri, ulusal ve yerel düzeydeki yönetimlerin çok sayıda verinin kullanım kolaylığı içinde kullanıcılara sağlanmasında yardımcı olmuştur. Açık kaynak kodlu yazılımların yaygınlaşması ise büyük miktarda verinin işlenmesi ve analizi için gerekli olanakları sağladığı gibi veri erişiminde kullanılacak programların ihtiyaca yönelik şekilde yeniden dizayn edilmesine imkân vermiştir.

Son yıllarda ise internetin ve ağların yaygın kullanımı sayesinde büyük miktarda veriyi toplayabilecek ve işleyebilecek teknolojilerin gelişmesiyle birlikte yeni bir meydan okuma ile karşı karşıya kaldığımızı söyleyebiliriz. Halihazırda belirli kamu hizmetlerinin sunulmasında, fabrikaların üretim süreçlerinin yönetiminde ve bazı özel sektör hizmetlerinin sunulmasında Endüstri 4.0 olarak adlandırılan uygulamalara rastlanmaktadır. Makine sensorlerinin entegrasyonu, yazılımlar, bulut bilişim ve depolama sistemleri gibi yenilikler gelişmiş analitiğe dayalı olarak büyük veri setlerinin sorgulanmasından elde edilen sonuçları bir geri besleme olarak örgütün operasyonel süreçlerini yürütmek için bir yöntem olarak kullanma fırsatını ortaya çıkarmıştır (Gilchrist, 2016: 4). Böylece makineler ve kullandığımız araba, buzdolabı, fabrika robotları gibi şeyler (things), işlevlerini yerine getirirken topladıkları verileri analiz etmekte, sorunları öngörerek hızlıca çözebilmekte,

ileri düzeyde kişiselleştirilmiş ürün ve hizmet anlayışını geliştirmektedir.

Makinelerin öğrenmesi ve eşyaların interneti deyimleriyle ifade edilen dönüşümün kamu hizmetleri ve politikaları üzerinde oluşturacağı etki ise henüz kapsamlı olarak ele alınmış değildir. Tüm bu gelişmelerin derinlemesine incelenmesi bu makalenin sınırlarını aşmaktadır. Ancak büyük verinin işlenmesinde kritik role sahip olan veri madenciliği konusu ve veri madenciliğinin kamu hizmetlerine olan etkilerinin değerlendirilmesi yoluyla bu makalenin ilgili yazına bir katkı sağlayacağı düşünülmektedir.

3. VERİ MADENCİLİĞİNİN TANIMLANMASI VE KAVRAMSAL ÇERÇEVE

Örgütsel amaçlara bağlı olarak işlemlerin yapılandırılmamış bir şekilde kaydedilmesine “veri” adı verilir (Barutçugil, 2002: 57). Veri, kavramsal bir çerçeve ve yapı içerisinde bulunan bir tür enformasyondur (Aktan ve Vural, 2005: 6). Bu durumda veri, işlenmemiş enformasyon parçacıkları iken, enformasyon organize edilmiş veri seti, bilgi ise anlamlı olan enformasyonlara işaret etmektedir (Bhatt, 2001: 68-75).

Geleneksel veriden farklı olarak büyük hacimli (volume), yüksek hızlı (velocity) ve çok çeşitlikteki (variety) veri setlerine “büyük veri” adı verilir (Foster vd., 2017: 3). Günümüzde sağlık verileri, alışveriş verileri, borsa ve finans verileri, sosyal medya verileri, kamusal işlemler sonucu üretilen veriler, özel sektör girişimleri neticesinde sağlanan veriler ve internetin kişisel kullanımı gibi yollarla oluşturulan veriler oldukça büyük çapta olup, gelişen bilişim teknolojileri, donanım ve yazılım çözümleri sayesinde veritabanları, veri ambarları ve bulut bilişim kaynaklarında saklanabilmektedir.

Saklanan veriler geleneksel yöntemler kullanılarak insan yeteneğiyle analiz edilemez büyüklüklere ulaşmıştır. Bu nedenle özel yazılımlar ve programlar

kullanılarak çözümlenmek gerekmektedir. Bu kapsamda Google tarafından geliştirilmiş GFS dosya sistemi ve MapReduce programlama tekniği, basit sunucular üzerinden büyük verileri işlemek amacıyla kullanılacak uygulamaları çalıştırmak üzere Hadoop ekosistemi ortaya çıkmıştır. Hadoop Distributed File System (HDFS) olarak adlandırılan bir dağıtık dosya sistemi ve MapReduce özelliklerini bir araya getiren, Java ile geliştirilmiş açık kaynaklı bir altyapıdır. Hadoop, HDFS ve MapReduce bileşenlerinden oluşan bir yazılımdır. Büyük veri üzerinde uygun yazılımlar, algoritmalar ve teknolojiler kullanılarak toplanan verilerin işlenmesi, aralarındaki ilişkinin tespiti, birlikteliklerin belirlenmesi ve örüntülerin çıkarılması mümkün hale gelmiştir. Bu noktada Spark, Storm, FLink, Apache Ranger, Knox Gateway, Kafka, R, Cascading, Elasticsearch, HBase, Hive, Pig, Cassandra ve MongoDB olmak üzere programlama dilleri ve NoSQL veritabanı sistemleri geliştirilmiştir.

Büyük verinin analiz edilmesinde önemli bir araç olan veri madenciliği, 1990'lı yıllardan itibaren veri depolama araçları, barkod ve sensor teknolojilerinin gelişmesiyle beraber hem uygulamada hem de yazında daha sık kullanılır hale gelmiştir. Veri madenciliği, daha önceden bilinmeyen, geçerli ve uygulanabilir; ancak büyük miktardaki verinin geniş veri tabanlarından elde edilmesi, verinin değerli bilgiye dönüştürülerek karar verme süreçlerinde kullanılması olarak tanımlanabilir (He vd., 2013: 123; Silahtaroglu, 2013: 12). Deyim ilk olarak 1978 yılında Leamer tarafından, klasik istatistik yöntemlerle çözülmesi zor problemlerin doğrusal regresyon kullanılarak çözülebileceğini göstermek için kullanılmıştır. Ancak günümüzde istatistiğin yanında veri analizi ve yönetim bilişim sistemleri gibi bilim dallarında çalışanlar tarafından da kullanılmaktadır (Usama ve Paul, 1997: 102). Ayrıca sosyal bilimlerin çeşitli dallarındaki uzmanlar veri madenciliğinin bir türü olan metin madenciliği tekniklerini araştırma

süreçlerine entegre ederek toplumsal sorunları incelemektedir.

Yapılandırılmamış şekilde elde edilen veriler veri tabanlarında toplanarak yapılandırılmış veriler halinde dönüşmekte, tablolar halinde belirli bir formatta saklanmaktadır. Veri madenciliği çalışmaları kurumların veri tabanlarında ve veri ambarlarında yer alan yapılandırılmış veriler üzerinden tanımlama ve tahmine yönelik hedeflerle gerçekleştirilmektedir. Büyük veri ise genellikle birbirinden farklı, çoklu ilişkisiz veri kümelerinin birleştirilmesi neticesinde oluşan büyük miktarda yapısal olmayan veriye işaret etmektedir. Bu verinin işlenmesi, gizli enformasyonun kısıtlı zaman içinde toplanması ve büyük veriye has kodlama teknikleri ile çözümlenmesini gerektirmektedir.

4. VERİ MADENCİLİĞİNİN TÜRLERİ VE UYGULAMA SÜRECİ

4.1. Veri Madenciliğinin Türleri

Veri madenciliğinin tahmin etme ve tanımlama olmak üzere iki temel amacı bulunmaktadır. *Tahmin*, veritabanındaki bazı değişkenleri kullanarak bilinmeyeni öngörmeyi veya gelecekteki değerlerini hesaplamayı içerir. *Tanımlama* ise, veri içindeki örüntüleri meydana çıkarmayı amaçlar. Tahmin etmeye dayanan modeller, “bir işlemde dolandırıcılık olup olmadığı” veya “bir müşteriden ne kadar gelir elde edilebileceği” gibi soruların yanıtlarını ararken, tanımlamaya dayalı modeller veriler arasında gizli kalmış örüntülerden “çocuk bezi alan bir müşterinin mama alma olasılığı diğer müşterilerden üç kat fazladır” gibi sonuçlar çıkarmayı hedefler (Silahtaroglu, 2013: 49).

Tanımlayıcı veri madenciliği (*descriptive data mining*), birliktelik kuralları ve kümeleme gibi veri içindeki belli başlı ilişkileri bulmayı amaçlar. Tahmin edici veri madenciliği (*predictive data mining*), tarihsel verilerden yola çıkarak belirlenmiş hedef değişkenin değerini öngörmeye çalışır. Bu hedef ayrık olduğunda, bu

işlemin sınıflandırma ile yapılması tercih edilir. Kredi başvurularında ödeme koşullarının tahmin edilmesi, müşterilerin sigorta üzerinden talep ettikleri alacakların birer dolandırıcılık işlemi olup olmadığı bu tarz işlemlere örnek verilebilir. Aynı zamanda, devamlı değişen hedefin tahmin edilmesi regresyonun görevidir (Baesens vd., 2009: 16).

Veri madenciliği uygulamalarının kullanıldığı iki alan metin madenciliği ve web madenciliğidir. *Metin madenciliği*, büyük kapasiteli dokümanların analiz edilerek metin tabanlı verilerden gizli bilgiye ulaşılmasıdır. Metin madenciliği e-postalar, raporlar, konuşma dökümleri, saha notları ve makaleler gibi metin içerikli veriler üzerinde işlemler gerçekleştirilmesini amaçlar. *Web madenciliği* ise yapılandırılmamış web içerikleri, linkler ve web erişim istatistikleri gibi internet üzerindeki kalitatif ve kantitatif verilerin analiz edilmesiyle ilgilenir (Tan ve Yu, 2003: 239)

4.2. Veri Madenciliği Teknikleri

Veri madenciliği birliktelik, kümeleme ve sınıflandırma olmak üzere üç işlemi içerir. *Sınıflandırma*, verileri sınıflandırmak için kullanılan veri analizi modelidir. Bir tıbbi araştırmacı, hastasının hangi tedaviyi alması gerektiğini tahmin edebilmek için kanser verilerini analiz etmek istediğinde, kredi başvurularının “güvenli veya “riskli” olarak belirlenmesi için bir model gerektiğinde veri analiz işlemi için sınıflandırmaya ihtiyaç duyulacaktır (Han vd., 2012: 328). *Kümeleme*, birbiri ile yüksek benzerlik gösteren nesnelerin birden çok grup veya kümede toplanması, benzerlik göstermeyen nesnelerin ise diğer kümelerde toplanması işlemidir (Han vd., 2012: 443). Eldeki veri setleri üzerinden olayların birlikte gerçekleşme durumlarını inceleyen tekniğe *birliktelik kuralları* adı verilir. Birliktelik kuralları yoluyla veriler arasındaki ilginç ilişkilerin tespit edilmesi amaçlanır. Elde edilen verinin her geçen gün artması neticesinde kurumlar bu verilerden yararlanmak suretiyle birliktelik ilişkilerini keşfetmek ve karar alma

süreçlerinde bu bilgiden yararlanmak istemektedir (Özekes, 2003: 76).

4.3. Veri Madenciliği Uygulama Süreci

Veri madenciliği uygulamaları için tanımlanması, verinin tanımlanması, verinin hazırlanması, modelleme, değerlendirme ve uygulama olmak üzere CRISP-DM metodolojisini esas alan 6 aşamadan meydana gelmektedir (Chapman vd., 2000: 10). İlk aşamada örgütün hedeflerine yönelik alınacak kararlarda ulaşılması istenen hedeflerin iyi anlaşılması, belirlenen hedeflere ulaşmak için ise problemlerli sahaların tespiti aşaması *işin tanımlanması* aşamasıdır. Bu aşamada, problemlerli sahalar veri madenciliği uygulamalarında kullanılmak üzere problem tanımına dönüştürülür. Neyi ölçmeye veya öngörmeye çalıştığımıza karar vermek, amacı açıkça belirlemek ve amaca giden süreci geliştirmek gerekir.

Veri madenciliğinde problem tespiti sonrası, işlenmek için problemlerli sahaya ait *verilerin hazırlanması* gerekmektedir. Veriyi anlama olarak tanımlanan bu aşamada veri madenciliğinde kullanılacak veriler ilişkisel veritabanları (*RDBMS, Relational Database Managment System*), veri ambarları (*datawarehouses*), işlemsel veritabanları (*transactional databases*), uzamsal veritabanları (*spatial databases*), metin veritabanları, multimedya veritabanları ve internet olmak üzere farklı kaynaklardan elde edilebilir.

Veri hazırlanması aşamasında eldeki veri yığımından küçük hacimli bir örnek elde edilerek veri madenciliği uygulamalarından etkili sonuçlar elde edilmeye çalışılır (Oğuzlar, 2003: 74). Veri madenciliği uygulamalarında elimizdeki veriler, kullanacağımız algoritma ile uyumlu olmayabilir. Bazı algoritmalar sadece sayısal değerler ile çalışırken, bazıları kategorik değer kullanabilir. Veri ön işleme aşaması, veri madenciliğinin en önemli aşamasıdır. Modelleme aşamasında verinin yeniden düzenlenmesi ihtiyacı doğabilir. Veri ön işleme (indirgeme) yöntemleri, veri birleştirme veya veri küpü (*data aggregation, data cube*), boyut indirgeme

(*dimension reduction*), veri sıkıştırma (*data compression*) ve kesikli hale getirmeden (*discretization*) oluşmaktadır.

Veri madenciliği; sınıflandırma, kümeleme gibi teknikleri iyi yapılandırılmış bir teknoloji alanı haline gelmiştir. Bu alanda yeni algoritmalar üretilmesi ve var olan algoritmaların iyileştirilmesi yönelik fikirler geliştirilmeye devam edilmektedir. Veri madenciliği, veri setlerinden kullanacağı *modele karar verme* ya da *modeli veriye uydurma* ile ilgilendirilir. Kullanımına karar verilen modeller, tüm interaktif veriden bilgi elde etme süreci neticesinde ulaşılabilecek kullanışlı bilgiyi etkiler ve kullanıcının karar vermesine yardımcı olur (Ralha ve Silva, 2012: 11645). Uygun modelleme tekniği seçilerek hayata geçirildikten sonra, sonuçları optimize etmek için model ayarları gözden geçirilir. Veri madenciliğinde problemi çözmek için farklı teknikler kullanılabilir. Bazı tekniklerin kullanılması, veri tipi için uygun olmayabilir. Bu durumda veri hazırlama aşamasına geri dönülerek gereksinimlere uygun hale getirilir (Larose, 2005: 7).

Uygulama kullanıma sunulmadan önce belirlenen hedeflere ulaşılıp ulaşılmadığı, bir veya daha fazla model denemek suretiyle *değerlendirilerek karar verilir* (Larose, 2005: 7). Modelin eksiksiz olarak değerlendirilmesi ve iş amaçlarına uyup uymadığına emin olmak için, model oluşturulana kadar yürütülen adımların tekrar gözden geçirilmesi büyük önem taşımaktadır. Buradaki temel amaç, analiz süresince gözden kaçan önemli bir noktanın var olup olmadığını belirlemektir. Değerlendirme aşaması sonucunda veri madenciliğinde uygulanacak adımların yürütülmesi ve belgelenmesi sağlanır.

Geliştirilen modellerin kullanılması *uygulama aşamasında* gerçekleştirilir. Modelin oluşturulması, veri madenciliği uygulamasının tamamlandığı anlamına gelmez. Örgütün farklı bölümlerindeki süreçler üzerinde modelin çalıştırılması, modelin doğruluğunu görmek açısından gereklidir (Larose, 2005: 7).

4.4. Kamu ve Özel Sektörde Veri Madenciliğinin Uygulama Alanları

Veri madenciliği kamu ve özel sektörde başta bankacılık, sigorta ve sağlık olmak üzere farklı alanlarda kullanılmaktadır. Bazı kullanım alanlarına örnek vermek gerekirse, bankacılık sektöründe risk analizi ve usulsüzlük tespiti, müşteri kazanma ve kredi kartı dolandırıcılıklarının tespiti maksadıyla kullanılmaktadır. Sigortacılık sektöründe müşteri kaybı sebeplerinin belirlenmesi ve riskli müşteri profilinin belirlenmesi maksadıyla uygulanmaktadır. Sağlık sektöründe ise tıbbi teşhis, uygun tedavi süreçleri, ürün geliştirme ve sonuçların tahmini için tercih edilmektedir (Rygielski ve vd., 2002: 488).

Ayrıca vergi kaçaklarının önlenmesi amacıyla yapılan vergi beyanlarının doğruluğunun otomatik olarak incelenmesi ve yanlış/hatalı beyanların belirlenmesi, internet üzerinden toplanan veriler ve kamudaki diğer verilerin entegrasyonu sonucunda bilgi güvenliği risklerinin daha iyi bir şekilde analiz edilmesi ve güvenlik kameraları, kişisel konum bilgisi sağlayan cep telefonu, sensor gibi araçlarla toplanan verilere dayalı güvenlik uygulamalarının geliştirilmesi gibi çok farklı konularda büyük veri toplanabilir ve veri madenciliği uygulamaları kullanılarak öngörü ve tahmin analizleri yapılabilir (Kalkınma Bakanlığı, 2015: 132).

Veri madenciliği ile kamu örgütleri, vatandaşın ihtiyaçlarını ve davranışlarını anlama, olayları önceden tahmin etme, bütçelerini ve mevcut insan kaynaklarını en optimum şekilde kullanma olanağına kavuşmaktadır. Bu sayede, veriden elde edilen bilgi kolay anlaşılır grafikler ile görüntülenebilir ve bilgi kullanılarak daha hızlı ve daha etkili kararlar alınabilir. Kamu örgütleri veri madenciliği uygulamaları vasıtasıyla, hilekârlık tespiti ve engelleme, uygunsuz ödemelerin tespiti, gelir ve vergi toplama, kamu güvenliği, asayiş ve ülke güvenliği konularında politikalar geliştirebilir. Ayrıca veri madenciliği uygulamaları ile oluşturulan modellere dayalı olarak usulsüz ödemeler ve kaynak

israfları tespit edilerek kaynak mali kaynak tasarrufu sağlanabilir.

Kamu hizmetleri ve politikaları için veri madenciliği uygulamalarının kullanım alanları açısından alanyazın incelendiğinde çeşitli kamusal sorunların çözümüne yönelik örnekler üzerinde çalışıldığı görülmektedir. Bunlar arasında e-devlet ve kolluk kuruluşlarının verimliliğini ve etkinliğini artırmak, karar vericilere zamanında destekleyici bilgi sağlamak için otomatik metin raporlarını analiz ederek metin analizi ve sınıflandırma tekniklerinin nasıl kullanılabileceği (Ku ve Leroy, 2014: 534), Çin’de kaçan suçluların yakalanma alanlarında sınıflandırma modeli ile siber izleme ve özellikle kaçan suçluların sorgularının karşılaştırılmasıyla derinlemesine araştırmalar yapmak üzere öneriler sunulması (He vd., 2011: 123), ihale başvuruçuları arasından kartel oluşumunu tespit etmek ve önlemek (Ralha vd., 2012: 11642), yayınlanan hassas verilerle ilgili gizliliğin korunması çalışması (Prakash ve Singaravel, 2015: 1), e-devletin yapılanması için arazi kayıtlarının sistematik ve güvenli bir şekilde güncellenerek devletin birimlerinin daha iyi kamu hizmeti vermesi ve bunları vatandaşlara yönelik hale getirmesi (Hanumanthappa vd., 2012: 208), İngiltere’nin kişisel vergilendirme sistemi (Brown vd., 2006: 666) ve polislerin düz metin olarak hikaye formatında hazırladıkları raporlardan önemli bulguları çıkarmak için metin madenciliği uygulamalarının kullanılması (Keyvanpour vd., 2011: 872) öne çıkmaktadır.

5. TÜRKİYE’DE KAMUDA BÜYÜK VERİ VE VERİ MADENCİLİĞİNE İLİŞKİN POLİTİKA VE PROJELER

Büyük veri kullanımının özel sektöre sağladığı rekabet avantajı karşısında, kamu örgütleri sağlık, eğitim, güvenlik, enerji, savunma, haberleşme ve ulaşım gibi hizmetlerin sunumunda etkinlik ve verimliliğin artırılması, daha etkili ve kaliteli kamu politikası oluşturulması amaçları doğrultusunda kullanmaya

yönelmektedir. Türkiye'de de özellikle merkezi hükümet düzeyinde bakanlıklar ve diğer bazı kamu kurumlarının büyük verinin yanında veri madenciliği uygulamalarından yararlandığı görülmektedir. Nitekim büyük veri ve veri madenciliğine ilişkin ulusal düzeydeki politika belgeleri ve eylem planlarında çeşitli hedeflerin dile getirilmesi, hükümet programları ve kalkınma planlarında amaçların belirlenmesi ve bakanlık stratejik planlarında faaliyet ve projelere yer verilmesi Türkiye'de kamu kurumlarının hizmet ve politikalarına belirli bir yansımının olduğunu göstermektedir.

5.1. Türkiye'de Yürütülen Başlıca Büyük Veri ve Veri Madenciliği Projeleri

Türkiye'de veri madenciliği alanında çalışmalara başlanmış olmakla birlikte, hayata geçen büyük veri uygulamaları sayısal olarak sınırlı kalmıştır. Özel sektörde haberleşme, perakende, bankacılık gibi müşteri bilgisinin ve davranışlarının yakından takip edildiği alanlarda hizmet gösteren şirketler veri analizi çalışmaları yürütmektedir. Bu sektörlerde faaliyet gösteren şirketler müşterileriyle ilgili topladıkları büyük hacimli veriyi kullanarak

müşteri davranışlarını tahmin etme, müşterilerine sunulan hizmetlerin kalitesini artırma, müşteri memnuniyetini ve bağlılığını artırma, karar alma mekanizmalarına yardımcı olacak öngörüler üretme, ürün tasarımlarını müşteri beklentilerine göre şekillendirme, ürün fiyatlandırmasını optimize etme, detaylı müşteri segmentasyonu yapma ve ürünlerini sundukları kanalların verimliliğini artırma gibi alanlarda "veri madenciliği" çalışmaları yapmaktadır (Kalkınma Bakanlığı, 2015: 132). Kamu kurumlarında ise vergi, sosyal güvenlik, sağlık ve emniyet gibi bazı alanlarda büyük miktardaki veri üzerinde yapılan analizler yoluyla verimliliği artırmak, kayıp-kaçak oranlarını düşürmek ve hizmet kalitesini iyileştirmek için çalışmalar gerçekleştirilmektedir (Kalkınma Bakanlığı, 2015: 56).

Türkiye'de kamu kurumları tarafından yürütülen büyük veri ve veri madenciliği projelerinin bir envanteri bulunmamaktadır. Bununla birlikte, Türkiye Bilişim Derneği (TBD) tarafından yakın dönemde yayımlanan bir rapor, üç kamu kurumu tarafından yürütülen büyük veri projelerini derlemiştir (bkz. Tablo 1).

Tablo 1: Türkiye'de Bazı Kamu Kurumları Tarafından Yürütülen Büyük Veri Projeleri

Kurum	Proje ve Faaliyet Adı	Proje ve Faaliyet Açıklaması
Sosyal Güvenlik Kurumu	e-Bildirge Sistemi	Tüm sigorta prim tahsillerinin ve işyeri tescil kayıtlarının takibinin gerçekleştirildiği sosyal güvenlik projesi
	MEDULA	Vatandaşlara ait tüm sağlık ödemelerinin belirlenen kurallar çerçevesinde yürütülmesine yönelik bir proje
	Aylık Tahsis ve Diğer Uygulamalar	Emeklilere yönelik Aylık Tahsis Uygulaması başta olmak üzere 1500 adet olgunlaşmış eski (legacy) uygulamadan oluşan bir uygulama grubu
	ALO 170 Projesi	Çalışma ve Sosyal Güvenlik Bakanlığı, Sosyal Güvenlik Kurumu ve Türkiye İş Kurumu Genel Müdürlüğü tarafından sunulan tüm hizmetlerle ilgili olarak bilgilendirme yapan ve çözüm üreten bir iletişim merkezi
	Veri Ambarı Projesi	Sağlık, Sigortalama, Tahsis (Emeklilik) rapor grupları ile ortalama 500 standart rapordan on binlerce farklı tertipte (kombinasyonda) sorgu yapılarak rapor üretilebilen bir veri ambarı
Milli Eğitim	MEBBİS	Öğretmenler, okul yöneticileri gibi Milli Eğitim Bakanlığı'na bağlı personelden oluşan kullanıcı gruplarının giriş yapabildiği modüllerden oluşan bir yazılım

Kurum	Proje ve Faaliyet Adı	Proje ve Faaliyet Açıklaması
Bakanlığı	e-Okul	Bir öğrencinin okula kaydından başlayıp, mezuniyetine kadar olan tüm süreçle ilişkin veriyi içeren bir sistem
	e-YAYGIN	Yaygın eğitim kurumları tarafından sürdürülen seminer, faaliyet ve kurslara ilişkin tüm iş ve işlemlerin internet ortamında yürütüldüğü otomasyon sistemi
	ALO 147 Projesi	Bakanlığın görev ve sorumluluklarıyla ilgili her türlü talep, şikâyet, görüş ve öneri, ihbar, bilgi edinme soruları ile bakanlığın vermiş olduğu tüm hizmetler hakkında bilgiyi etkin ve hızlı bir biçimde sunabilmek ve sorunların çözüme kavuşturulabilmesi amacıyla kurulan iletişim merkezi
Sağlık Bakanlığı	e-Nabız Projesi	Muayene, tetkik ve tedavi bilgileri başta olmak üzere tüm sağlık verilerinin yönetebildiği, tıbbi özgeçmişe tek bir yerden ulaşılabilen bir kişisel sağlık kaydı sistemi
	UBYS Projesi (Sağlık.NET)	Her bireyin kendi bilgilerine erişebildiği, bireyin doğumundan önce başlayıp tüm yaşamı boyunca sağlığıyla ilgili verilerden oluşan işlevsel bir veri tabanının, yüksek bant genişlikli ve tüm ülkeyi kapsayan bir iletişim omurgasında paylaşılması ve tele-tıp uygulamalarına varan teknolojilerin mesleki pratikte kullanılmasını temel alan elektronik kayıt sistemi
	MHRS Projesi	Vatandaşların Merkezi Hekim Randevu Sistemi'ni arayarak canlı operatörlerden veya web üzerinden kendilerine istedikleri hastane ve hekimden randevu alabilecekleri bir uygulama
	Aşı Takip Sistemi	Tüm ülke genelindeki aşuların stok kontrolünden, taşınmasının yapıldığı ve barındırıldığı her noktadaki ısı maruziyetine, uygulanan kişi ve yerine kadar her türlü veri ve bilginin tutulduğu sistem

Kaynak: (Türkiye Bilişim Derneği, 2016: 19-23) kaynağından tablolaştırılmıştır.

5.2. Üst Politika Belgelerinde Büyük Veri ve Veri Madenciliğinin Yeri

Genel olarak e-devlete, bilgi ve iletişim teknolojilerinin yaygınlaştırılması, enformasyon altyapısı ve bu alandaki yatırımların planlanması gibi günümüz büyük veri ve veri madenciliği uygulamalarına temel oluşturan girişimler 1999 yılında yayınlanan Ulusal Enformasyon Altyapısı Ana Planı (TUENA), e-Türkiye Girişimi Eylem Planı (2000) ve e-Dönüşüm Türkiye Projesi Kısa Dönem Eylem Planı (2003-2004) gibi politika belgelerinde düzenlenmiş ve takip edilmiştir.

Daha yakın döneme geldiğimizde ise örneğin 65. Hükümet Programı, büyük veri ve veri madenciliğine odaklanan politika ve amaçları içermektedir. Dijital Türkiye Projesi, Entegre Kent Yönetimi Bilgi Sistemi Projesi, Kamu Bulut Projesi ve

Kamu Entegre Veri Merkezi Projesinin gerçekleştirilmesi, kamu kurumlarının topladıkları ve ürettikleri verilerin, bilgi güvenliği tedbirlerini almak suretiyle paylaşımının sağlanması ve kamuda büyük veri uygulamalarının hayata geçirilmesi bu kapsamdaki hedeflerden bazılarıdır (65. Hükümet Programı, 2016).

İçinde bulunduğumuz dönemi kapsayan 10. Kalkınma Planı (2014-2018) incelendiğinde ağırlıklı olarak e-devlet kapsamında politika ve hedefler belirlenmekle birlikte, büyük veri ve veri madenciliğine yönelik önerilere de rastlanmaktadır. Katma değerli hizmet üretimi için kamu sektörü bilgisinin paylaşımı ve yeniden kullanımının sağlanması; açık kaynak kodlu yazılımlar, büyük veri, bulut bilişim, yeşil bilişim, mobil platform, nesnelerin interneti gibi ürün, hizmet ve yönelimlerin değerlendirilerek kamu için uygun olabilecek çözümlerin hayata geçirilmesi

(s.55); bulut bilişim hizmetlerinin gelişebilmesi ve yaygınlaşması için gerekli yasal ve idari düzenlemelerin yapılması ve akıllı uygulamaların sağlık, ulaştırma, bina, enerji ile afet ve su yönetimi gibi alanlar başta olmak üzere kullanımı (s.97) bunlar arasında öne çıkmaktadır (Kalkınma Bakanlığı, 2013).

65. Hükümet Programı ve 10. Kalkınma Planına uygun olarak 2016 yılında hazırlanan Ulusal E-Devlet Stratejisi ve Eylem Planı (2016-2019) doğrudan büyük veri ve veri madenciliğine ilişkin politika ve hedefler içermesi bakımından önemli bir politika dokümanıdır. Ulusal E-Devlet Stratejisi e-devlet ekosisteminin etkinliği ve sürdürülebilirliğinin sağlanması, altyapı ve idari hizmetlere yönelik tak sistemlerin hayata geçirilmesi, kamu hizmetlerinde e-dönüşümün sağlanması ile kullanım, katılım ve şeffaflığın artırılmasından oluşan dört stratejik amaç üzerine inşa edilmiştir. Geliştirilen dört stratejik amaç ile de büyük veri ve veri madenciliği arasında ilişki kurmak mümkündür. Kamu Entegre Veri Merkezlerinin kurulması, kamu bulut bilişim altyapısının oluşturulması, akıllı kentler programı geliştirilmesi, kent yönetimi bilgi sistemi geliştirilmesi, kamu verisinin paylaşılması, açık veri paylaşım portalının oluşturulması, kamu verilerinin açık veriye dönüştürülmesi ve paylaşılması, kamuda büyük veri ve nesnelerin interneti politikalarının geliştirilmesi gibi eylem ve uygulamalar bu çerçevede belirlenmiştir. Ayrıca yurtiçi ve yurtdışı örneklerinin incelenerek hangi hizmetlerin büyük veri ve nesnelerin interneti kullanılarak sağlanabileceğinin tespit edilmesi, buna yönelik politika ve stratejilerin geliştirilerek projeler üretilmesi, ihtiyaç duyulan konularda mevzuat çalışmalarının yapılarak uygulanması, kamu çalışanlarının eğitimi ve farkındalık artırıcı çalışmaların gerçekleştirilmesi öngörülmektedir (Ulaştırma, Denizcilik ve Haberleşme Bakanlığı, 2016).

2015-2018 yıllarını kapsayan ve ikinci Bilgi Toplumu Stratejisi ve Eylem Planında ü büyük verinin ekonomik değere dönüşmesini sağlamak, bu amaçla sağlık,

sosyal güvenlik, vergi, güvenlik ve istatistik gibi alanlar başta olmak üzere kamuda büyük veri uygulamaları geliştirmek, kamuda büyük veri pilot uygulaması gerçekleştirmek gibi eylem ve politikalar belirlenmiştir (Kalkınma Bakanlığı, 2015). Ulusal politika belgelerinde büyük veriye ilişkin uygulamalar öne çıkmakla birlikte, veri madenciliği teknik ve modellerinin kullanımını gerektiren hedef ve projelerin öngörüldüğü anlaşılmaktadır.

5.3. Bakanlıkların Büyük Veri ve Veri Madenciliği Politikalarını Belirlemeye Yönelik İçerik Analizi

Stratejik planlar, üst politika belgelerinde geliştirilen strateji, amaç ve politikaların somut hedeflere, faaliyet ve projelere dönüştürüldüğü, gerçekleşme süreçlerinin takip edildiği ve ulaşılan başarının değerlendirildiği belgelerdir. Bu yönüyle Türkiye'deki bakanlıkların stratejik planları, büyük veri ve veri madenciliğine ilişkin politika, hedef ve projelerin bir fotoğrafını çekmek amacıyla içerik analizi yöntemiyle incelenmiştir. İçerik analizinin yürütülmesinde yöntem olarak büyük veri, açık veri, veri ambarı, veri madenciliği gibi konuyla ilgili anahtar kelimeler üzerinden metin içerik taraması gerçekleştirilmiştir. Ayrıca büyük veri ve veri madenciliğiyle ilişkili olabilecek, ancak anahtar kelime taramasından ulaşılamayacak örtük hedef ve projelerin tespiti için ayrıntılı bir okuma yapılmıştır. Bu çerçevede, büyük veri ve veri madenciliği ile ilişkisi bulunan strateji ve faaliyetler ile bunların ilişkili olduğu amaç ve hedefler Tablo 2'de gösterilmektedir.

Ancak burada üç önemli kısıttan söz etmek gerekir. Yukarıda içeriği incelenen 10. Kalkınma Planı 2014, Bilgi Toplumu Stratejisi ve Eylem Planı 2015 ve Ulusal e-Devlet Stratejisi ve Eylem Planı 2016 yıllarında yayımlanmıştır. Ancak Tablo 2'de yer verilen diğer bakanlıkların birçoğu stratejik planlarını bahsedilen temel üst politika belgelerinden önce hazırlanmıştır. Dolayısıyla yeni geliştirilen eylem planlarındaki hedefleri dikkate almaları mümkün değildir.

İkinci kısım ise, bakanlıklar dışında merkezi idare kurumlarının stratejik planlarının bu çalışmanın kapsamı dışında bırakılmış olmasıdır. Merkezi idarenin bağlı, ilgili ve ilişkili kuruluşlarının büyük veri ve veri madenciliğine yönelik politika ve hedefleri bu araştırma kapsamına alınmamıştır. Örneğin Çalışma ve Sosyal Güvenlik Bakanlığı'nın ilgili kuruluşu olan Sosyal Güvenlik Kurumu'nun büyük veri projeleri yürüttüğü Tablo 1'den anlaşılmaktadır, ancak 2017-2021 dönemini kapsayan stratejik planı incelendiğinde bakanlığın

doğrudan büyük veri ve veri madenciliğine ilişkin hedef ve strateji belirlemediği görülmektedir.

Üçüncü olarak, büyük verinin toplanması ve analizine ilişkin bir takım politika, hedef ve faaliyetlerin stratejik planda yer almaması ihtimali de göz ardı edilmemelidir. Plan döneminden önce başlanan ve sürdürülen büyük veri ve veri madenciliği projeleri veya veri güvenliği kapsamında açıklanmaması gereken faaliyetlere stratejik planda da yer verilmediği tahmin edilmektedir.

Tablo 2: Bakanlıkların Büyük Veri ve Veri Madenciliğine İlişkin Amaç ve Hedefleri

Bakanlık Stratejik Planı	Stratejik Amaç	Stratejik Hedef	Hedefe Yönelik Strateji ve Faaliyetler
Bilim, Sanayi ve Teknoloji Bakanlığı (2013-2017)	Ekonominin bütün kesimlerinde verimlilik ve temiz üretim/eko-verimlilik ile ilgili gösterge ve istatistikler oluşturmak, ölçümler yapmak ...	Politika belirleme ve araştırma çalışmalarında kullanılmak üzere temiz üretim/eko-verimlilik ile ilgili makro düzeyde gösterge ve istatistikler oluşturulacak ...	Temiz üretim/eko-verimlilik ile ilgili istatistiklerin hesaplanmasına yönelik veri altyapısı çalışmalarının yapılması ve güncellenmesi
Ekonomi Bakanlığı (2013-2017)	İkili, bölgesel ve çok taraflı ticaret ve yatırım ilişkilerini ülkemizin çıkarları doğrultusunda yürütmek	İkili, bölgesel ve çok taraflı stratejilerimizin oluşturulmasında nicel ve nitel analiz yöntemlerinin etkin kullanımı sağlanacaktır.	Ekonomi ve dış ticaret konusunda modelleme çalışmaları ... ekonomi politika uygulamaları ile ticaret politikası araçlarının etkinliğinin tespiti ve geliştirilmesi için nicel ve nitel analiz yöntemleri ile etki analizi mekanizmaları tesis edilecektir.
Enerji ve Tabii Kaynaklar Bakanlığı (2015-2019)	Enerji dışı hammadde tedarik güvenliği	Ülkemiz maden kaynaklarının aranmasının uluslararası standartlarda yapılması ve raporlandırılmasına ilişkin altyapı çalışmaları tamamlanacaktır.	[Maden kaynaklarıyla ilgili] Kapsayıcı bir veri envanterinin diğer paydaşların katılımı ile oluşturulması, bu kapsamda tüm jeolojik bulguların ... yer aldığı bir veri bankasının oluşturulması amacıyla yasal ve fiziki altyapının geliştirilmesi (fiziki ve elektronik veri tabanı şeklinde) ...

Bakanlık Stratejik Planı	Stratejik Amaç	Stratejik Hedef	Hedefe Yönelik Strateji ve Faaliyetler
Gıda, Tarım ve Hayvancılık Bakanlığı (2013-2017)	Üretimden tüketime kadar, uluslararası standartlara uygun gıda güvenilirliğini sağlamak.	Güvenilir gıda konusunda toplumda duyarlılığı artırmak	Alo Gıda Hattına gelen başvurular doğrultusunda elde edilen bilgiler veri havuzunda toplanarak risk haritası çıkarılmakta ve hangi ürünlerin hangi bölgelerde riskli olduğu saptanarak denetim ve izleme programlarının daha etkin olarak yapılması sağlanmaktadır.
Sağlık Bakanlığı (2013-2017)	Birey ve topluma erişilebilir, uygun, etkili ve etkin sağlık hizmetleri sunmak	Sağlık hizmet sunumunun izlenmesi, değerlendirilmesi ve kanıta dayalı karar almak için sağlık bilgi sistemlerini geliştirmek	Sağlık hizmetlerini planlamak amacıyla oluşturulan “Karar Destek Sistemi” için veri ambarları kurmak ve veri madenciliği uygulamalarını iyileştirmek
Ulaştırma Denizcilik ve Haberleşme Bakanlığı (2017-2021)	Kurumsal kapasiteyi geliştirmek	Kurumsal ihtiyaçları etkin, verimli ve güvenli bir şekilde karşılamak amacıyla bilişim alt yapısını güçlendirmek	Büyük verinin toplanması, organize edilmesi ve analizi ile birlikte, stratejik kararlarının alınabilmesi, ihlal olayları olmadan önce tahmin edilebilmesi amacıyla, yapısal verilerin yanı sıra yapısal olmayan verilerin de kurumların karar destek sürecinde kullanılması
İçişleri Bakanlığı (2015-2019)	Hizmet sunumunda bilgi teknolojilerini etkin, verimli ve güvenli olarak kullanmak	Nüfus ve vatandaşlık hizmetlerinin sunumunda bilgi teknolojilerinin kullanımını yaygınlaştırmak	Açık kaynak kodlu yazılımlar, büyük veri, bulut bilişim, yeşil bilişim, mobil platform, nesnelerin interneti gibi ürün, hizmet ve yönelimler değerlendirilerek kamu için uygun olabilecek çözümler hayata geçirilecektir.
Çevre ve Şehircilik Bakanlığı (2015-2017)	Sağlıklı, güvenli ve mekân kalitesi yüksek yerleşmeleri oluşturmak için; mekânsal planlama, kentsel dönüşüm proje ve uygulamaları ile altyapı hizmetlerini yürütmek, ulusal coğrafi bilgi sistemini kurmak	Yatırım ortamını iyileştirmek, kamu hizmetlerini etkinleştirmek ve verimliliği arttırmak için Coğrafi Bilgi Sistemleri çalışmaları yürütülecektir.	Kent Bilgi Sisteminin yaygın kullanımını sağlamak ve coğrafi bilginin üretimi ve paylaşımında çağdaş bilgi teknolojilerini kullanmak amacıyla yerel yönetimlere hizmet veren bulut bilişim altyapısı kurulması ve idame ettirilmesi hedeflenmektedir.

Bakanlık Stratejik Planı	Stratejik Amaç	Stratejik Hedef	Hedefe Yönelik Strateji ve Faaliyetler
Gümrük ve Ticaret Bakanlığı (2013-2017)	Gümrük ve ticarete ilişkin yönlendirici politikaları ilgili kurum ve kuruluşlarla işbirliği ve eşgüdüm halinde belirlemek	Politika geliştirme ile hızlı ve nitelikli karar almaya yardımcı olacak veri, enformasyon ve bilgi yönetimini geliştirmek ve kurumsallaştırmak	Veri, enformasyon ve bilgi ile desteklenen politikalar geliştirmek
Kalkınma Bakanlığı (2014-2018)	Küresel ve ulusal gelişmeler ile temel eğilimleri yakından izleyerek, ... kısa, orta ve uzun vadeli tahmin, analiz ve araştırmalar yapmak.	Mikro veri temelli politika ve etki analizi çalışmaları yapılacaktır	Bu kapsamda, kamu politikalarının etkilerini ölçmek amaçlı geliştirilmiş ve uluslararası bilimsel standartlara uygun niceliksel metodolojilerle mikro veriye dayalı analizler yapılacaktır.
Millî Eğitim Bakanlığı (2015-2019)	Bütün bireylere çağın gerektirdiği bilgi, beceri, tutum ve davranışın kazandırılması ile girişimci, yenilikçi ... bireylerin yetişmesine imkân sağlamak	Bütün bireylerin bedensel, ruhsal ve zihinsel gelişmelerine yönelik faaliyetlere katılım oranını ve öğrencilerin akademik başarı düzeylerini artırmak	Öğrencilerin açık öğretim okullarında ortalama kayıtlı kalma süresi gibi önemli verilerin toplanmasına yönelik izleme sistemi kurularak veriye dayalı olarak açık öğretim sisteminin niteliği geliştirilecek...

Yukarıda yer verilen kısıtlar altında büyük veriden yola çıkarak politika ve proje geliştirmeye yönelik Orman ve Su İşleri Bakanlığı, Çalışma ve Sosyal Güvenlik Bakanlığı, Kültür ve Turizm Bakanlığı, Gençlik ve Spor Bakanlığı ve Maliye Bakanlığı dışındaki tüm bakanlıkların doğrudan ya da dolaylı strateji ve/veya faaliyet belirlediği anlaşılmaktadır. Büyük veri ve veri madenciliğine ilişkin stratejik amaç ve hedef değil, strateji ve faaliyet geliştirilme eğilimi tespit edilmektedir. Bununla birlikte, geliştirilen strateji ve faaliyetlerin büyük kısmının dolaylı olarak büyük veri ve veri madenciliğiyle ilişkisi kurulabilmekte, sadece Sağlık Bakanlığı, Ulaştırma, Denizcilik ve Haberleşme Bakanlığı ile İçişleri Bakanlığı doğrudan büyük veri ve veri madenciliğine atıf yapmaktadır.

6. SONUÇ

Büyük veri kullanımı ve veri madenciliği uygulamalarının özel sektöre sağladığı rekabet avantajını; kamu kurumları hizmet sunumu, karar verme ve politika yapımını iyileştirme yönünde değerlendirmekte ve bu maksatla büyük veri kullanımına yönelik projelere yönelmektedir. Özel sektörde olduğu gibi kamu kurumları için de avantajlar sunan büyük uygulamalar, birtakım zorluklar da içermektedir. Veri madenciliği ve büyük veri projeleri genellikle ciddi bir maliyeti gerektirmektedir. Ayrıca verilerin toplanması, depolanması ve analizi ise verinin niteliği, bilişim altyapısının yetersizliği, donanımlı insan kaynağının bulunmaması gibi nedenlerle her zaman başarılı olmayabilmektedir. Veri güvenliği, kurumlar arası veri entegrasyonu ve veri paylaşımındaki sorunlar da büyük veri konusunda ulusal ve kurumsal düzeyde politika geliştirmeyi zorlaştırmaktadır.

Bu çalışmada, üst politika belgeleri ve bakanlık stratejik planları büyük veri ve veri madenciliğine açısından içerik analizi yöntemiyle incelenmiş ve bazı sonuçlara ulaşılmıştır. Buna göre bakanlıkların büyük kısmının temel görevlerini yerine getirmede büyük veriyi toplama, analiz etme ve bunu hizmete dönüştürme konusunda hedeflere sahip olduğu görülmektedir. Bu durum, Türkiye'de bakanlıkların büyük veri ve veri madenciliğinin ortaya çıkardığı fırsatların farkında olduklarını ve bu kapsamda hedef ve politikalara sahip olduklarını göstermektedir. Bununla birlikte, sadece üç bakanlıkta doğrudan büyük veri ve veri madenciliği ifadeleri faaliyet ve stratejiler içinde kullanılmaktadır. Elbette bu kapsamda geliştirilen üst politika belgeleri nispeten yeni olduğundan, bakanlıkların bir sonraki stratejik planlama döneminde bu konuya çok daha fazla önem ve öncelik verecekleri tahmin edilebilir.

Planlar incelendiğinde bakanlıkların kendi görev alanlarına ilişkin veri üretme, farklı kurumlarla veri paylaşma, çeşitli konularda veri tabanı geliştirme ve veri tabanlarını entegre etme gibi konularda sorun

yaşadıkları sıklıkla ifade edilmektedir. Bu durum büyük veri ve veri madenciliği uygulamalarının bakanlık hizmet alanlarında yaygınlaşmasının önündeki engellerden arasında öne çıkmaktadır.

Büyük veri ve veri madenciliğine yönelik kamu yönetimi alanında yapılan çalışmaların sayısının artırılmasına ihtiyaç bulunmaktadır. Bu çalışmanın, alana bir katkı yapması ve gelecek araştırmalara yol açması beklenmektedir. Bu bağlamda, merkezi idaredeki diğer kamu kurumları ile yerel yönetim kuruluşlarında yürütülen büyük veri ve veri madenciliği uygulamalarının, bu kapsamda geliştirilen proje ve faaliyetlerin yürütülmesinde karşılaşılan zorlukların ve iyi örneklerin saha çalışmaları ve örnek olay araştırmalarıyla incelenmesine ihtiyaç bulunmaktadır. Ayrıca ulusal düzeyde, sektörel bazda ve yerel yönetimler için bütüncül bir büyük veri ve veri madenciliği politika ve stratejisinin geliştirilmesine, bu çerçevede ihtiyaç duyulan kaynakların belirlenmesine ve nitelikli insan gücünün yetiştirilmesine yönelmek gerekmektedir.

KAYNAKÇA

1. 65. HÜKÜMET PROGRAMI (2016). file:///Users/macbookair/Downloads/6_5._H%C3%BCK%C3%BCmet_Program%C4%B1_kapakl%C4%B1.pdf, 14.10.2017.
2. ALTUNIŞIK, R. (2015). "Büyük Veri: Fırsatlar Kaynağı mı Yoksa Yeni Sorunlar Yumağı mı?", Yıldız Social Science Review, 1 (1): 45-76.
3. ATAN, S. (2016). "Veri, Büyük Veri ve İşletmecilik", Balıkesir University The Journal of Social Sciences Institute, 19 (35): 137-153.
4. BAESENS B., MUES, C., MARTENS, D. ve VANTHIENEN, J. (2009), "50 Years of Data Mining and OR: Upcoming Trends and Challenges", Journal of The Operational Research Society, 60: 16-23.
5. BARUTÇUGİL, İ. (2002). Bilgi Yönetimi, Anı Yayıncılık, İstanbul.
6. BHATT, G. D. (2001). Knowledge Management in Organizations, Journal of Knowledge Management, 5(1): 68-75.
7. BROWN, J., COOPER, C. ve PIDD, M. (2006). "A Taxing Problem: The Complementary Use of Hard and Soft OR in The Public Sector, European Journal of Operational Research, 172: 666-679.
8. CHADWICK, A. ve MAY, C. (2003), "Interaction Between States and Citizens in the Age of the Internet: "E-Government" in the United States, Britain, and the European Union, Governance, 16 (2): 271-300.

9. CHAPMAN P., CLINTON, J., KERBER, R., KHABAZA, T., REINARTZ, T., SHEARER, C. ve WIRTH, R. (2000). CRISP-DM 1.0 Step-by-step Data Mining Guide, SPSS Inc.
10. ÇAKIREL, Y. (2016). “İşletmelerde Büyük Veri”, Kırklareli Üniversitesi İktisadi ve İdari Bilimler Fakültesi Dergisi, 5 (1): 52-62.
11. DENHARDT R.B ve DENHARDT, J.V. (2000). “The New Public Service: Serving Rather Than Steering”, Public Administration Review, 60 (6): 549-559.
12. ERBAY, H. ve KÖR, H. (2016). “Büyük Veri ve Büyük Verinin Analizi”, Uluslararası Bilim ve Teknoloji Konferansı, 3-6 Ekim, Ankara, Türkiye.
13. EYÜPOĞLU, C., AYDIN, M.A., SERTBAŞ, A., ZAİM, A.H. ve ÖNEŞ, O. (2017). “Büyük Veride Kişi Mahremiyetinin Korunması”, Bilişim Teknolojileri Dergisi, 10 (2): 177-184.
14. FOSTER, I., GHANI, R., JARMIN, R.S., KREUTER, F. ve LANE, J. (2017). Big Data and Social Science A Practical Guide to Methods and Tools, CRC Press Taylor & Francis Group, Boca Raton.
15. GILCHRIST, A. (2016). Industry 4.0: The Industrial Internet of Things, APress.
16. HAN, J., KAMBER M., ve PEI, J. (2012). Data Mining Concepts and Techniques, Elsevier, London.
17. HANUMANTHAPPA, M., PRAKASH, B.R. ve KUMAR, M. (2012). “Applications of Data Mining in e-Governance: A Case Study of Bhoomi Project”, ICDEM 2010, LNCS 6411, 208-218.
18. HARRISON, T. M., GUERRERO, S., BURKE, B. G., COOK, M., CRESSWELL, A., HELBIG, N., HRDINOVA, J., PARDO, T. (2012). “Open Government and E-Government: Democratic Challenges from A Public Value Perspective”, Information Polity, 17: 83-97.
19. HE, Y., LIU, J. ve LIN, L. (2011). “A Research on Application of Data Mining in Cybre-Pursuit”, International Journal of Management Science and Engineering Management, 6 (2): 123-129.
20. HO, A. T. (2002). “Reinventing Local Governments and the E-government Initiative”, Public Administration Review. 62 (4): 434-444.
21. HOOD, C. (1991). “A Public Management for All Seasons?”, Public Administration, 69 (1): 3-19.
22. HUGHES, O. E. (2003). Public Management and Administration: An Introduction, Palgrave Macmillan, New York.
23. IŞIKLI, R. (2014). “Büyük Veri, Epistemoloji ve Etik Tartışmalar?”, AJIT-e: Online Academic Journal of Information Technology, 5 (17): 89-122.
24. KALKINMA BAKANLIĞI (2013). 10. Kalkınma Planı 2014-2018, <http://www.kalkinma.gov.tr/Lists/Kalkinma%20Planlar/Attachments/12/Onuncu%20Kalk%C4%B1nma%20Plan%C4%B1.pdf>, 15.10. 2017.
25. KALKINMA BAKANLIĞI (2015). Bilgi Toplumu Stratejisi ve Eylem Planı 2015-2018, <http://www.bilgitoplumustratejisi.org/tr/doc/8a9481984680deca014bea4232490005>, 15.10. 2017.
26. KEYVANPOUR, M. R., JAVŞDEH, M. ve EBRAHİMİ, M. R. (2011). “Detecting and Investigating Crime by Means of Data Mining: A General Crime Matching Framework”, Procedia Computer Science 3: 872-880.
27. KU, C. H. ve LEROY, G. (2014). “A Decision Support System: Automated Crime Report Analysis and Classification for E-government”, Government Information Quarterly, 31 (4): 534-544.

28. LAROSE, D. T. (2005). *Discovering Knowledge in Data: An Introduction to Data Mining*, John and Wiley Sons Incorporated, USA.
29. MILLER, D. (2016). "What is social media?", *How the World Changed Social Media*, (Ed.) COSTA, E., HAYNES, N., MCDONALD, T., NICOLESCU, R., SİNANAN, J., SPYER, J., VENKATRAMAN, S. ve WANG, X. UCL Press, London.
30. OĞUZLAR, A. (2003). "Veri Ön İşleme", *Erciyes Üniversitesi İktisadi ve İdari Bilimler Fakültesi Dergisi*, 21: 67-76.
31. ÖZBİLGİN, İ. G. (2015). "Kamuda Büyük Veri ve Uygulamaları", AB2015 Akademik Bilişim Konferansı, 31 Ocak-06 Şubat, Eskişehir, Türkiye.
32. ÖZEKES, S. (2003). "Veri Madenciliği Modelleri ve Uygulama Alanları", *İstanbul Ticaret Üniversitesi Fen Bilimleri Dergisi*, 2 (3): 65-82.
33. PRAKASH, M. ve SINGARAVEL, G. (2015). "An Approach for Prevention of Privacy Breach and Information Leakage in Sensitive Data Mining", *Computers and Electrical Engineering*, 45: 134-140.
34. RALHA C. G. ve SILVA, C. V. S. (2012). "A Multi-Agent Data Mining System for Cartel Detection In Brazilian Government Procurement", *Expert Systems with Applications*, 39: 11642-11656.
35. RYGIELSKI, C., WANG, J. C. ve YEN, D. C. (2002). "Data Mining Techniques for Customer Relationship Management", *Technology in Society*, 24: 483-502.
36. SİLAHTAROĞLU, G. (2013). *Veri Madenciliği (Kavram ve Algoritmaları)*, Papatya Yayıncılık, İstanbul.
37. TAN, A. H. ve YU, P. S. (2003). "Guest Editorial: Text and Web Mining", *Applied Intelligence*. 18: 239-241.
38. THOMPSON, J. (2008). "Don't Be Afraid to Explore Web 2.0", *Phi Delta Kappa International*, 89 (10), s.711-718.
39. TÜRKİYE BİLİŞİM DERNEĞİ (2016). *Büyük Veri Uygulamaları Çalışma Grubu 4*, Ankara: TBD. www.kamu-bib.org.tr/CG4-Buyuk-Veri-Uygulamalari-2016.pdf, 13.10.2017.
40. ULAŞTIRMA, DENİZCİLİK VE HABERLEŞME BAKANLIĞI (2016). *2016-2019 Ulusal e-Devlet Stratejisi ve Eylem Planı*, <http://www.edevlet.gov.tr/wp-content/uploads/2016/07/2016-2019-Ulusal-e-Devlet-Stratejisi-ve-Eylem-Planı.pdf>, 13.10.2017.
41. USAMA, F. ve PAUL, S. (1997). "Data Mining and KDD: Promise and Challenges", *Future Generation Computer Systems*, 13: 99-1