

40. BÖLÜM / CHAPTER 40

Bilgi Bilimin Kavramsal Çerçevesinde Veri Yönetimi ve Veri Analitiği Uygulamaları

Data Management and Data Analytics Applications in the Conceptual Framework of Information Science

Özgür KÜLCÜ¹

¹Prof. Dr., Hacettepe Üniversitesi, Edebiyat Fakültesi Bilgi ve Belge Yönetimi Bölümü, Türkiye
E-mail: ozgurkulcu@hacettepe.edu.tr

DOI: 10.26650/B/SS53.2024.015.40

ÖZ

Bilgiyi kaydetme, saklama ve tekrar kullanma becerisi bizi tüm canlılardan farklılaştırdı. Süreçte bilgi kaynağından özgürleştikçe daha akışkan ve etkileşimli hale geldi. Bilgi hizmetleri de kaynaktan ilişkilere ve hatta sentezlere yöneldi. Geleneksel bilgi hizmetleri yaşamda ağırlığını koruyor. Öte yandan bilgi, iletişim, yönetme ve doğrulama aracına dönüşüyor. Bilginin yapı taşı olan veri ve veri kaynaklarına dönük araştırmalar hızla artıyor.

Çalışmada bu çerçevede önceleri bilgi taşıyıcıları ardından bilginin kendisinin yönetimine dönük teorik ve uygulamalı çalışmalar anlatılmakta, bilgi ve belge yönetimi alanının sanat ve insanlık bilimleri, sosyal bilimler, davranış bilimleri ve doğa bilimleri ile ilişkileri açıklanmaktadır. Bilgi bilimin değişen kapsamı içerisinde, verinin ortaya çıkışından bilgiye dönüşümüne içeriğin kapsamı, formatı, kullanımı ve onu çevreleyen karmaşık tüm konuların bir bütün olarak değerlendirmesi gerekmektedir. Bu kapsamda çalışmada betimsel teknikler kullanılarak bilgi bilim çerçevesinde veri yönetimi, veri analitiği, iş zekâsı ve makine öğrenme çalışmalarının neler olduğu anlatılmakta, bilgi hizmetleri kapsamında kullanılacak veri analitiği bileşenleri tanıtılmaktadır.

Bilgi bilim çerçevesinde veri yönetimi, veri analitiği, iş zekâsı ve makine öğrenme çalışmaları bilgi hizmetleri kapsamında kullanılacak veri analitiği bileşenleri arasındadır. Veri analitiği, bilgi sistemleri ve yazılım uygulamaları konusunda teknik alt yapıya sahip olmayan sıradan insanlar için bile öğrenilebilir ve uygulanabilir bir araştırma alanı olmaya başlamıştır. Makine öğrenmede yeni yaklaşımlar, bilgisayara dayalı görüş sistemleri ve doğal dil işleme alanındaki gelişmeler bilgi bilime yeni ufuklar açmaktadır. Çalışmada bilgi bilimde veri yönetimi temel bileşenleri tanımlanmaya, bilgi hizmetlerinde kullanılacak veri analitiği teknikleri açıklamaya çalışılmaktadır.

Anahtar kelimeler: Bilgi bilim, bilgi ve belge yönetimi, veri yönetimi, veri analitiği, iş zekâsı

ABSTRACT

The ability to record, store and reuse information differentiated us from all living things. In the process, it became more fluid and interactive as it became free from its source of information. Information services also moved from source to relationships and even syntheses. Traditional information services maintain their weight in life. On the other hand, information becomes a means of communication, management and verification. Research on data and data sources, which are the building blocks of knowledge, is increasing rapidly.

In this context, theoretical and applied studies on the management of information itself, and then the information carriers, are explained in this framework, and the relations of the field of information and records management with the arts and humanities, social sciences, behavioral sciences and natural sciences are explained. In the changing scope of information science, it is necessary to evaluate the scope, format, use of content and all the complex issues surrounding it as a whole, from the emergence of data to its transformation into knowledge. In this context, using descriptive techniques, data management, data analytics, business intelligence and machine learning studies are explained within the framework of information science, and data analytics components that can be used within the scope of information services are introduced.

Within the framework of information science, data management, data analytics, business intelligence and machine learning studies are among the data analytics components that can be used within the scope of information services. Data analytics has become a learnable and applicable research field even for ordinary people who do not have a technical background in information systems and software applications. New approaches in machine learning, computer-based vision systems and developments in natural language processing open new horizons for information science. In this study, it is tried to define the basic components of data management in information science and to explain data analytics techniques that can be used in information services.

Keywords: Information science, information and records management, data management, data analytics, business intelligence

Extended Abstract

Preliminary studies in the field of information management were mostly made on information resources, then theoretical and applied studies were carried out on the management of information itself. In this context, it was closely related to the arts and humanities, then social sciences, behavioral sciences, and finally natural sciences. In the changing scope of information science, it is necessary to evaluate the scope, format, use of content and all the complex issues surrounding it as a whole, from the emergence of data to its transformation into knowledge.

Traditionally, management of cultural heritage, librarianship, archiving, museology have been included in the arts and humanities. Information systems and records management, bio-informatics, service development in information centers were among the subjects that converged with social and behavioral sciences over time. Recently, information retrieval, information architecture, data management, bibliometrics, information/document engineering have been among the research topics of information science, natural sciences and mathematics.

Information science is essentially concerned with the social aspects of knowledge. It focuses on the development of services related to the transformation of individual knowledge into social knowledge, access and use principles. It creates the conditions for the reuse of information. The building block of knowledge, effective management of data, use of technological tools in information services, and even contributing to the development of related technologies are within the scope of the study of information science. In that context carrier resources (books, electronic media, etc.) are an intermediary. The main purpose is to ensure that the data or information content, whose integrity and validity has not been compromised, can be accessed to authorized persons on time and when necessary, and to protect it without damaging its integrity, regardless of the medium or carrier.

Many fields or subjects such as organizations, services and processes in organizations, subject specialties, geographic information services, historical research, information technology management and policies can be included in the scope of information research. Increasingly, data management and data analytics applications are increasing their weight in information research.

Especially with the 2010s, big data studies have started to transform from theoretical researches to technical applications with the introduction of cloud computing and artificial intelligence applications. Different types of libraries, archives and information centers naturally rank among the most important data sources. In this context, institutional information and document centers come to the fore. Big data in libraries even though it doesn't appear directly in data sources? This question is commonly asked today. Research shows that the demand for big data is increasing, especially in digitalized libraries. However, it is seen in the literature that big data studies in libraries are limited.

Big data in information centers; It can be used effectively in resource purchasing decisions, personalized reader services, effective process analysis and development of information services. In this context, data mining and data management applications, data visualization and analysis auxiliary tools can be used. Information centers, as the place where human information is collected, are among the most important potential sources of big data research.

The production and use of electronic documents and the social and industrial trends related to the subject seem to be among the topics of discussion of different fields. The question of what could be the future of computational archival science as a new field by examining the theory and applications of information and records management is one of the research topics of scientists working on the subject today.

On the other hand, the public opinion formed towards the use of public data for social benefit; It is important not only in terms of accountability and transparency, but also in terms of effective use of resources, strengthening local governments, and ensuring technical and social development. For the effective use of national digital public content, artificial intelligence technologies, developing information and database management applications, and data analysis/data mining studies should be examined. In the future, more investments are required to ensure the integration of management and administrative systems with business processes and information technology. Enterprise information management systems have to cover not only the management of information and document resources, but also data management applications. All processes, from the creation of any content produced or provided on corporate systems to its archiving and re-evaluation within the scope of data analytics applications, are evaluated within this scope.

In this context, in this study, data management, data analytics, business intelligence and machine learning studies are explained within the framework of information science, and data analytics components that can be used within the scope of information services are explained. Data analytics has become a learnable and applicable research field even for ordinary people who do not have a technical background in information systems and software applications. New approaches in machine learning, computer-based vision systems and developments in natural language processing open new horizons for information science.

Information science development, which is associated with the beginning of recorded history, took place depending on the transformations in humanity's relationship with knowledge. The ability to record, store and reuse information differentiated us from all living things. In the process, it became more fluid and interactive as it became free from its source of information. Information services also moved from source to relationships and even syntheses. Traditional information services maintain their weight in life. On the other hand, information becomes a means of communication, management and verification. Research on data and data sources, which are the building blocks of knowledge, is increasing rapidly. Machine learning and artificial intelligence are based on collecting, cleaning, processing, synthesizing, using and protecting the generated data. As mentioned in the study, information is growing rapidly by taking all these developments into the general framework of science. In these conditions, where we are and how much we invest in what is very important. What we will do both academically and in practice will carry us to the future with new dimensions or we will limit ourselves within the framework of traditional information services.

1. Giriş

Bilginin temel varlık nedeni etkileşimdir. Süje ve obje arasındaki etkileşimin ürünü olan bilgi, etiketlendiğinde ve kaydedildiğinde anlamlı hale gelir. Bilginin paylaşıldıkça ve yeni bilgilerle harmanlandıkça yaşamı dönüştüren güce dönüştüğünü fark eden insanoğlu uygarlığı inşa etmeye başarabilmiştir. Bireysel bilginin toplumsal bilgiye dönüşümünde aracı rolü bilgi bilim üstlenmiştir. Tarihsel süreçte bilgi merkezleri bilgiye erişim, araştırma ve öğrenme gereksinimlerinin nasıl karşılanacağı sorusuna yanıt aramışlardır. Bilginin kaydedildiği taşıyıcı araçlar olarak kil tabletler, kâğıt, manyetik bantlar, yarı iletken ortamlarda bilginin form yapısı, tanımlaması, depolanması, erişimi ve korunması bilgi bilimin uğraşı alanları arasında yer almıştır. Günümüzde kaynaktan ve mekândan özgürleştikçe bilginin kendisinin ve yapı taşı olan verinin nasıl yönetileceği pek çok disiplini içerisine alan yeni bir fenomene dönüşmeye başlamıştır. Bu kapsamda teorik ve uygulamalı çalışmaların yapıldığı bilgi ve belge yönetimi alanı, sanat ve insanlık bilimleri, sosyal bilimler, davranış bilimleri ve doğa bilimleri ile yakın ilişki içerisinde. Bilgi bilimin değişen kapsamı içerisinde, verinin ortaya çıkışından bilgiye dönüşümüne içeriğin kapsamı, formatı, kullanımı ve onu çevreleyen karmaşık tüm konular bir bütün olarak değerlendirilmek durumundadır. Son dönemlerde bilgi bilimin doğa bilimleri ve matematik alanına giren araştırma başlıkları arasında bilgi erişim, bilgi mimarisi, veri yönetimi, bibliyometri, bilgi/belge mühendisliği ön plana çıkmaktadır. İnsan bilgisinin toplandığı yer olarak bilgi merkezleri aynı zamanda büyük veri araştırmalarının en önemli potansiyel kaynakları arasında yer almaktadır. Çalışmada bu kapsamda bilgi bilimin temel dinamikleri ve geleneksel rolleri dışlanmadan değişen yönleri ele alınmaktadır. Bilginin organizasyonu, doğru ve güvenilir bilgiye erişim kadar veri yönetimi ve veri analitiğini anlamak da önemlidir. Alanımızın dönüşümünün bir parçası olarak bilgi taşıyıcısı araçlar ve mekanların yönetiminden bilginin kendisinin yönetimine dönük süreç devam etmektedir. Bilgi hala kavramsal olarak oldukça güçlü. Öte yandan bilginin yapı taşı olarak veri, güvenilir ve işlenebilir kaynak olarak önümüzde yeni ufuklar açmaya aday. Bu koşullarda gerek akademik çalışmalarda gerekse uygulamada çezeğimiz yeni ufuklar bizi yeni boyutlarla geleceğe taşıyabilir ya da geleneksel bilgi ve belge hizmetleri çerçevesinde kendimizi sınırlandırabiliriz. Çalışma bu kapsamda bir çerçeve çizmeyi amaçlamaktadır.

2. Bilgi Bilimin Gelişimi ve Günümüzde Aldığı Boyutlar

Bilgi bilim, yüzyılı aşan teorik ve uygulamalı birikimini değişen koşullarda yönetim, iletişim, ekonomi, istatistik, sistem mühendisliği ve veri yönetimi alanlarıyla harmanlayarak geleceğe uzanmaktadır. Önceleri bilgi taşıyıcıları ardından bilginin kendisinin yönetimine

döntük teorik ve uygulamalı çalışmaların yapıldığı bilgi ve belge yönetimi alanı, sanat ve insanlık bilimleri, sosyal bilimler, davranış bilimleri ve doğa bilimleri ile yakın ilişki içerisinde olmuştur (Bates, 2015). Aşağıdaki şekilde yer aldığı gibi bu kapsamda sanat ve insanlık bilimleri, sosyal bilimler, davranış bilimleri ve doğa bilimleri ile matematik alanları kesişebilmektedir. Geleneksel çerçevede kütüphanecilik, arşivcilik ve müzecilik hizmetleri kaynağa ve kaynağa erişime odaklanmaktadır. Bu kapsamdaki çalışmalar sanat ve insanlık bilimlerine daha yakın olarak tanımlanmaktadır. Kaynak merkezli bilgi yönetiminden ilişki merkezli bilgi yönetimine geçişle (Tonta, 2012, s.301), geleneksel kütüphane hizmetlerinin duvarların ötesinde bilgi hizmetlerine dönüşmesini sağlamıştır (Çakın, 2012). Bu gelişmeler bilgi ve belge yönetimi alanını önce sosyal bilimler ve davranış bilimlerine yakınlaştırmış, ardından bilgi sistemleri ve veri yönetimi ile doğa bilimleri ve matematik alanlarına yakın araştırmalar yürütülmeye başlanmıştır.

Çalışmada bilgi bilimin aldığı yeni boyutları inceleyebilmek için betimsel araştırma yöntemi kullanılmıştır yararlanılmıştır. Bu kapsamda literatür incelemesi ve eleştirel yansıtma ve sorgulama yoluyla yaşamın kalitesini artırmak için önceden planlanmış, düzenlenmiş ve işbirliğine dayalı sistematik incelemeler şeklinde tanımlanan eylem araştırması tekniklerinden yararlanılmıştır (Uzuner, 2005; Mills, 2003; Johnson, 2002)

Tablo 1. Bilgi Bilim Alanında Araştırmaların Bilim Alanlarına Dağılımı (Higgins, 2018, s.1322)

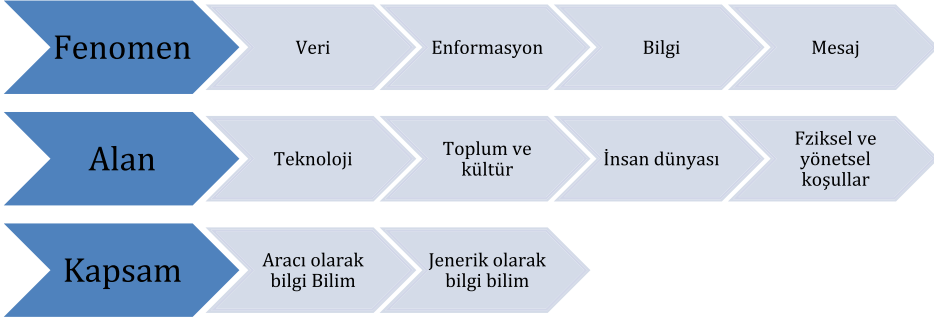
Okuma alışkanlığı Sanat kütüphaneciliği Yayıncılık çalışmaları Dijital değer yönetimi Sosyal epistemoloji Görsel işitsel arşivler Analitik bibliyografya Kültür enformatiği Müze enformatiği	Bilgi politikaları Kütüphane Koleksiyonlar El yazmaları Teorik kütüphanecilik Semantik Diplomatik Kütüphane yönetimi Sosyal enformatik Legal enformatik İş süreç yönetimi Jeneolojik arşivler	Bilgi davranışları Belge yönetimi Özel kütüphaneler Sosyal bilimler Veri arşivleri İş enformatiği	Kütüphane otomasyonu Yönetim bilgi sistemleri Kamusal arşiv yönetimi İnsan bilgisayar etkileşimi Bio-enformatik Tıbbi sağlık enformatiği	Bilgi erişim Veri tabanı yönetim sistemleri Bilgi mimarisi Elektronik yayıncılık Dijital kütüphaneler Veri madenciliği Bilim müzeleri Veri depolama Geo-enformatik Kimya enformatiği
Sanat ve İnsanlık Bilimler	Sosyal ve Davranışsal Bilimler	Doğa Bilimleri ve Matematik		

Yukarıdaki şekilde yer aldığı gibi sanat ve insanlık bilimleri kapsamında kültürel mirasın yönetimi, kütüphanecilik, müzecilik diplomatik ve jeneolojik araştırmalar, tarih ve arşiv araştırmaları, bilgi sosyolojisi, sosyal enformatik konuları bilgi bilimin alanına girmiştir. Sosyal ve davranış bilimleri ile yakınsanan konular arasında bilgi sistemleri ve belge yönetimi, bio-enformatik, bilgi merkezlerinde hizmet geliştirme vb. yer almıştır. Bilgi bilimin doğa bilimleri ve matematik alanına giren araştırma başlıkları arasında ise bilgi erişim, bilgi mimarisi, veri yönetimi, bibliyometri, bilgi/belge mühendisliği yer almıştır (Higgins, 2018, s.1322).

Geçmişten günümüze bilgi bilimin temel konuları arasında aşağıdaki başlıklar sıralanabilir:

- Kütüphane ve bilgi bilim
- Arşiv bilimi
- Bibliyografya
- Belge yönetimi
- Bilgi (Knowledge) yönetimi
- Enformatik
- Bilgi sistemleri
- Sosyal çalışmalar ve enformasyon (Bates, 2015; Higgins, 2018, s.1322).

Bilgi bilim özünde bilginin toplumsal yönleriyle ilgilenir. Bireysel bilginin toplumsal bilgiye dönüşümü, erişim ve kullanım ilkeleriyle ilgili hizmetlerin geliştirilmesine yönelir (Shera, 1968, Külcü, 2018). Bilginin tekrar kullanımına dönük koşulları oluşturur. Bilginin yapı taşı verilerin etkin yönetimi, bilgi hizmetlerinde teknolojik araçların kullanımı, hatta ilgili teknolojilerin gelişimine katkı sağlama bilgi bilimin çalışma kapsamında yer alır. Aşağıdaki şekilde yer almayan taşıyıcı kaynaklar (kitap, elektronik ortam vb.) bir araçtır. Asıl amaç ortamı ya da taşıyıcı ne olursa olsun bütünlüğü ve geçerliliği bozulmamış veri ya da bilgi içeriğini, zamanında ve gerektiğinde yetkilendirilmiş kişilere eriştirebilmek ve bütünlüğü bozulmadan korunmasını sağlamaktır.



Şekil 1. Bilgi bilimin kapsamı (Higgins, 2018, ss.1323)

Yukarıdaki şekilde kaynak, uygulama boyutları ve teorik kapsamın açıklanmaktadır. Bu kapsamda bilgi içeriğinin üretim, sağlama, düzenleme, yayım/dağıtım, indeksleme, depolama, erişim, dönüştürme ve kullanım aşamalarından oluşan yaşam döngüsü çerçevesinde değerlendirilmesi gerekir. Konu uzmanlıkları, bilginin organizasyonu, kullanıcı araştırmaları, bibliyometri ve tarihi çalışmalar, belge yönetimi, epistemolojik araştırmalar, bilişsel çalışmalarda geliştirilen yöntem ve tekniklerden yararlanılmaktadır.



Şekil 2. Bilgi bilimin etkileşimleri (Higgins, 2018, ss.1323)

Bilgi araştırmaları kapsamına organizasyonlar, organizasyonlarda hizmetler ve süreçler, konu uzmanlıkları, coğrafi bilgi hizmetleri, tarihi araştırmalar, bilgi teknolojileri yönetimi ve politikalar gibi pek çok alan ya da konu girebilmektedir. Giderek veri yönetimi ve veri analitiği

uygulamaları bilgi arařtırmaları ierisinde ađırlıđını artırmaktadır. Bu kapsamda ařađıdaki bařlıkları da gz nne olarak bilgi bilimin alıřma alanının yenido tanımlanmasında yarar vardır.



Şekil 3. Bilgi arařtırmalarının kapsamı (Bates 2015)

Yukarıdaki řekil bilgi bilimin kaynaktan sistemlere ya da srelere ynelen yeni kapsamını ortaya koymaktadır. Bu kapsamda giderek kapsamı geniřleyen organizasyonlarda bilgi ynetimi sistemlerini ncelikle incelemekte yarar grlmektedir.

3. Organizasyonlarda Bilgi Ynetimi Sistemleri

Bilgi toplumuna dnřm hızı ve bilgi teknolojilerinin geliřimi gz nne alındıđında, bilgi ynetiminin dnya ekonomisinin geleceđinde bařat rol oynayacađını sylemek mmkndr. Organizasyonlarda bilgi ynetimi dijital teknoloji uygulamaları, bilgisayar programları ve programlama dilleri, masast uygulamalar, veri tabanı ynetimi, grselleřtirme yazılımları ve grnt iřleme uygulamaları ile birlikte kresel bir ađ ekonomisinin omurgasını oluřturmaktadır. Bu kapsamda ařađıdaki bařlıkları kurumsal bilgi ynetimi sistemlerinin bileřenleri olarak tanımlanmaktadır (Eroshkin, Kameneva, Kovkov, Sukhoruko, 2017).

3.1. Bilgi yönetiminde kullanılan modern kavram, metodoloji ve yaklaşımlar

- BPM (Business Process Management- İş Süreci Yönetimi): Kurumsal iş süreçlerinin yönetimi için geliştirilen çözümleri tanımlamaktadır.
- BPM (Business Process Modeling- İş Süreci Modelleme): İş süreçlerini modellenmesine dayanmaktadır. “İş Süreçleri Yönetimi” uygulamalarının bir alt sistemi olarak kullanılmaktadır.
- TQM (Total Quality Management; Toplam Kalite Yönetimi): Toplam kalite yönetimi, tüm organizasyonel süreçlerin sürekli kalite iyileştirme yaklaşımıyla ele alınmasını gerektirmektedir.
- BPR (Business Process Reengineering- İş Süreçlerinin Yeniden Yapılandırılması): İş süreçlerinde yeniden yapılanmada devrim niteliğinde değişikliklerin yer aldığı bir yaklaşım olarak tanımlanmaktadır.
- CPI (Continuous Process Improvement- Sürekli Süreç İyileştirme): İş süreçlerinin sürekli iyileştirilmesine yönelik bir yaklaşımdır.
- BSC (Balanced Scorecard): Strateji haritaları veya dengeli puan kartlarına dayanmaktadır.

3.2. Kurumlarda modern kavramların, metodolojilerin ve yaklaşımların uygulandığı sistemler ve araçlar

- BI (Business Intelligence- İş Zekâsı): İş zekasına dayalı bilgi sistemlerinin temel DSS’ye (Decision Support System; Karar Destek Sistemi)- destek sistemleri karar verme) dayanır. İş zekası uygulamaları bilimsel literatürde yer aldığı gibi KDD (Knowledge Discovery in Databases- Veritabanlarında Bilgi Keşfi), OLAP (Online Analytical Processing; Çevrimiçi Analitik İşleme) ve EIS (Executive Information Systems – Uzman Bilgi Sistemleri) gibi - çok boyutlu analitik veri işleme ve veri madenciliği tekniklerini uygulamayı gerektirir (Eroshkin, Kameneva, Kovkov, Sukhoruko, 2017, s. 610).
- BPMS (Business Process Management System- İş Süreci Yönetim Sistemi): İş süreçlerinin yönetimi, BPM, TKY kavramlarının uygulanmasını sağlamak ve iş süreçlerini modellemek için farklı notasyon ve dilleri kullanmaya dayanır.
- CPM (Corporate Performance Management; Kurumsal Performans Yönetimi): Ku-

rumun performans yönetim sistemidir. EPM (Enterprise Performance Management; Kurumsal Performans Yönetimi), yani işletmenin performans yönetimi sistemine dayanır. Bilimsel literatürde, bu kavram BPM (Business Performance Management; İş Performansı Yönetimi) olarak da adlandırılmaktadır.

- PPM (Project Portfolio Management- Proje Portföyü Yönetimi): Proje tabanlı portföy yönetimine dayanır.
- KM (Knowledge Management- Bilgi Yönetimi): İş stratejilerinin başarılı bir şekilde uygulanabilmesi için gerekli olan bilgi yönetimi sistemlerini ele alır.

3.3. Operasyonel yönetim bilgi sistemleri

- ECM (Enterprise Content Management; Kurumsal İçerik Yönetimi): Organizasyonun bilgi kaynaklarının yönetimi, “elektronik belge yönetim sistemi (ERMS), PM (Proje Yönetimi) sistemlerinin bütünleştirildiği sistemlerdir.
- ERP (Enterprise Resource Planning; Kurumsal Kaynak Planlaması) - kurumsal kaynak planlaması;
- CRM (Customer Relationships Management- Müşteri İlişkileri Yönetimi) - müşteri ilişkileri yönetimi);
- SRM (Supplier Relationships Management- Tedarikçi İlişkileri Yönetimi) - Tedarikçilerle ilişki yönetimi;
- SCM (Supply Chain Management- Tedarik Zinciri Yönetimi) - tedarik zinciri yönetimi;
- MES (Manufacturing Execution System) sistemi operasyonel (üretim yeri) üretim süreçlerinin kontrolü (üretim-bakım);
- WMS (Warehouse Management System; Depo Yönetim Sistemi - depo yönetim sistemi);
- CMMS (Computerized Maintenance Management System - Bilgisayarlı Bakım Yönetim Sistemi) - bilgisayarlı bakım yönetimi (onarım);
- EAM (Enterprise Asset Management- Kurumsal Varlık Yönetimi) - CMMS sistemlerinin dönüştürülmesi, EAM stratejisinin uygulanması (ekipmanın güvenilirliğini veya ekipmanın işletme parametrelerini, masrafları artırmadan bakım, onarım ve lojistik destek maliyetlerini azaltmak)

- SCADA (Supervisory Control and Data Acquisition- Denetleyici Kontrol ve Veri Toplama) - Denetleyici kontrol ve veri toplama, teknolojik işlemlerin otomatik kontrol sistemlerine benzer, teknolojik işlemlerin otomatik kontrolü için programlar.

3.4. BPM (Business Process Management; İş Süreci Yönetimi) kavramını uygulayan gösterim ve modelleme dilleri

- BPMN (Business Process Model and Notation- İş Süreci Modeli ve Notasyonu) iş süreçlerini modellemek için grafiksel bir gösterimdir.
- EPC (Event-driven Process Chain- Olay odaklı Süreç Zinciri) olay odaklı süreç zinciridir. Grafik gösterimi, iş modellemesi için kullanılan grafik türü.
- BPEL (Business Process Execution Language- İş Süreci Yürütme Dili), iş süreçlerinin resmileştirilmesi ve bunlar arasındaki etkileşim protokolleri için XML tabanlı bir dildir. İş süreçlerinin yürütülmesi için kullanılmaktadır.

3.5. Tüm yaşam döngüleri dikkate alınarak modern ve gelişmiş yönetim konsept ürünü ya da hizmet inşası

- PLM (Product Lifecycle Management- Ürün Yaşam Döngüsü Yönetimi), ürün üretiminin ve yaşam döngüsü yönetimidir.
- BIM (Building Information Model- Yapı Bilgi Modeli) tam yaşam döngüleri dikkate alınarak yapı tasarımıdır;
- BLM (Yapısal Yaşam Döngüsü Yönetimi- Yapısal yaşam döngüsünün yönetilmesi anlamına gelir) (Eroshkin, Kameneva, Kovkov, Sukhoruko, 2017, s. 610-611).

4. Veri Araştırmaları ve Bilgi Merkezleri

Büyük veri kavramı ilk olarak 2001 yılında Laney tarafından geleneksel veri yönetim araçları tarafından işlenmesi mümkün olmayan verileri tanımlamak için kullanılmıştır (Laney, 2001).

Özellikle 2010'lu yıllarla birlikte büyük veri çalışmaları bulut bilişim ve yapay zekâ uygulamalarının devreye girmesiyle hızla teorik araştırmalardan teknik uygulamalara dönüşmeye başlamıştır (Li, Jiao, Zhang, Xu, 2019).

Farklı tür kütüphaneleri, arşivler ve bilgi merkezleri doğal olarak en önemli veri kaynakları arasında sıralanmaktadır. Bu kapsamda kurumsal bilgi ve belge merkezleri ön plana çıkmaktadır. Doğrudan veri kaynakları arasında görünmese de kütüphanelerde büyük ver var mı? Sorusu günümüzde yaygın olarak sorulmaktadır (Li, Jiao, Zhang and Xu, 2019, s.23).

Araştırmalar özellikle dijitalleşen kütüphanelerde büyük veri talebinin arttığını göstermektedir. Bununla birlikte literatürde kütüphanelerde büyük veri çalışmalarının sınırlı kaldığı görülmektedir. Bu tür çalışmaların yeterli sayıda olmamasının nedenleri arasında, veritabanı sağlayıcılarının kütüphanelerin veri depolama ve işleme gereksinimlerini karşılaması gösterilmektedir (Xu, Du, Wang ve Liu, 2017). Ayrıca bilgi sistemlerinde büyük miktarda dijital veri işlenmesine karşın, bu verilerin çok az oranı daha sonraki araştırmalarda kullanılabilir durumdadır. Bu çerçevede verilerin tekrar işlenebilir hale getirilmesi önemli bir iş olarak görülmektedir. Öte yandan mevcut bilgi sistemlerinin çoğu büyük verileri depolama, işleme veya analiz etme yeteneğine sahip değildir. Bu tür sistemler artan miktarda veri ile başa çıkakta dahi güçlük çekebilmektedir. Giderek artan sayıda kullanıcı verisi, web içeriği, görsel işitsel malzemeler ve veri tabanları, yazışmalar, idari ve mali işlemlerle ilgili bilgi sistemlerindeki kayıtların büyük veri araştırmaları için kullanılabilir kılınması yapısal ve teknik bir süreç olduğu kadar hukuki koşullarının da iyi tanımlanması gerektirmektedir (Salo, 2010). Ayrıca, farklı veri kümeleri arasındaki bağlantının nasıl yapılandırılacağı bu çerçevede düşünülmek durumundadır.

Bilgi merkezlerinde büyük veri; kaynak satın alma kararları, kişiselleştirilmiş okuyucu hizmetleri, etkin süreç analizi ve bilgi hizmetlerinin geliştirilmesinde etkin olarak kullanılabilir. Bu kapsamda veri madenciliği ve veri yönetimi uygulamaları, veri görselleştirme ve analizi yardımcı araçlarından yararlanılabilmektedir (Wu, Su ve Deng, 2013). İnsan bilgisinin toplandığı yer olarak bilgi merkezleri büyük veri araştırmalarının en önemli potansiyel kaynakları arasında yer almaktadır.

4.1. Büyük Verilerin Dijital Kütüphaneler Üzerindeki Etkisi

Kütüphanelerde veri yönetimi uygulamaları geliştirilirken büyük verinin depolanması ve işlenmesindeki güçlükler, aşırı bilgi artışının yarattığı sorunlar, bilgi kaynaklarının verimli kullanımı ile zorluklar ve eldeki bilginin geçerliliği ve güvenilirliği ile ilgili belirsizlikler da düşünülmek durumundadır (Carole, 2008). Kütüphanelerde veri yönetimini sadece geleneksel bilgi yönetimi yaklaşımlarıyla gerçekleştirmek mümkün değildir. Teknolojik yetenekler açısından, büyük veri işleme bilgi ve becerilerine sahip veri mühendisleri, büyük verileri modelleyebilen ve analizler gerçekleştirebilen veri analistleri ve bu veriyi etkili biçimde yönetebilen veri yöneticilerine ihtiyaç artmaktadır. Tüm bu işlemler ve süreçler modern bilgi yönetimi uygulamalarının bir parçası olarak da düşünülebilir. Bu kapsamda veri yöneticileri, veri arşivcileri, alan ve konu uzmanları yetiştirilmek durumundadır (Boulton, 2014).

Bilgi profesyonelleri uzunca yıllardır bilimsel bilginin toplamasına ve analiz edilmesine önemli katkılar sağladılar. Büyük veri çağında bilgi profesyonelleri uzmanlaşmış veri profes-

yonellerine dönüşmek durumundalar. Veri profesyonelleri için beceri ile ilgili bilgiler arasında açık veri lisanslama anlaşmaları, fikri mülkiyet hakları, kişisel ve hassas veriler, veri sahipliği, veri maskeleme ve erişim yetkilendirmeleri, veri yönetimi planları, kaynak kullanımı, veri analizi, büyük veri dağıtımı ve ilgili mimariler, kurumsal veri tabanları, dijital arşivlerin yönetimi, veri referansı ve veri yayımı yer almaktadır (Li, Jiao, Zhang and Xu, 2019, s.24).

Diğer yandan veri içeriğinin nasıl kullanılacağı ve değerlendirileceği ayrı bir araştırma konusudur. Veriyi doğru okumak ve değerlendirmek için veri okuryazarlığı ilkelerinin de belirlenmesi gerekmektedir. Veri okuryazarlığı, büyük veri çağında bilgi okuryazarlığının yeni bir boyutu olarak kabul edilebilir (Huang ve Li, 2016).

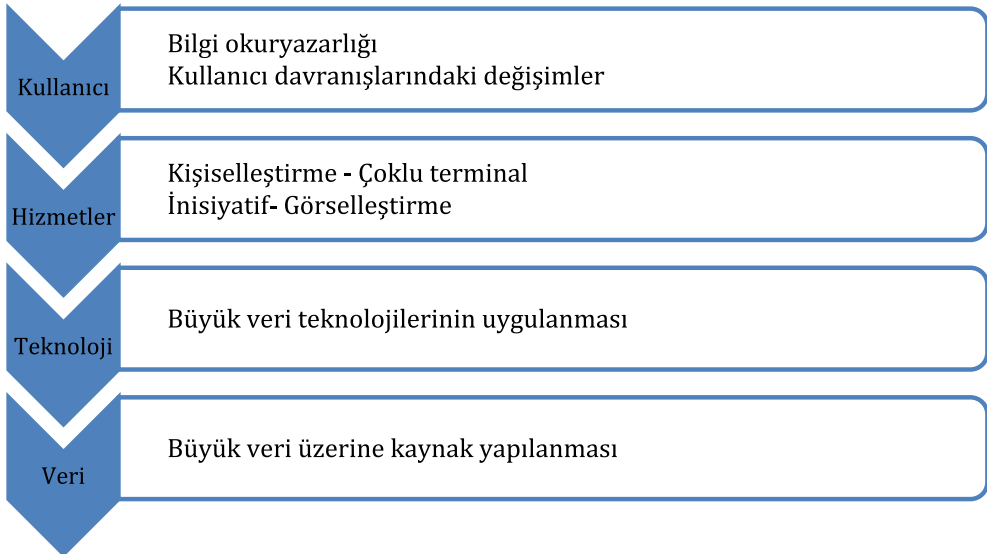
Hizmet kaynağı perspektifinden bakıldığında, geleneksel dijital bilgi merkezleri genellikle sadece veri kaynaklarının sağlanmasında rol üstlenmekte ve daha çok bireysel bilginin toplumsal bilgiye dönüşmesinde öncü bir rol oynamaktadırlar. Ancak, bireysel kullanıcı ilgi alanlarına ve veri içeriği gereksinimlerine yeterince dikkat etmemektedirler. Büyük veri teknolojilerinin hızla gelişmesiyle, bilgi merkezleri geleneksel kitle hizmeti modelinden kişiselleştirilmiş hizmet modeline geçiş yaşamaktadırlar. Bu değişim bilgi merkezi kullanıcılarının bilgiyi daha etkili bir şekilde edinmelerini ve bilgi kaynaklarının optimizasyonunu kolaylaştırmaktadır (Wang, 2015). Bu koşullarda kişiselleştirilmiş kullanıcı bilgisi gereksinimlerini karşılama yeteneklerinin geliştirilmesi gerekmektedir. Kullanıcılar artık bilgi içeriğini bağımsız olarak etiketlemek, bilgiyi yeniden düzenlemek, yaygın olarak paylaşmak ya da dönüştürmek ve çevrimiçi referans araçlarıyla diğer kullanıcılarla etkileşimler geliştirmekle daha fazla ilgilenmektedirler.

Bilgi merkezlerinde kişiselleştirilmiş hizmetlere duyulan ihtiyaç ve çok miktarda verinin sürekli olarak üretilmesi, gerekli bilgiye erişimi zorlaştırmaktadır. Aşırı bilgi yüklemesi sorunu, büyük ölçekli veri kaynaklarından ilgili bilginin tespit edilmesi ve gereksiz bilgi tarama maliyetlerini azaltılması noktasında bilgi profesyonellerinin üreteceği çözümlere gereksinim vardır. Yine sürekli artan veri miktarı ile birlikte farklı veri kaynaklarının ilişkilendirilmesi de gerekli bir konu haline gelmektedir. Örneğin sosyal veriler, kullanıcılar arasındaki ilişkilendirmeler, kullanıcılar ve kaynaklar arasındaki ilişkilendirmeler ve farklı kaynaklar arasındaki ilişkilendirmeler ön plana çıkmaktadır. Bu kapsamda semantik uygulamaların geliştirilmesine ihtiyaç duyulmaktadır.

Bilgi merkezleri kişiselleştirilmiş hizmetler sunarak kullanıcı hizmetlerinin çeşitliliğini de büyük ölçüde artırabilirler. Tamamen işlevsel ve kullanıcı tarafından özelleştirilebilir dijital kütüphane hizmet sisteminin tasarımı, bilişsel ve davranışsal tasarım süreçlerini içerir. Buna

ek olarak, etkili, kullanıcı merkezli bir insan-bilgisayar arayüzü kullanmak, kullanıcı memnuniyetinin artırmak yanında kullanıcılara kişiselleştirilmiş daha bağımsız araştırma ortamı yaratabilir (Ferran, Mor ve Minguillón, 2005). Bilgi merkezleri kamunun sosyal hizmet sisteminin önemli bir bileşenidir. Kişiselleştirilmiş hizmet için artan kullanıcı talebini karşılamak için bilgi merkezlerinin teknolojik gelişmelere ayak uydurmaları ve sistem yapısını optimize etmek için büyük veri uygulamalarını hayata geçirmeleri gerekmektedir.

Bilgi merkezlerinde büyük veri öncelikle bilgi ve belge kaynakları, kullanıcı verileri, satın alma ya da ödeme verileri, tarama geçmişi ve idari ve mali süreçlere ilgili belgelerden oluşur (Wang, 2015, Li, Jiao, Zhang and Xu, 2019, s.25).



Şekil 4. Büyük veri çağında bilgi merkezlerinin dönüşümü ve yönelimleri (Zhang and Xu, 2019, s.25)

4.2. Veri Araştırmaları ve Arşivleri

On yıllardır arşivciler, kâğıt belgeler için geliştirilen arşiv teorileri ve yöntemlerini kullanarak elektronik belgelere değer biçiyor, koruyor ve sağlıyorlar. Bununla birlikte, elektronik belgelerin üretimi ve kullanımı ile konuya ilişkin sosyal ve endüstriyel eğilimler farklı alanların tartışma konular arasında görünmektedir. Bilgi ve belge yönetimi teori ve uygulamaları incelenerek yeni bir alan olarak hesaplamalı arşiv biliminin (computational archival science; CAS) geleceği ne olabilir sorusu günümüzde konu üzerine çalışan bilim insanlarının araştırma başlıklarından birisini oluşturmaktadır (Marciano ve diğerleri, 2018).

Temel veri kaynakları arasında yer alan arşivler iş ve sosyal yaşamla ilgili analitik çalışmaların da beşiği konumundadır. Lemieux'in (2012) Robert Allen, Ian Anderson ve Mitchell

Whitelaw'la birlikte yürüttükleri bir projede arşiv içeriklerinin görselleştirme ve görsel analitik konularına ne denli önemli içeriğe kaynaklık ettiğini göstermiştir. Duranti ve Michetti (2015) geleneksel arşivciliğin yerini disiplinler arası yaklaşımlara bırakması gerektiğini dile getirmektedir. Geleneksel arşivlemeyi ileri boyuta taşımak isteyen arşivciler çoklu provenans tezini ortaya koymuşlardır (Cunningham, 2016)

Dijital belgelerin artışı ile dijital belgelerin oluşturulması, işlenmesi, korunmasının, güvenilirliği, korunması ve erişilebilirliği önemli bir sorun haline gelmiştir. CAS'ın bir bütün olarak bu sorunlara çözüm üretme amacıyla doğduğu dile getirilmektedir. Bu kapsamda Lemieux (2014), Kanada'da finansal elektronik kayıtların bağlamının prototip bir grafik gösterimini oluşturmuştur. Thibodeau (2016), grafik teorisinin arşiv provenansın etkin uygulanmasında kullanılabileceğini göstermiştir. Ayrıca blokzinciri uygulamaları arşiv kaynaklarının korunması ve güvenliği olduğu kadar depolanması ve erişimine dönük de yeni yaklaşımların geliştirilmesine olanak verecek gibi durmaktadır (Marciano ve diğerleri, 2018, s. 13-14).

Elektronik belgeler genellikle benzersiz analiz türlerine ve bağlantılara olanak veren mekansal ve zamansal (spatial and temporal) özelliklere sahiptir. Bu niteliklerin arşiv kayıtlarında ortaya çıkarılması, yeni erişim stratejilerinin geliştirilmesi açısından önemlidir (Heard ve Marciano, 2011). Mekansal nitelikler ile on milyonlarca devlet kaydını ilişkilendirmek mümkündür. Örneğin Winling, Connolly, Nelson ve Marciano'da (2017) ile Travis ve arkadaşları (2016), tarihi arşiv içeriğini coğrafi kodlama ve mekansal haritalama ile entegre etmişlerdir. Dingwall, Marciano, Moore ve McLellan (2007)'de geçici boyut, bir kayıt tutma sistemi içinde büyük miktarlarda coğrafi verilerin yakalanabileceğini ve uzun süreli koruma amacıyla zamansal gelişimin depolanabileceğini göstermişlerdir. 2007 e-Legacy projesi, ile Kaliforniya eyaletinin jeo-uzamsal e-kayıtlarının harita üzerinde gösterimi gerçekleştirilmiştir. 2003'te ICAP projesi ile kayıtların versiyonlarını karşılaştırabilme ve geçmiş sorguları etkileştirilebilmek için mekânsal ve zamansal erişim sistemleri geliştirilmiştir (Marciano ve diğerleri, 2018, s. 16).

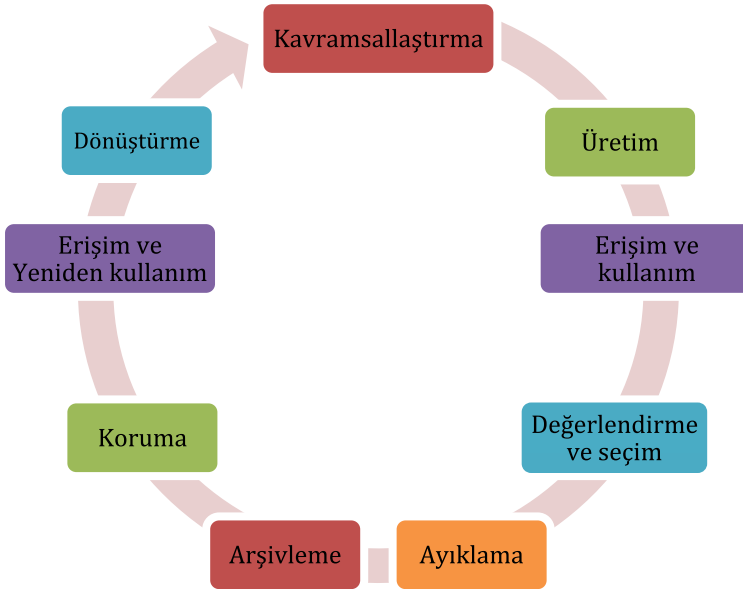
4.2.1. Arşivsel Verilerin Yönetiminde Dijital Kürasyon

Dijital kürasyon genel olarak dijital içeriğin uzun vadede erişilebilir ve kullanılabilir olmasını sağlamak için geliştirilen teknik, idari, mali ve ekolojiiyi yaklaşımları tanımlamaktadır. Kişisel bilgisayarlar ve mobil cihazlar dijital bilgiyi zaman ve mekândan bağımsızlaştırmıştır. Katlanarak artan dijital içeriğin etkin kullanımı ve ardından üretimine neden olan amaçlar dışında farklı nedenlerle kullanılmak üzere veriye dönüşmesi dijital içerik için yeni bir yaşam döngüsüne işaret etmektedir. Dijital küratörlük özellikle 2000'li yıllarla birlikte bilgi bilimi

şemsiyesi içinden yeni bir disiplin olarak ortaya çıkmıştır (Abbott, 2008; Digital Curation Center, 2005). Disiplinin kökeni iki paralel erken bilgi bilim araştırma odağına dayanmaktadır. Belge yönetimi ve arşivcilik dijital uygulamalarla çalışma yönünü elektronik malzemelerin korunmasına odaklamıştır (Hirtle, 2008, s. 125). İlerleyen dönemlerde kütüphaneler de analog malzemelerin dijitalleştirilmesine yönelmişti (Hirtle, 2008, ss.124-125). Bilgi yönetimi uygulamaları elektronik ortama kaydıkça, dijital kütürlük bilgisayar bilimleri ile bilgi bilim arasında köprü görevi görmeye başlamıştır (Higgins, 2017).

Günümüzde bellek kurumları, uzun süreli koruma için artan elektronik doküman içeriğini nasıl yöneteceği konusunda sorunlar yaşamaktadırlar. Geleneksel dijital koruma modelleri genellikle hem veri hem de meta veri içeriğinin bir bilgi paketleri üzerine inşa edildiği Açık Arşiv Bilgi Sistemi (Open Archival Information System) gibi referans modellere dayanır. Belgeler, görüntüler, sesler, videolar vb. belirli dosya formatlarında saklanır. Aktarılan dosyalar standart formatlarda olsalar bile, standartların uygulanması garanti edilemez ve kullanılan yazılıma bağlı olarak sonuçlar farklı olabilir. Kullanılan yazılımlar, içeriği sağlayan kurumlar tarafından kontrol edilmez. Transferlerin uygunluk testleri yapılır, ancak yine de uygunluk testinde sağlayıcısıyla temas kurmazsa sonuçlarda istenilen başarı sağlanamayabilmektedir (Fresa, 2015).

Büyük Britanya'daki Dijital Kütürlük Merkezi (DCC), dijital kütasyon uygulamalarının döngüsünü aşağıdaki adımlarla tanımlamaktadır.



Şekil 5. Dijital kütasyon yaşam döngüsü modeli (Fresa, Justrell, Prandoni, 2015, ss.196)

Kavramlaştırma: Veri yakalama yöntemleri ve depolama seçenekleri dahil olmak üzere dijital nesnelerin oluşturulması için planlamalar yapılmalıdır.

Üretim: Dijital nesnelere üretin ve idari, açıklayıcı, yapısal ve teknik meta veriler eklenmelidir.

Erişim ve kullanım: Yetkilendirilmiş kullanıcıların, parola korumalı olarak kamuya açık veya seçili içeriğe kolay erişebilmeleri sağlanmalıdır.

Değerlendirme ve seçim: Uzun süreli koruma gerektirenleri dijital objeleri değerlendirilmesi ve seçimi.

Ayıklama: Uzun süreli koruma için seçilmemiş dijital objelerin belirlenmesi. Bu içerik için güvenli bir imha planları oluşturulmalıdır.

Arşivleme: Uzun süre koruma kapsamındaki içeriği dijital arşivler ya da veri merkezlerine transfer edilmelidir.

Koruma: Dijital nesnelerin uzun süreli korunmasının yanı sıra geçerlilik niteliklerinin korunmasının sağlanması için gerekli işlemlerin yapılması. Doğrulama prosedürlerinde başarısız olan dijital nesnelerin tasfiyesi gerçekleştirilmelidir.

Erişim ve yeniden kullanım: Verileri, ilgili standartlarda belirtildiği şekilde güvenli bir şekilde saklanması. Verilerin tekrar kullanımları ve erişilebilirliklerini denetlenmesi. Halka açık, kısıtlı erişimli ya da gizli malzemelerin statülerini belirli sürelerde tekrar değerlendirilmesi.

Dönüştürme: Gerekirse orijinalden yeni dijital nesnelerin oluşturulması. Örneğin farklı bir formata geçirmek vb. gibi (Fresa, Justrell, Prandoni, 2015, ss.196-197).

4.2.2. Arşivsel Veri Kaynağı Olarak Kamusal Açık Veri

Kamusal açık veri kapsamına kamu sektöründe üretilen ya da sağlanan içeriğin yönetimi girmektedir. Bu kapsamda InterPARES Trust¹ projesi Avrupa ortaklarından biri olan UCL²'nin araştırma bulguları ilgili çekicidir. UCL'ce 2013'te başlatılan araştırmalarda bilgi ve belge yönetimi disiplininin kamu organlarınınca üretilen ve açık devlet verisi kapsamında verilerin yönetimindeki sorumlulukları araştırılmıştır. Araştırmalarda, İngiltere'de kamu sektörünün farklı kesimlerinden dört örnek kurum incelemesi yapılmıştır. Bu kapsamda Avrupa Genel Veri Koruma Düzenlemesi (European General Data Protection Regulation, GDPR) ve İngil-

1 <http://interparestrust.org/> <https://interparestrust.ai.org/>

2 <https://www.ucl.ac.uk/information-studies/department-information-studies-faculty-arts-humanities>

tere Bilgi Yönetimden Sorumlu Ofisin (UK Information Commissioner's Office (ICO, 2016) çalışmaları referans olarak alınmıştır. Günümüzde vatandaşlar kamu süreçlerinde daha fazla hesap verebilirlik ve şeffaflık talep etmektedirler. Bu değişen iklimde, açık devlet veri politikasının bütünsel olarak değerlendirilmesi son derece önemlidir (Janssen ve diğerleri, 2012). Açık devlet verileri kamusal bilgiye erişim ve sınırlandırmalar, bilgi edinme özgürlüğü, bağlı veri, veri yönetimi uygulamaları ve programlama gibi çok çeşitli yönlerle değerlendirilmek durumundadır. Devlete ait görevler giderek daha fazla ticari ve üçüncü sektör organları tarafından yürütüldüğünden bu verinin telif ve erişim yetkilendirmeleri üzerinde ayrıca durulması gerekmektedir (Shepherd, Bunn, Flinn, Lomas, 2019, s.153).

Açık devlet verilerinin yönetimine dönük aşağıdaki temel konular ön plana çıkmaktadır:

- Açık devlet verileri nelerdir, hangi kategorilerden oluşur.
- Kamusal verinin sahibi niteliğindeki kişi ya da kurumların sorumlulukları nelerdir.
- Veri koruma ve mahremiyete ilişkin yasal ve idari düzenlemeler nelerdir.
- Tıbbi kayıtlar ve hasta verileri de dahil olmak üzere, açık devlet verilerini ve kamu sektörü bilgilerine erişime dönük mevcut durum nedir.
- Açık devlet verilerinin gereksinimler ve beklentiler çerçevesinde sağlanması, düzenlenmesi, depolanması, erişimi, güvenliği ve sürekliliğine dönük hangi stratejiler ve uygulamaların geliştirilmesine gereksinim vardır (Shepherd, Bunn, Flinn, Lomas, 2019, s.155).
- Açık devlet verileri sağlama ve kamu yükümlülüklerine uyma konusunda mevcut örgütsel uygulamaları ve yapılanmalar nelerdir (Shepherd, Bunn, Flinn, Lomas, 2019, s.156).

Belge yönetimi ve dijital kürasyon uygulamalarına benzer şekilde açık verilerin etkin yönetimi için yaşam döngüsü modellerinin geliştirilmesi önemli görülmüştür. Bu kapsamda örnek olarak aşağıdaki şekilde yer alan İngiltere Milli Arşivi tarafından geliştirilen Veri Arşivi Yaşam Döngüsü Modeli verilebilir.



Şekil 6: İngiltere’de veri arşivi yaşam döngüsü modeli (Wyk and Walt, 2018, s.3)

- Veri üretimi: Veri yönetimi planlarının oluşturulması, veri sağlama, veri depolama, veri üretimi.
- Veri işleme: Veri depolama, üstveri oluşturma, veri temizleme, veri doğrulama, Veri anonimleştirme.
- Veri analizi: Veri depolama, üstveri oluşturma, veri temizleme, veri doğrulama, veri anonimleştirme, veri işleme ve analizi, veri görselleştirme.
- Veri koruma: Veri arşivleme, veri koruma, üstveri oluşturma, çıktılarla veriyi ilişkilendirme.
- Veri erişim: Veri paylaşımı, veri yayımlama, çıktılarla veriyi ilişkilendirme.
- Veri Tekrar kullanımı: Veri yönlendirme ve kullanım, veri alıntıları.

Devlet organları tüm toplumdaki bilgi kaynaklarının en büyük üreticisi, sahibi ve kullanıcısı olduğu gibi bu bilginin kontrolünde söz sahibi en temel aktörlerdendir. Bu verilerin değerlendirilmesi, içerik madenciliği ve çeşitli bilgi servislerinin yapılandırılması toplumsal gelişimin itici güçlerinden birisi haline gelmektedir. Ocak 2009’da ABD’de yasallaşan Memorandum for the Heads of Executive Departments and Agencies (İcra Daireleri ve Ajans

Başkanları Memorandumu) devlet verilerinin halkın kullanımına açılmasına dönük ilk ciddi hareketler arasında sayılmaktadır. 20 Eylül 2011’de Brezilya, Endonezya, Meksika, Norveç, Filipinler, Güney Afrika, İngiltere, ABD ve diğer 8 ülke birlikte “G8 Open Data Charter” (G8 Açık Veri Tüzüğü) imzaladı ve Açık Hükümet Ortaklığı (OGP) (the Open Government Partnership, OGP) kuruldu. Ardından 60’tan fazla ülke OGP’nin bir parçası oldu, çoğu ülke devlet verileri için belirli planlar oluşturmak için web siteleri (data.gov) kurdu. Bu değerli devlet verileri kamu istatistikleri ve ekonomik göstergeler, ulusal istatistik kuruluşlarının yayımladığı nüfus sayımı ve diğer demografik sosyal veriler, sağlık kuruluşlarından gelen veriler, yargı kurumlarından gelen suç istatistikleri, kalkınma ajanları ve diğer kamu kurumlarından gelen sanayi, ticaret, eğitim, finans, bilimsel araştırma, ulaşım ve benzeri her türlü değerli kamu verisidir. Bu verilerin önemli bir kısmı halihazırda e-devlet sistemlerinde çeşitli biçimlerde ve şekillerde var olan verilerden oluşmaktadır. Ancak genellikle e-devlet sistemleri kişisel konular dışında bu verilerin halkın erişimine izin vermediğinden, verilerin kullanımı oldukça güçtür. Bu durum kaçınılmaz olarak büyük miktarda dijital kaynağın etkin kullanılamamasına yol açmaktadır.

Öte yandan kamusal verilerin toplumsal fayda sağlamaya dönük kullanımına dönük oluşan kamuoyu; sadece hesap verilebilirlik, şeffaflık açısından değil kaynakların etkin kullanımı, yerel yönetimlerin güçlendirilmesi, teknik ve sosyal gelişimin sağlanması açısından önemlidir. Örneğin Çin’de, 2015 yılında özellikle elektronik ortamda bilgi hizmetlerinin geliştirilmesi ve etkin kullanımının sağlanması adına yürürlüğe konulan yasal düzenlemelere dayanarak, kamu kurumlarında bilgi kaynaklarının tek elden idaresine dönük yeniden yapılanmaya gidilmiştir. Bu kapsamda 2016 yılında devlet verilerinin paylaşılması ve yönetimine dönük ilkeler, 2017 yılında da devlet bilgi sisteminin entegrasyonu ve paylaşımının uygulama modeli yayımlanmıştır (2016, No.5; 2017, No39) (Sixin, Yayuan, Jiang, 2017, s. 525). Bu kapsamda özellikle ulusal arşiv stratejilerinin geliştirilmesi önemli bir yere sahiptir. Gelişmelere paralel olarak kurulan Ulusal Arşiv Yönetim Ajansı özellikle farklı kamu kurumlarında üretilen arşiv içeriğinin merkezileştirilebilmesi için önemli çalışmalar yapmıştır.

Çin’de kamu kurumları öncelikle ulusal dijital arşivle entegrasyonu sağlayacak kendi IT departmanlarını oluşturmuşlardır. Güncel belge işlemlerinin yönetimi ulusal düzenlemelere dayanarak kamu kurumlarının sorumluluğu altındadır. Denetimi ise merkezi sistemce yürütülmektedir (Sixin, Yayuan, Jiang, 2017, s. 527). Çinin büyük veri çağında sürekli olarak kamusal işlevlerini yeniden yapılandırmaya çalıştığından söz edilmektedir. Ulusal Büyük Veri Bürosu ve Ticari Bulut Veri Merkezi (Big Data Bureau and the Commercial Cloud Data Center) Çin’in çeşitli bölgelerinde kendisine bağlı merkezler oluşturmaktadır. Bu kapsamda,

büyük veriye dayanarak bilgi kaynakları yönetimi, merkezi entegrasyon, sosyal ağlar ve pazarlama ana ilerleme konuları arasındadır (Sixin, Yayuan, Jiang, 2017, ss. 528-529).

Kamusal bilgi kaynakları farklı özelliklere sahip birçok konu ve dosyanın depolandığı heterojen bir yapıdadır. Bu nedenle e-devlet sistemlerinden tek bir dijital ulusal arşive bu verilerin transferi oldukça zordur. Bu nedenle, e-devlet bilgi kaynaklarının otomatik olarak e-devlet sisteminden büyük veri merkezine ve/veya dijital arşivlere nasıl dağıtılacağı, çözülmesi gereken kilit sorunlardan biridir. Tarihsel ve sosyal nedenlerle koruma değerine sahip dijital kamu içeriğinin e-devlet sistemlerinden kalıcı dijital arşivlere aktarılması önemli bir konudur. Aktarım işlemi çok fazla içeriğin aynı anda otomatik sistemlerce sınıflandırılması ve standart bir katalog formunda yerleştirilmesi olarak da düşünülebilir. Ancak tüm veri kaynaklarının ve üstverilerinin buna uygun harmanlama ve bütünleştirme (aggregation) özelliklerine sahip olması gerekir. Oluşan ulusal dijital kamu içeriğinin etkin kullanımı için yapay zekâ teknolojileri, gelişen bilgi ve veri tabanı yönetimi uygulamaları ve veri analizi/veri madenciliği çalışmalarının incelenmesi gerekir. Gelecekte yönetim ve idari sistemler ile iş süreçleri ve bilgi teknolojisi entegrasyonunu sağlamaya dönük daha fazla yatırım yapılması gerekmektedir.

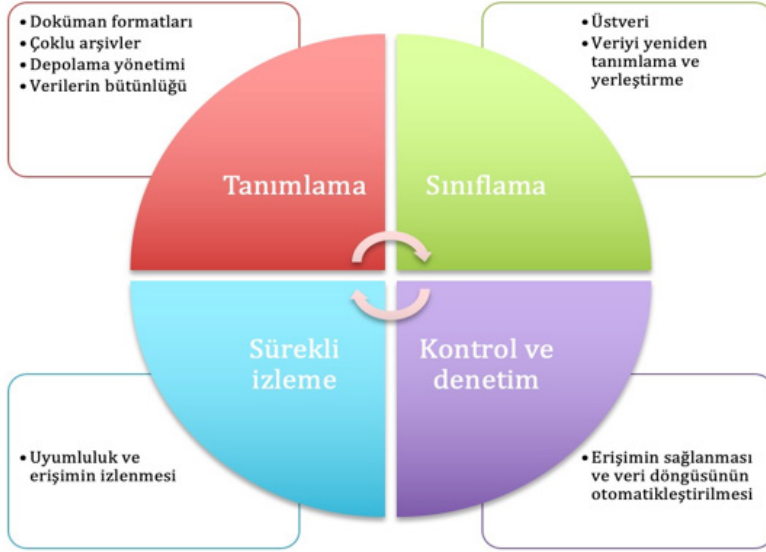
4.3. Veri Kaynağı Olarak Kamusal Bilgi ve Belgeler

Kurumsal bilgi yönetim sistemleri sadece bilgi ve belge kaynaklarının yönetimini değil veri yönetimi uygulamalarını da kapsamak durumundadır. Kurumsal sistemler üzerinde üretilen ya da sağlanan her türlü içeriğin oluşumundan arşivlenmesine ve veri analitiği uygulamaları kapsamında tekrar değerlendirilmesine tüm süreçler bu kapsamda değerlendirilmektedir. Bu kapsamda kurumlar içerisindeki karanlık verilerin (dark data) tespiti ve değerlendirilmesi de içerik yönetimi uygulamalarının bir parçasıdır (Gopalakrishnan, 2010).

Kurumlar bilgi sistemlerinde ürettikleri verileri iş veya süreç tamamlandığında basılı ya da dijital arşivlere kaldırır ve nadiren bu bilgilere tekrar başvururlar. Genellikle arşivlenen kurumsal bilgilere tekrar erişim için gelişmiş arşiv sistemleri de kullanılmaz. Karanlık arşiv (dark archive) olarak da tanımlanan, kurumsal arşivlerde yer alan ancak tekrar erişimin oldukça güç olduğu kaynaklara erişim için bilgi mimarilerinin planlanması, bu verilerin analizine dönük sistemler geliştirilmesi gerektiği üzerinde durulmaktadır. Karanlık veriler, bir kuruluşun normal operasyonlarının bir fonksiyonu olarak toplanan ancak nadiren veya hiç analiz edilmeyen veya akıllı iş kararları vermek için kullanılmayan verilerden oluşur (Ajis, 2019).

Düzenli yönetilmeyen kurumsal veriler kurumları önemli güvenlik açıklarına da maruz bırakabilir. Karanlık verilerin önemli bir kısmı denetim olmadan yaygın şekilde dağıtılır ve çoğaltılır, bu verilerin güvenliği genellikle göz ardı edilir. Karanlık veriler genellikle dene-

timsiz arşivlerde yer aldığından çalınmaya veya ihlale açıktır. Bu verilerin ele geçirilmesi kuruluşların rekabet gücünün kaybına ve çeşitli zararların oluşmasına neden olabilir. Buna ek olarak karanlık verilerin ihtiyaç duyulduğunda erişiminin sağlanamaması bir iş süreçlerini yavaşlatırken, gereksiz olanların depolanması ekstra maliyetlere yol açabilir (Kooiman, 2005).



Şekil 7. Karanlık veri yönetimi modeli (Ajis and Baharin, 2019)

Kurumların açığa çıkarılmamış değerleri olarak görülen karanlık arşivlerin depolama ortamı sorunları da dahil olmak üzere tüm koşullarının tanımlanması oldukça önemlidir. Bu arşiv kaynakları için hangi verilerin saklanması gerektiğini belirleyen tanımlama süreci ardından belge formatlarına karar verme, veri içeriğine göre çoklu arşivleme ve uzun süre saklama, değeri olan içeriğe karar verme işlemleri gerçekleştirilmelidir.

İkinci aşama sınıflama prosedürlerini içerir. Bu kapsamda kurumsal veri havuzunda saklanan içeriğe uygun meta verilerin atanması gerekir. İçeriğin yerleştirileceği alanların belirlenmesi sınıflama sürecinde yapılmalıdır.

Sürecin üçüncü aşaması idari ve yasal koşullarla uyum ve eşgüdümün sağlanmasına dönük kontrol aşamasını oluşturur. İçeriğin yönetimine dönük tüm aşamalar erişilebilir formatlarda kayıt altına alınmalıdır.

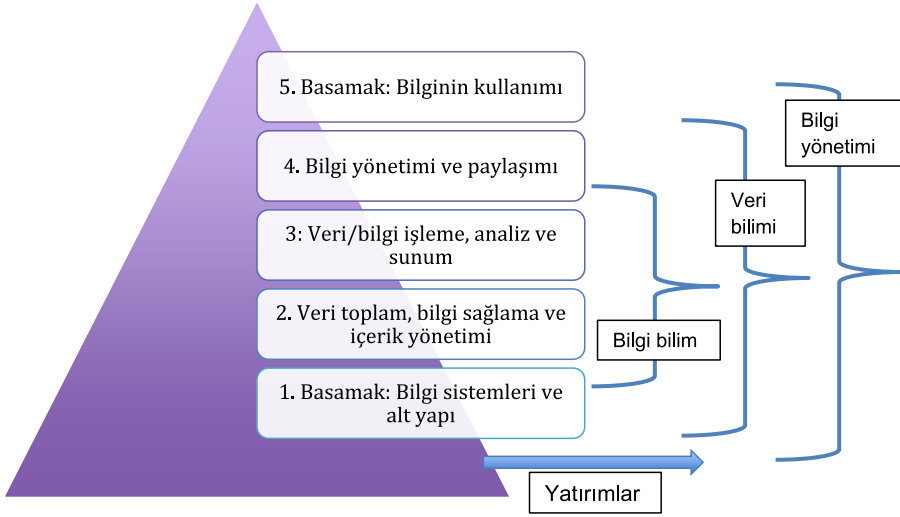
Son olarak, sistemler üzerinde içeriğin etkinliği ve erişilebilirliği sürekli olarak gözetim altında tutulmalıdır (Ajis and Baharin, 2019, s.36).

5. Veri Bilimi ve Veri Analitiği

Veri bilimi yeni nesil istatistik tekniklere dayanan, disiplinler arası bir araştırma metodolojisi sunan ve akıllı algoritmalarla işlenmemiş ham veriden anlamlı sonuçlar elde etmeye odaklanan bir disiplindir. Genel olarak veri bilimi veriyi çevresel koşullarla birlikte (organizasyonel ve sosyal etki alanları gibi diğer bağlamsal yönler dahil) incelemek için istatistik, bilişim, bilgi bilim, iletişim, yönetim ve sosyolojiyi sentezleyen yeni bir disiplinlerarası alan olarak tanımlanabilir. Bu tanıma göre veri bilimi = istatistik + bilişim + bilgi bilim + iletişim + sosyoloji + yönetim bilimidir (Longbing, 2017, 43:8). Veri biliminin çıktıları veri ürünleridir (Loukides 2011, 2012).

Bir veri ürünü, veriden elde edilebilir veri veya verilere dayanarak elde edilen keşif, tahmin, hizmet, öneri, karar verme, düşünme, model, mod, paradigma, araç veya sistem olabilir. Kullanım değerine sahip veri yerinde bilgi, doğru karar ve geliştirici güce dönüşebilir. Öte yandan sosyal medya platformları, yeni nesil medya uygulamaları hatta alışveriş siteleri gibi günlük yaşamımızdaki sistemler için de veri hem girdi hem de aynı zamanda işlenebilir çıktı kaynağıdır. Veri bilimi bize kurumsal, eğitimsel, etik, toplumsal, kültürel, ekonomik, politik, siber-fiziğe dönük yorumlarımıza katkı sağlarken yeni paradigmlar, eğilimler, stratejiler ve politikalar geliştirmemize yardımcı olmaktadır (Fairfielda ve Shteina 2014). Veri artık yaşamın her alanında, veri işleme, ölçüm ve doğrulama ihtiyacı da doğal olarak hızla artmaktadır (Ayankoya ve diğerleri. 2014). Giderek daha popüler hale gelen veri hizmetleri, veri akışını ve büyük veri alanını önemli ölçüde güçlendirmektedir. Bu tür sistemler arasında giyilebilir ürünler, Nesnelerin İnterneti (IoT), mobil ve sosyal uygulamalar yer almaktadır.

Günümüzde veri doğrulama ve veri ölçümü, herhangi bir zamanda, herhangi bir yerde, herhangi bir şekilde, genişlikte, derinlikte, çeşitlilikte ve hızda gerçekleştirilmektedir. Bu kapsamda nicel ve niteli veri kaynaklarını tanımakta yarar vardır. Nicel veri kaynakları arasında; çalışma hayatına, günlük yaşama, spora, eğlenceyi ve sosyalleşmeye ilişkin veriler yer alabilmektedir. Nitel veri kaynakları arasında; biyolojik sistemlere, fiziksel olaylara, davranışsal ve duygusal koşullara, bilişsel, çevresel, kültürel, ekonomik, sosyolojik ve politik sistemler sıralanmaktadır (Kelly 2012). Yine sağlık ve tıbbi alanlarda hem geleneksel tıp hem de sağlık bakım verileri (genomik, proteomik, mikrobiyomik, metabolomik, vb.) bu kapsamda sayılmaktadır.



Şekil 8. Bilgi işlem sürecinde veri bilimi (Longbing, 2017)

Veri bilimi dönemselsel popüler bir konu başlığı olarak kalmamış, inanılmaz bir inovasyon ortamı yaratmış ve yeni ekonomik fırsatlar oluşturmuştur. Veri araştırmaları, başlangıçta özel sektöre ait veri odaklı girişimlerden son zamanlarda büyük devlet kurumlarına ve akademik organizasyonlara yayılan geniş bir etki alanına sahip olmaya başlamıştır. Bilgi ve iletişim sektörünün dev şirketleri büyük veri odaklı projeler yürütürken Avrupa Birliği ve Birleşmiş Milletler gibi uluslar üstü organizasyonlar konuya ilişkin stratejiler geliştirmektedirler (USNSF, 2012, AB Commission, 2014). Veri biliminin gelişimi geleneksel veri odaklı çalışan mühendislik alanlarını yeniden şekillendirmekle kalmayıp aynı zamanda sosyal bilimlerin de veri mühendisliği etkisinde yeniden şekillendirilmesine yol açmaktadır (Yiu 2012; Labrinidis ve Jagadish 2012). Bu yeniden yapılanma verinin işlenmesini teknik bir alanın ötesine taşıyarak veriyi keşfetme, anlama ve yorumlaya dönük çok yönlü çalışmaların önünü açmaktadır (Longbing, 2017, 43.1-43.2). Veri biliminin, geleneksel bilimsel yaklaşımların kuram, deney ve yoruma dayalı, genellikle ampirik araştırmalardan oluşan çerçevesine dördüncü bir paradigma eklediğinden söz edilmektedir. Veri odaklı bilimsel yaklaşımların bilim ve teknolojinin, ekonominin yönlendirilmesine, yarının dünyasını yeniden şekillenmesine yol açacağı konuşulmaktadır (Hey ve diğerleri 2009).

Veri araştırmaları ve veri yönetimine dönük girişimler sadece belli çıkar grupları ya da sosyal toplulukların değil ülkelerin ve uluslararası organizasyonların da odağında yer almaktadır. ABD’de bu kapsamda girişimler arasında Büyük Veri Araştırma Girişimi (USNSF 2012), Dijital Verilerde İstikrarlı Çalışma Grubu (CSNSTC 2009), DARPA’nın XDATA Programı (DARPA 2016), ABD NSF Büyük Veri Araştırma Fonu (USNSF 2012). Ve yine US Big Data

Research Initiative (USNSF 2012), Interagency Working Group on Digital Data (CSNSTC 2009], DARPA's XDATA Program (DARPA 2016), USA NSF Big Data Research Fund (USNSF 2012) sayılabilir. Amerika Birleşik Devletleri (ABD) Büyük Veri Araştırma Girişimi (USNSF 2012), büyük veri çağına dönük temel bilimsel araştırmaları ve alt yapıyı desteklemeyi hedeflemektedir. Kamuya açık dijital bilimsel verinin tüm yelpazesine güvenilir, etkili bir erişim sağlayan kapsamlı ve şeffaf, geliştirilebilir, genişletilebilir sistemler, politikalar ve stratejiler geliştirmeyi amaçlamaktadır (USD2D 2016)

İngiltere'de İngiltere'nin Büyük Veri ve Enerji Verimli Hesaplama UK's Big Data and Energy Efficient Computing (UK., 2016) çalışmaları yapılmıştır. İngiltere'nin Araştırma Konseyleri tarafından finanse edilen Büyük Veri ve Enerji Verimli Bilgi İşlem girişimi (UK 2016), araştırmacıların, kullanıcıların ve endüstrinin ihtiyacı olan bilimsel keşif ve geliştirme için fırsatlar yaratmak amacıyla geliştirilmiştir.

Avustralya'da Kamu Hizmeti Büyük Veri Stratejisi [BM 2010], Veri Analitik Konusunda Hükümet Mükemmeliyet Merkezi (AU 2016) ve Public Service Big Data Strategy (UN 2010), Whole-of-Government Centre of Excellence on Data Analytics oluşturulmuştur. Avustralya Kamu Hizmeti Büyük Veri Stratejisi (BM 2010), "büyük verilerin kullanımıyla ilgili kurumlara sunulan fırsatları, olanakları ve tehditleri ortaya koymayı amaçlamaktadır (AGIMO 2013).

Kanada'da Büyük Verilerden Yararlanma (Capitalizing on Big Data) Stratejisi geliştirilmiştir. Kanada'nın Büyük Veriden Yararlanmasına İlişkin Politika Çerçevesi (CA 2016) "bir yönetim kültürü oluşturmayı hedeflemektedir

Çin'de Büyük Veri Rehberi, Çin Büyük Veri Üzerine Bilgisayar Federasyonu Görev Gücü (CCF-BDTF 2013), Çin Ulusal Bilim Vakfı büyük veri programı (CNSF 2015) ve Big Data Guideline, China Computer Federation Task Force on Big Data (CCF-BDTF 2013), China National Science Foundation big data program geliştirilmiştir. Çin'de yayımlanan kılavuzları büyük veri araştırma ve uygulamalarının geliştirilmesini, büyük veri geliştirme ve uygulama için genel bir koordinasyon mekanizması oluşturulmasını. Çin ayrıca IoT ve büyük veriler için ulusal bir stratejik plan hazırlamıştır (CNSF. (2015).

Avrupa Birliği; Veri Odaklı Ekonomi (European Commission, 2014). Avrupa Komisyonu Horizon 2020 Büyük Veri Özel Kamu Ortaklığı (Horizon 2014 ve Data-driven Economy) (European Commission, 2014). European Commission Horizon 2020 Big Data Private Public Partnership bu kapsamda önemli içeriklere sahiptir. Avrupa Birliği'nin Gelişmekte Olan Veri Odaklı Bir Ekonomiye Doğru İletişimi [EU 2014], geleceğin veri odaklı ekonomisini ortaya çıkarmak için bir eylem planıdır. Avrupa'da veri odalı dayalı ekonomi, veriler üzerinde araştırma ve inovasyonu teşvik etmeyi, özellikle Avrupa genelinde KOBİ'ler için bilgi ve sermaye kullanı-

minin artırılmasını amaçlamaktadır. Bu çerçevede 2015 yılında, Avrupa Veri Bilimi Akademisi, EDSA (European Commission, 2014) kurulmuştur.

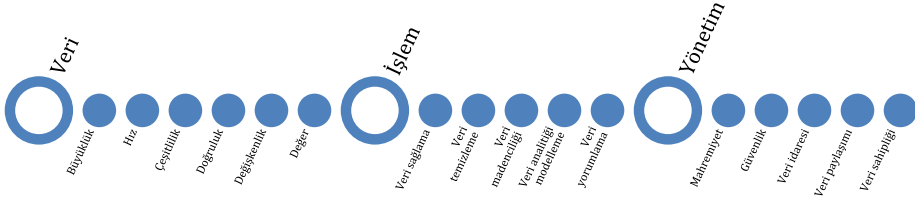
Birleşmiş Milletler BM Küresel Etki Projesi (UN Global Pulse Project; UN 2010) bu kapsamda değerlendirilebilir. Birleşmiş Milletler (BM) Küresel Etki Projesi, Birleşmiş Milletler Genel Sekreteri'nin büyük veriler konusunda amiral gemisi bir inovasyon girişimidir. (Longbing, 2017, 43:11). Misyonu, sürdürülebilir kalkınma ve insani faaliyetler için büyük veri yenilikçiliğinin keşfedilmesi, geliştirilmesi ve ölçeklendirilmesinin hızlandırılmasıdır (UN 2010).

Öte yandan bir kavram olarak büyük verinin nasıl tanımlanacağına ilişkin belirsizlikler sürmektedir. Kavramın 1990'ların ortalarında başlayan veri çalışmalarını tanımlamak için ortaya çıkışından (Diebold, 2012) sonra konu üzerine çalışan ticari firmaların da büyük veri kavramının gelişiminden rolü olduğu düşünülmektedir (Gandomi ve Haider, 2015, s138). Verinin büyüklüğü tek başına büyük veri çalışmalarının ana eksenine oturmamaktadır. Laney (2001)'nin verinin hacmi (volüme), çeşitliliği (variety) ve hızının (velocity) veri yönetimin zorluklarının üç boyutu olduğunu tanımlamasının ardından Üç V'ler, büyük veriyi tanımlamak için ortak bir çerçeve oluşturmuştur (Chen, Chiang ve Storey, 2012; Kwon, Lee ve Shin, 2014). Büyük veri, bilgilerin sağlanması, depolanması, dağıtılması, yönetimi ve analizini sağlamak için ileri teknikler ve teknolojiler gerektiren büyük hacimli, yüksek hızda, karmaşık ve değişken verileri tanımlayan bir terim olarak adlandırılmaktadır (TechAmerica Vakfı Federal Büyük Verileri) Komisyon, 2012).

Büyük veri kapsamında: Hacim, verilerin büyüklüğünü ifade eder. Büyük veri boyutları terabaytlarda ve hatta petabaytlar boyutunda olabilir. IBM'in 2012 ortasında yaptığı bir araştırma, 1144 katılımcının yarısından fazlasının bir terabayt üzerindeki veri kümelerinin büyük veri olarak kabul edildiğini ortaya koymuştur (Bir terabayt, yaklaşık 16 milyon Facebook fotoğrafını depolayacak kadar 1500 CD veya 220 DVD'ye sığacak kadar veri anlamına gelmektedir). Facebook'un saniyede bir milyona kadar fotoğrafı işlediğini bildirmektedir. Bir petabayt, 1024 terabayta eşittir. 2015 yılındaki bir çalışmada Facebook'un 20 petabayttan fazla depolama alanı kullanarak 260 milyar fotoğraf sakladığı ortaya konulmuştur (Gandomi ve Haider, 2015). Ancak büyük veri hacimlerinin tanımları görecelidir ve zaman ve veri türü gibi faktörlere göre değişiklik gösterebilir.

Büyük veriler çok sayıda platformlardaki içeriği (örneğin makineden makineye iletişim, sosyal medya siteleri, sensörler, ağlar), multimedya içeriği (örneğin videolar, görüntüler ve ses dosyaları), metinsel içerik (yapılandırılmış, yarı yapılandırılmış ve yapılandırılmamış) kapsamına alır.

Büyük veri çalışmaları aşağıdaki kapsamda üç başlık altında toplanmaktadır.



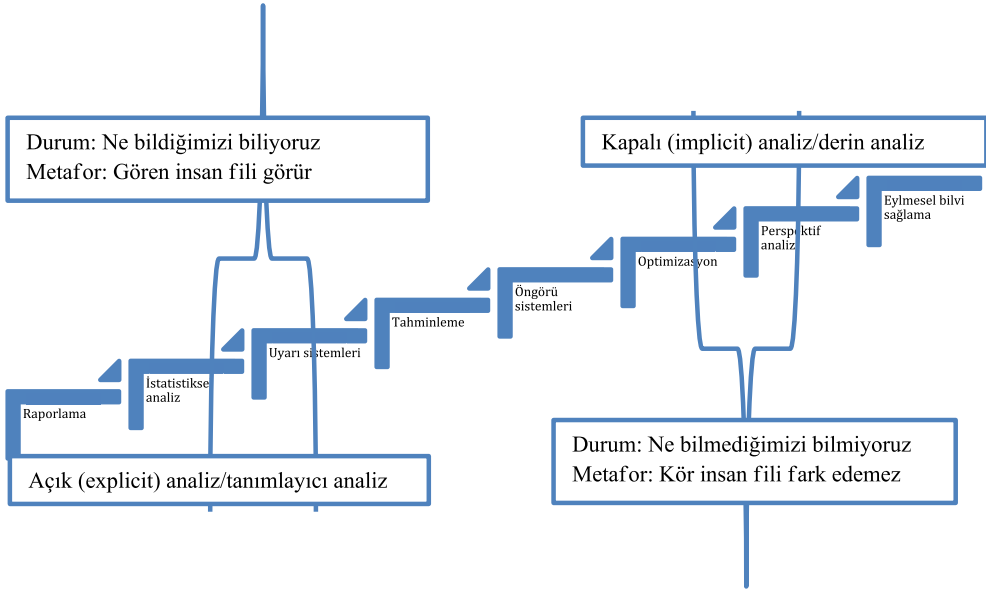
Büyük veri hala gelişen bir kavramdır. Ancak günümüzde artık büyük veri endüstri ve akademi dünyası için katma değerli bir araştırma konusuna dönüşmüştür. Büyük veri analitiği tekniklerdeki önemli yenilikler henüz gerçekleştirme aşamasındadır. Literatürde de büyük veri analitiği çalışmamalarının sınırlı kaldığı, bu araştırmaların önemli bir kısmının gerçek zamanlı veri analitiğine kayacağı, makine öğrenme ve yapay zekâ teknolojilerini destekleyeceği üzerinde durulmaktadır (Eckerson, 2010).

4.1. Veri Analitiği

Veri analitiği çalışmaları kapsamında neyin analiz edileceği, analiz edilecek verilerin özelliklerinin ne olduğu ve hangi analiz araçlarının kullanılacağına açıklanması gerekir. Bu noktada:

- Verilerin geçmişten günümüze ve geleceğe yaşam döngüsü,
- Açık (bilinen) veri analitiği, dolaylı olarak reaktif anlayış (bilinmeyen koşullar için) ve proaktif (erken tahmin ve müdahale koşulları) koşullar,
- Veri araştırmalarında (tanımlayıcı ve öngörücü analitikler yoluyla) öngörülebilir analiz araçlarının tanımlanması gerekir.

Aşağıdaki şekilde yer aldığı gibi veri analitiği ve analitik verinin kullanımı üzerine modeller geliştirilmelidir. Modellerde geçmişe ait verileri iş araştırmaları, modellemeler, deneysel tasarımlar, sürece ilişkin nasıl ve neden sorularına yanıt vermek için kullanılabilir. Güncel veriler var olan süreçleri doğru yapılandırmak için son derece önemlidir. Geleceğe ilişkin veriler ise daha çok tahminleme ve planlama süreçlerinde kullanılır. Gelecek süreçlerin tahminlenmesi proaktif önlemlerin alınabilmesi için de gereklidir (Longbing, 2017, 43.18)

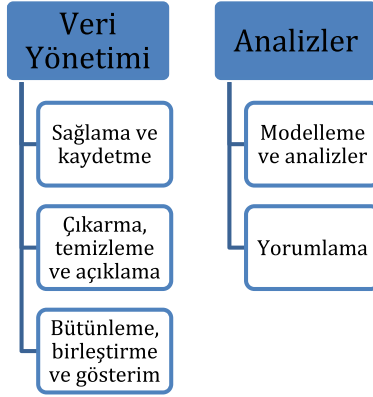


Şekil 10. Açık ve kapalı analiz spektrumu ve gelişimi (Longbing, 2017)

Veri analizi geleneksel araç ve tekniklerle (örneğin, klasik istatistiksel, matematiksel veya mantıksal teoriler) ortaya çıkarılamayan içeriğin derinlemesine anlaşılması ve keşfedilmesini sağlamaktadır. Bu kapsamda tanımlayıcı analiz ve tahmine dayalı analitik çalışmalara imkân veren yazılım platformları ya da Knime, Orange ya da Tableau benzeri programlardan yararlanılabilmektedir. Yine tanımlayıcı analizlerde ortaya konulamayan nedenler, süreçler ve tahminler derin analiz teknikleriyle açıklanabilmektedir (Longbing, 2017, 43).

Önceden üretilen ve depolanan veri kaynakları, üzerinde analitik bir çalışma yapılmadığı sürece değersizdir. Büyük verileri olarak nitelendirilen verilerin %95'inin yapılandırılmamış olması bunların analizini daha değerli kılmaktadır. Büyük verinin öncelikli potansiyel değeri, karar verme sürecini etkinleştirmektir. Kurumların kanıta dayalı karar vermeyi mümkün kılmak için yüksek hacimli verileri hızlı analiz edebilen farklı veri türleriyle çalışabilecek araçlara ihtiyacı vardır. Bu kapsamda çalışmalar aşağıdaki şekilde görülebileceği gibi 5 aşamada yürütülebilir (Labrinidis ve Jagadish, 2012; Uslu, 2016). Bu beş aşama iki ana alt süreci oluşturur: veri yönetimi ve analitik. Veri yönetimi, veri toplama ve depolama ve analiz için hazırlama ile veri analizi için süreçleri ve destek teknolojilerini içerir. Veri analizleri, büyük veriden anlamlı sonuçlar elde etmek için kullanılan teknikleri ifade eder. Bu nedenle, büyük veri analitiği, genel olarak büyük veriden anlamlı sonuçlar çıkarma sürecinde bir alt süreç olarak görülebilir.

Aşağıda yapılandırılmış ve yapılandırılmamış veriler için veri analitiği uygulamaları ve-
rilmektedir.



Şekil 11. Veri işleme kapsamı (Fresa, Justrell, Prandoni, 2015, s.196)

Veri analitiği, bilgi sistemleri ve yazılım uygulamaları konusunda teknik alt yapıya sahip olmayan sıradan insanlar için bile öğrenilebilir ve uygulanabilir bir araştırma alanı olmaya başlamıştır. Makine öğrenmede yeni yaklaşımlar, bilgisayara dayalı görü sistemleri ve doğal dil işleme alanındaki gelişmeler (Business Intelligence, 2019; Skeina, 2017, s.16 Fresa, Justrell, Prandoni, 2015, ss.196-197) bilgi hizmetlerine de yeni ufuklar açmaktadır.

Veri analizi süreci ve bu sürecin sonunda verinin anlamlı bilgiye dönüşümü aşağıdaki aşamaları içermektedir:



Şekil 12. Veri analizi kapsamı (Uthayasankar, Kamal, Irani, Weerakkody, 2017, s. 266)

Tanımlayıcı Analitik (Descriptive Analytics): Tanımlayıcı analitiği, tarif edici olarak da ifade edebiliriz. Veriyi (iş uygulamalarında işlemi) tanımladığımız, onunla ilgili bilgi edindiğimiz analitik çalışmalar tanımlayıcı analitik olarak ifade edilir. Veri hakkında tanımlayıcı bilgileri ortaya koymaktadır. Ne oldu? sorusuna cevap aldığımız uygulamalardır. İstatistik, veri madenciliği, iş zekası uygulamalarında sıklıkla kullanılır.

Teşhis/Öngörücü (Diagnostic Analytics): Öngörücü analitik uygulamaları, veriye ya da içeriğe ilişkin tanı/teşhis ve öngörülerin yapılmasına imkân veren teknikleri içerir. İçerisinde örüntü (pattern), veri madenciliği (data mining), korelasyon uygulamaları vardır. Genellikle neden oldu? sorusuna yanıt vermek için kullanılır.

Tahmine Dayalı Analitik (Predictive Analytics): Genellikle tahminleme geçmişe ya da geleceğe yönelik olarak belirli bir durum olması/olmaması halinde satış, karlılık, performans gibi metriklerin nasıl değişkenlik göstereceğini belirlemede kullanılır. Ne olacak en karakteristik sorusudur (Massaro, A., Vitti, V., Galiano, A., Morelli, A., 2019).

Tahmine dayalı analiz, tarihsel ve güncel verilere dayanarak gelecekteki sonuçları öngören çeşitli teknikleri içerir. Tahmin edici analitikler neredeyse tüm disiplinlere uygulanabilir. Bu alanda çalışmalar jet motorlarının yüzlerce sensörden gelen veriyi analiz edilmesinden, müşterilerin satın alma eğilimlerine kadar çok geniş bir alana yayılır. Temelde, tahmine dayalı analitik, modelleme yapmayı ve veriler arasındaki ilişkileri tespit etmeyi amaçlar. Tahmine dayalı analiz teknikleri iki gruba ayrılır. Ortalamaların taşınması gibi bazı teknikler, sonuç değişkenlerindeki tarihsel kalıpları keşfetmeyi amaçlar. Doğrusal regresyon gibi teknikler, sonuç değişkenleri ile açıklayıcı değişkenler arasındaki karşılıklı ilişkileri yakalamayı amaçlar. Bu tür çalışmalarda genel olarak regresyon ve makine öğrenme teknikleri (örneğin karar ağaçları ve sinir ağları) kullanılır. Tahmine dayalı analitik teknikler popülasyondan küçük bir örnek alınır ve örneğe ilişkin sonuca dayanarak büyük kümeye ilişkin genellemelere ulaşılmaya çalışılır (Fan, Han ve Liu, 2014). Bu kapsamda veri bilimciler aşağıdaki sorular öncelikle sorgulamaktadırlar:

- Belirli bir veri setinden neler öğrenebilirsiniz?
- Toplum gerçekte ne bilmek istiyor?
- Bilmek toplum için ne anlama geliyor? (Skeina, 2017, s.16).

Bazı veri kümeleri, bir veri tabanlarında veya elektronik tablo programındaki tablolar gibi güzel bir şekilde yapılandırılmıştır. Yapılandırılmamış veriler daha çok gerçek zamanlı veri sağlayıcılardan elde edilir. Yapılandırılmış veriler genellikle bir matrisin satırlarının farklı öğeleri veya kayıtları temsil ettiği ve sütunların bu öğelerin farklı özelliklerini temsil ettiği bir matris ile temsil edilir. Örneğin, bir ülkenin şehirleri ile ilgili veriler, her şehir için eyalet, nüfus ve alan gibi özellikleri temsil eden sütunlarla birlikte bir satırı temsil eder.

Sosyal medyada herhangi bir konuda atılan mesajlar ilgili veri kümesi gibi yapılandırılmamış bir veri kaynağıyla karşı karşıya kaldığımızda ise, ilk adımımız genellikle onu yapılandırmak için bir matris oluşturmaktır (Skeina, 2017, s. 29).

5.2. Veri Bilimi Kapsamında Makine Öğrenme

Günümüzde bilgi yönetimindeki en heyecan verici gelişmeler güçlü ve yeni algoritmalarla heyecan verici yeni uygulamalarla karşımıza çıkan makine öğrenmesi etrafında dönmektedir. Bu gelişmeler birkaç nedenden ötürü gerçekleşmiştir. İlk olarak, mevcut veri hacmi ve bilgi işlem gücü, makine öğrenme sistemlerinin eski yaklaşımları kullanarak bile ilginç şeyler yapmaya başladığı sihirli bir eşiği geçmiştir. Bu durum aynı zamanda daha iyi ölçeklendirme yöntemleri geliştirme konusunda daha fazla yatırım yapılmasına yol açmıştır.

Makine öğrenme temelde aşağıdaki yaklaşımlara dayanır:

Güç ve açıklık: Makine öğrenme yöntemleri, destekledikleri modellerin zenginliği ve karmaşıklığı bakımından farklılık gösterir. Doğrusal regresyon doğrusal fonksiyonlara dayanır. En yakın komşu metotları ise rasgele eğrileri yaklaştırmaya yetecek kadar parçalı doğrusalların ayırımı mümkün olur.

Yorumlanabilirlik: Derin öğrenme gibi güçlü yöntemler genellikle tamamen aşlamayan modeller üretir. Uygulamada çok doğru bir sınıflandırma sağlayabilirler. Buna karşılık, doğrusal regresyon modelindeki en büyük katsayılar en güçlü özellikleri tanımlar ve en yakın komşuların kimlikleri bu analogilere olan güvenimizi bağımsız olarak belirlememizi sağlar.

Kullanım kolaylığı: Bazı makine öğrenme yöntemleri göreceli olarak az sayıda parametre veya karar içerir. Doğrusal regresyon ve en yakın komşu sınıflandırması bu açıdan oldukça anlaşılabilir bir yapıdadır. Buna karşılık, destek vektör makineleri (SVM'ler) gibi yöntemler algoritma performansını uygun ayarlarla optimize etmek için çok daha fazla içerik sunar.

Eğitim hızı: Metotlar, modelin gerekli parametrelerine ne kadar hızlı uydukları konusunda büyük farklılıklar gösterir; bu da pratikte ne kadar deney verisi kullanabileceğinizi belirler. Geleneksel doğrusal regresyon yöntemleri, büyük modeller için uygun olmayabilir. Buna karşılık, en yakın komşu araması, uygun arama veri yapısını oluşturmanın dışında neredeyse hiç eğitim zamanı gerektirmez.

Tahmin hızı: Yeni bir sorgulama için sınıflandırma kararlarını ne kadar hızlı yaptıkları ile ilgilidir. Doğrusal regresyon hızlıdır, sadece girdi kayıtlarındaki alanların ağırlıklı bir toplamını hesaplar. Buna karşılık, en yakın komşu araştırması, önemli miktarda eğitim testine karşı q testini gerektirir. Genelde veri eğitim oldukça hızlıdır.

Tablo 2. Makine Öğrenmede En yaygın Kullanılan Modeller ve Özellikleri Skeina, 2017, ss. 351-353)

	Açıklama gücü	Tercüme kolaylığı	Kullanım kolaylığı	Eğitim hızı	Tahmin hızı
Lineer regresyon (Linear regression)	5	9	8	9	9
En yakın komşu (Nearest neighbor)	5	9	10	10	2
Naive Bayes	4	8	9	9	8
Kara ağacı (Decision tree)	8	8	7	7	8
Support vector machine	8	6	7	7	7
Boosting grafik modeli	9	6	6	6	6
Grafik modeller	9	8	4	4	4
Derin Öğrenme	10	3	3	3	7

5.3. Derin Öğrenme

Derin öğrenme makine öğrenmede ayrılan heyecan verici bir gelişmedir. 1980'lerin popüler bir yaklaşımı olan ve daha sonra tamamen modası geçmiş olan sinirsel ağlarına dayanır. Özellikle 2010'lu yıllarla birlikte çok katmanlı (derin) ağlar, bilgisayara dayalı görü sistemleri ve doğal dil işlemedeki gelişmeler derin öğrenmenin gelişimini tetiklemiştir. Google'ın Tensor-Flow gibi yeni açık kaynaklı yazılım çerçeveleri, ağ mimariler, büyüklük sırasına göre verinin eğitimi hızlandırmak için tasarlanan özel amaçlı işlemciler, bu geliştirmeleri kolaylaştırmıştır.

Derin öğrenmeyi diğer yaklaşımlardan ayıran temel farkı sadece teknik ve mühendislik süreçlerine dayanmamasıdır. Bir sinir ağındaki her katman, genel olarak bir önceki katmanın çıktısını girdi olarak kabul eder ve ağıın tepesine doğru ilerlerken aşamalı olarak daha yüksek özellikler sunar. Bu durum ham girdiden nihai sonuca kadar olan bir anlayış hiyerarşisinin tanımlanmasını sağlar, daha organik ve özgün sonuçların olmasını tetikler (Skeina, 2017, s. 377).

5.4. Bilgi Kaynağı Olarak Metin Analizi

Metin analizi (metin madenciliği, text mining), metinsel verilerinden anlamlı bilgiler çıkarmak için gerçekleştirilir. Sosyal ağ kaynakları, e-postalar, blog yazıları, çevrimiçi forum içerikleri, anket cevapları, kurumsal belgeler, arşiv kaynakları, kitaplar, haberler ve çağrı merkezi kayıtları, organizasyonlar tarafından tutulan metin verisi örnekleridir. Metin analizi, istatistiksel analiz, hesaplamalı dil bilimi ve makine öğrenimini içerir. Metin analizi, işletmelerin büyük miktarlarda insan kaynaklı metni, kanıt temelli karar vermeyi destekleyen anlamlı özetlere dönüştürmelerini sağlar. Örneğin, finansal analizlerden çıkarılan bilgilere dayanarak

borsa tahmininde metin analizleri kullanılabilir (Chung, 2014). Aşağıda metin analizi yöntemlerinin kısa bir derlemesini yer almaktadır.

Bilgi çıkarma (IE) teknikleri, yapılandırılmamış metinden yapılandırılmış verinin çıkarılmasına dayanır. Örneğin, IE algoritmaları ilaç adı, dozu ve tıbbi reçetelerden sıklık gibi yapısal bilgileri çıkarabilir. IE'deki iki alt görev, Varlık Tanıma (ER) ve İlişki Çıkarma (RE) (Jiang, 2012). ER, adları metinde bulur ve bunları kişi, tarih, konum ve organizasyon gibi önceden tanımlanmış kategorilere sınıflandırır. RE, metindeki varlıklar (örneğin kişiler, kuruluşlar, ilaçlar, genler, vb.) arasındaki anlamsal ilişkileri bulur ve çıkarır. Örneğin, "1976'da Apple Inc.'in kurucusu Steve Jobs" cümlesi dikkate alındığında, bir RE sistemi FounderOf [Steve Jobs, Apple Inc.] veya FoundedIn [Apple Inc., 1976] gibi ilişkileri çıkarabilir.

Metin özetleme teknikleri otomatik olarak tek veya çoklu belgelerin kısa bir özetini oluşturur. Sonuçta ortaya çıkan özet, asıl metin (ler) deki anahtar bilgiyi taşır. Uygulamalar bilimsel ve haber makaleleri, reklamlar, e-postalar ve blogları içerir. Genel olarak özetleme iki yaklaşımı izler: çıkarıcı yaklaşım ve soyutlayıcı yaklaşım. Çıkarımlı özetlemede, orijinal metin birimlerinden bir özet oluşturulur (genellikle cümleler). Ortaya çıkan özet, orijinal belgenin bir alt kümesidir. Çıkarma yaklaşımına dayanarak, bir özet oluşturmak bir metnin göze çarpan birimlerinin belirlenmesini ve bir araya getirilmesini içerir. Metin birimlerinin önemi, metin içindeki yerleri ve sıklıkları analiz edilerek değerlendirilir. Ekstraktif özetleme teknikleri metnin "anlaşılmasını" gerektirmez. Buna karşılık, özet teknikler metinden anlamsal bilginin metinden çıkarılmasını içerir. Özetlerde, mutlaka orijinal metinde bulunmayan metin birimleri bulunur. Özgün metni ayrıştırmak ve özeti oluşturmak için soyut özetleme, gelişmiş Doğal Dil İşleme (NLP) tekniklerini içerir. Sonuç olarak, soyutlayıcı sistemler, çıkarıcı sistemlerden daha tutarlı özetler üretme eğilimindedir (Hahn ve Mani, 2000). Bununla birlikte, ekstraktif sistemleri, özellikle büyük veriler için kullanmak daha pratiktir.

Soru cevaplama (QA) teknikleri, doğal dilde verilen sorulara cevaplar sağlar. Apple'ın Siri'i ve IBM'in Watson'ı ticari KG sistemlerine örnektir. Bu sistemler sağlık, finans, pazarlama ve eğitim alanlarında uygulanmıştır. Özet özetlemeye benzer şekilde, KG sistemleri karmaşık NLP tekniklerine dayanır. Teknikler üç kategoriye ayrılır: bilgi edinme (IR) temelli yaklaşım, bilgiye dayalı yaklaşım ve hibrid yaklaşım. IR tabanlı KG sistemlerinde genellikle üç alt bileşen bulunur. Birincisi, soru türü, soru odağı ve sorgu oluşturmak için kullanılan cevap türü gibi ayrıntıları belirlemek için kullanılan soru işleme yöntemidir. İkincisi, soru işleminde formüle edilmiş sorguyu kullanarak ilgili önceden yazılmış bölümleri bir dizi mevcut dokümandan almak için kullanılan doküman işlemedir. Üçüncüsü, önceki bileşenin çıktısından aday cevapları çıkarmak, onları sıralamak ve QA sisteminin çıktısı olarak en yüksek

derecedeki adayı geri vermek için kullanılan cevap işlemedir. Bilgi temelli KG sistemleri, daha sonra yapılandırılmış kaynakları sorgulamak için kullanılan sorunun anlamsal bir tanımını oluşturur. Bilgiye dayalı KG sistemi, özellikle önceden yazılmış çok sayıda belgenin bulunmadığı turizm, tıp ve ulaşım gibi kısıtlı alanlar için kullanışlıdır. Bu tür etki alanları, IR tabanlı QA sistemleri için gerekli olan veri fazlalığına sahip değildir. Siri türü sistemler de bilgi temelli yaklaşımı kullanan QA sistemi örnekleridir. Karma QA sistemlerinde, IBM'in Watson'ı gibi, soru anlamsal olarak analiz edilirken, aday cevaplar IR yöntemleri kullanılarak oluşturulmaktadır.

Metin madenciliği kapsamında duygu analizi (fikir madenciliği) teknikleri, insanların ürünler, organizasyonlar, bireyler ve etkinlikler gibi varlıklara yönelik görüşlerini içeren doğrulanmış metni analiz eder. İş dünyası, müşterilerinin duyarlılık analizlerinin artmasına neden olan duyguları hakkında giderek daha fazla veri toplamaktadır (Liu, 2012). Pazarlama, finans ve siyasal ve sosyal bilimler, duyarlılık analizinin ana uygulama alanlarıdır. Duyarlılık analizi teknikleri ayrıca, belge düzeyinde, cümle düzeyinde ve en boy temelli olmak üzere üç alt gruba ayrılmıştır. Belge düzeyinde teknikler, tüm belgenin olumsuz mu yoksa olumlu bir duygu mu ifade ettiğini belirler. Varsayım, belgenin tek bir varlık hakkındaki düşünceleri içerdiği varsayılmaktadır. Belli teknikler bir dokümanı iki sınıfa ayırırken, negatif ve pozitif, diğerleri ise daha fazla duyarlılık içermektedir (Amazon'un beş yıldızlı sistemi gibi) (Feldman, 2013). Cümle düzeyinde teknikler, tek bir cümlenin, tek bir cümleyle ifade edilen bilinen bir varlık hakkındaki kutupluluğunu (olumlu uçtan olumsuz uca) belirlemeye çalışır. Cümle düzeyindeki teknikler önce öznel söylemleri objektif olanlardan ayırmaktadır. Dolayısıyla cümle seviyesi teknikleri, belge seviyesi teknikleriyle karşılaştırıldığında daha karmaşık olma eğilimindedir. Boyuta dayalı teknikler, bir belgedeki tüm duyguları tanır ve her bir hissini başvurduğu varlığın yönlerini tanımlar. Örneğin, müşteri ürün incelemeleri genellikle bir ürünün farklı yönleri (veya özellikleri) hakkında görüşler içerir. En boy temelli teknikleri kullanarak kullanıcı ya da alıcı, polarite açısından sınıflandırılmışsa, ürünün farklı özellikleri hakkında değerli bilgiler alabilir (Gandomi ve Haider, 2015, s140).

5.5. İş Zekâsı ve İş Analitiği

En genel tanımıyla iş zekâsı ve analitiği, kurumsal bir iş üzerine karar süreçlerini geliştirmek için bir dizi yöntem, teknoloji ve araçların kullanımınıdır (Foley ve Guillemette, 2010). Kurumsal bilgi sistemleri ve teknolojilerindeki gelişmeler, 1990'ların sonlarında iş zekâsı sistemlerinin ortaya çıkmasını sağlamıştır. İş zekâsı (BI) sistemleri organizasyonların, fırsatlarını, güçlü yönlerini ve zayıflıklarını anlamak için kullanılabilir. Bunun için geniş veri kaynaklarının oluşturulması entegrasyonu ve analizi gerekir. BI, karar süreçlerini destekleyen

bir bilgi sistemidir. BI sistemlerinin hedefleri arasındadır; a- daha fazla toplanmaya, sistematik entegrasyona ve yapılandırılmamış ve yapılandırılmış verilerin yönetimine yardımcı olmak, b- çok büyük miktarda veri ile uğraşmak (örneğin, “Büyük Veri”), c- yeni bilgileri keşfetmek için işlem becerilerini artırmak ve d- analiz çözümleri, geçici sorgular, raporlama ve tahminler sunmak yer almaktadır (Katerina ve Dimovski, 2019).

Günümüzde “iş zekâsı” terimi “iş analitiği” terimi ile birlikte kullanılmaya başlanmıştır. Aslında iş analitiği yeni ortaya çıkan bir terim değildir. 1940’larla birlikte bilimsel literatürde iş analitik kavramı kullanılmaya başlanmıştır. Bilgisayarlar karar destek sistemleri olarak 1960’ların sonunda itibaren hayatımıza girmeye başlamıştır. Bununla birlikte ilk dönemlerde veri kaynakları nispeten küçük ve yapılandırılmış kaynaklardan oluşmuş ve analitik çalışmalarının büyük çoğunluğu tanımlayıcı analitik çalışmalardan meydana gelmiştir. Tanımlayıcı analitik (işletme zekâsı (BI) veya işletme raporlaması olarak da adlandırılır), analitik taksonomisine giriş seviyesi olarak görülmektedir. Bu seviyedeki analitik faaliyetlerin çoğu, geçmiş ya da mevcut durumu bildirmek için ticari faaliyetleri özetleyen raporlar yaratma ile ilgilidir (Delen, 2014). Akademik camiada 2001’de Analytics 2.0 kavramının tartışılmaya başlanmasıyla veri bilimi hızla gelişmiştir. 2000’li yılların başında büyük veri ve ileri analitik (tahmine dayalı ve öngörücü) iş dünyasında ortaya çıkmaya başlamış ve geleneksel iş zekâsından gerçek zamanlı veya doğru zamanda karar destek verisine geçiş gerçek bir paradigma değişikliğini meydana getirmiştir (Delen, Moscato, Toma, 2018).

BI sistemleri, içeriğinde veri ambarları, çevrimiçi analitik işlem (OLAP) ve gösterge panoları (data warehouse, online analytical processing (OLAP) and dashboards) vardır. Bir veri deposu, derinlemesine analiz için birçok kaynaktan elde edilen doğru, temiz ve ayrıntılı verileri toplar, çevrimiçi analitik işlem (OLAP) ise gerçek zamanlı olarak çok boyutlu analizi destekler ve kullanıcıların toplama, filtreleme, yuvarlama ve açıklama gibi işlemleri yapmalarını sağlar. Ayrıca, veri görselleştirme ve performans yönetimi için ön uç uygulama olarak panolar (dashboard) kullanılır. BI sistemleri işletmenin temel performans göstergelerini takip etmek için kullanıcıların grafikler, çizelgeler, widget’lar ve geçici raporlar ve karar vericiler oluşturmasına olanak sağlar (Clark, Jones, ve Armstrong, 2007).

İş zekâsı uygulamaları aşağıdaki katmanlardan oluşur (Scholtz, Calitz ve Haupt, 2016)

- Tanımlama
- Dönüştürme ve yükleme
- Veri katmanı

- Raporlama katmanı
- Analiz katman
- İzleme katmanı
- Sunum katmanı

Tanımlama iş zekâsı çerçevesinde kullanılacak verilerin belirlenmesi ile ilgilidir. Dönüş-türme ve yükleme verinin hazırlanması aşamasıdır. Veri katmanı, kurumun tüm sürdürülebi-lirlik verilerini depolamak için kullanılacak veri tabanı yönetim sistemleri ve ambarlarının oluşturulması ile ilgilidir (Muntean ve diğerleri, 2013). Raporlama katmanı, geleneksel BI araçlarının bir parçasını oluşturan sorgulama ve raporlama özelliklerinden oluşur Özel sor-gulama özelliği, kullanıcıların filtrelemek ve istenen verileri seçmek için bir veri tabanı sor-gulamasını sağlar (Ong ve diğ., 2011). Belirli bir soruyu cevaplayacak bilgileri bulmak için sorgu alanları oluşturulabilir. Raporlama sorgulama alanlarındaki verinin değerlendirilmesi ile elde edilir (Scholtz, Calitz and Haupt, 2016, s.270).

Analitik katmanda, profesyonel analiz araçları geleneksel BI bileşenlerinin çok daha öte-sinde çözümler sunar. Bu kapsamda OLAP (Online Analytical Processing), veri madenciliği, öngörme ve öngörücü modelleme gibi analitik araçlar sağlar. Bu yetenekler bir kurumun karar alma niteliklerini geliştirmesine yardımcı olabilir. Tahmin, tarihsel verilere dayanarak gelecekteki sonuçları tahmin etme sürecidir. Bu yetenekler aynı zamanda iş analitiği alanının temel yetenekleridir. Bu yetenekler stratejik zekanın oluşturulmasını sağlayan temel bileşen-lerdir. Analitik katmanda oluşturulan bilgiler genellikle izleme katmanında yönetime sunulur (Kourentzes ve diğerleri, 2014). İzleme katmanı, bir kurumun yönetimine gerçek zamanlı stratejik bilgiler sağlayarak sürdürülebilirliğe katkı sağlar. Bu noktada performans gösterge panoları ve puan kartlarından (Dashboards and scorecards) yararlanılır (Yiğitbaşıoğlu ve Vel-cu, 2012). Sunum katmanı, tüm kullanıcıların diğer katmanlar tarafından sağlanan bilgilere mobil, masaüstü veya web uygulamaları aracılığıyla erişmelerini sağlayan bir platform sunar. Kurum, bu bilgileri kullanarak, kurumsal stratejide belirler,

Scholtz, Calitz and Haupt'a göre, (2016, s.272) Caraianni ve Dumitrana (2005) tarafından önerilen gösterge panolarının tasarımı ve geliştirilmesi için altı adım aşağıda sıralanmaktadır:

- Hedefleri belirleyin.
- Görev, yeterlilik ve sorumlulukların listesini tanımlayın.
- Göstergeleri ayarlayın.

- Tanımlanan göstergelere göre bilgi toplayın.
- Gösterge panelini geliştirin.
- Gösterge panelini değerlendirin (Scholtz, Calitz and Haupt, 2016, s.271).

Tablo 3. BI ve Veri Analitiği Kapsamında Veri Mimarisi (Scholtz, Calitz and Haupt, 2016, s.271).

İlk derece kodları	Teorik kategoriler	Teorik kapsam
Dış kaynaklarda değerli verinin seçimi Yapısal olmayan formlardaki büyük miktardaki veri İlgili bilgi arama Veri tutarsızlığı ve eksikliği Hedef örneği anlama	Sağlama kapasitesi	Uygun kapasite
İş gereksinimlerine göre tanımlanmış sorular Veri yorumlama Veriyi anlaşılır kılma	Asimilasyon kapasitesi	
Yeni bilginin gerektirdiği yeni yeteneklerin yapılandırılması Anlaşılır formlarda dönüşüm Yeni bilginin yeni yetenekler gerektirdiğini içselleştirme	Dönüşüm kapasitesi	
Eylem bilgisi ve öngörüler Kara verme sürecinde iyileşme Pazar ve müşteri/kullanıcının anlaşılması Yeni ürün ya da hizmet geliştirme Toplanan verilerin tekrar kullanımı	İşletme kapasitesi	
Değişen veri gereksinimlerini izleyen teknolojik gelişmeler Yeni alt yapı, mimari ve programlar Performans, ölçülebilirlik ve zamanında karşılama	Teknolojik değerler	İş zekâsı değeri
İnsan katılımı gerektiren veriyi anlamlandırma Kapsam bilgisi ile kombine edilmiş analitik bilgi ve beceri Uygun işgücünü bulma zorlukları	İnsan Değeri	
Güçlü sosyal taahhütler Bilginin yayılımı Bölümler arası işbirlikleri	İlişkisel Değer	

İş zekâsı uygulamaları kurumsal kararları ve performansı destekleyen önemli bir araçtır. BI sistemleri, Karar Verme Sistemleri (Decision Support Systems , DSS) ve İş Performansı Ölçüm Sistemleri (Business Performance Measurement Systems, BPMS) gibi büyük veritabanı mimarilerini de içerir (Khan and Quadri, 2012). Bu sistemler farklı yapılandırılmış ve yapılandırılmamış verileri entegrasyonu ile geliştirilebilir (Chen, Chiang and Storey, 2012). Ayrıca, Kurumsal Kaynak Planlama (Enterprise Resource Planning, ERP) platformları büyük verilere dayanarak doğru veritabanı tasarımını gerektiren bir iş zekâsı bileşeni olarak

düşünölmek durumundadır (Elragal, 2014). Arařtırmacılar hala, BI sistemlerinin başarılı bir şekilde benimsenmesi ve kullanılmasına yönelik stratejik ve taktiksel yaklaşımları tartışmaktadır (Katerina and Dimovski, 2019, ss.1-2).

6. Genel Deęerlendirme ve Sonu

Bilgi bilim özünde bilginin toplumsal yönleriyle ilgilenir. Bireysel bilginin toplumsal bilgiye dönüşümü, erişim ve kullanım ilkeleriyle ilgili hizmetlerin geliştirilmesine yönelir. Bilginin tekrar kullanımına dönük koşulları oluşturur. Bilginin yapı taşı verilerin etkin yönetimi, bilgi hizmetlerinde teknolojik araçların kullanımı, hatta ilgili teknolojilerin gelişimine katkı sağlama bilgi bilimin çalışma kapsamında yer alır. Bu kapsamda taşıyıcı kaynaklar (kitap, elektronik ortam vb.) bir araçtır. Asıl amaç ortamı ya da taşıyıcısı ne olursa olsun bütünlüğü ve geçerlilięi bozulmamış veri ya da bilgi içeriğini, zamanında ve gerektiğinde yetkilendirilmiş kişilere eriştirebilmek ve bütünlüğü bozulmadan korunmasını sağlamaktır.

Bilgi arařtırmaları kapsamına organizasyonlar, organizasyonlarda hizmetler ve süreçler, konu uzmanlıkları, coęrafi bilgi hizmetleri, tarihi arařtırmalar, bilgi teknolojileri yönetimi ve politikalar gibi pek çok alan ya da konu girebilmektedir. Giderek veri yönetimi ve veri analitięi uygulamaları bilgi arařtırmaları içerisinde aęırlığını artırmaktadır.

Özellikle 2010'lu yıllarla birlikte büyük veri çalışmaları bulut bilişim ve yapay zekâ uygulamalarının devreye girmesiyle hızla teorik arařtırmalardan teknik uygulamalara dönüşmeye başlamıştır. Bilgi ve belge merkezleri doğası gereęi en önemli veri kaynakları arasında yer almaktadırlar. Bilgi merkezlerinde büyük veri; kaynak satın alma kararları, kişiselleştirilmiş okuyucu hizmetleri, etkin süreç analizi ve bilgi hizmetlerinin geliştirilmesinde etkin olarak kullanılabilir. Bu kapsamda veri madencilięi ve veri yönetimi uygulamaları, veri görselleştirme ve analizi yardımcı araçlarından yararlanılması gerekmektedir. İnsan bilgisinin toplandıęı yer olarak bilgi merkezleri büyük veri arařtırmalarının en önemli potansiyel kaynakları arasında yer almaktadır.

Öte yandan kamusal verilerin toplumsal fayda sağlamasına dönük oluşan kamuoyu; sadece hesap verilebilirlik, şeffaflık açısından değil kaynakların etkin kullanımı, yerel yönetimlerin güçlendirilmesi, teknik ve sosyal gelişimin sağlanması açısından önemlidir. Gelecekte yönetim ve idari sistemler ile iş süreçleri ve bilgi teknolojisi entegrasyonunu sağlamaya dönük daha fazla yatırım yapılması gerekmektedir. Kurumsal bilgi yönetim sistemleri sadece bilgi ve belge kaynaklarının yönetimini değil veri yönetimi uygulamalarını da kapsamak durumundadır.

Geleneksel olarak el yazmaları ve basılı kaynaklardan kütüphane, arşiv ve bilgi/belge hizmetlerine, ardından bilginin kendisine ve aracı olarak elektronik hizmetlere yönelen bilgi bilimin, veriyi işlenebilir ve yönetilebilir bir meta olarak değerlendirmemesi düşünülemez. Yine modern yönetim sistemlerinde kurumsal işleyiş ve yazışmalar, idari bilgi sistemleri ve belge yönetimi, bilgi işleme ve iş zekâsı uygulamaları doğası gereği bilgi bilimin kapsamı alanında yer almak durumundadır. Geleceğe açılan pencerede farklılaşan uzmanlıklar bilgi hizmetlerinin kurumsal yapılarda nasıl örgütleneceği sorununu da gündeme getirmektedir. Günümüzde dağınık yapılanan kütüphane, dokümantasyon, arşiv/belge yönetimi, bilgi işlem, strateji, kalite süreç, veri yönetimi, bellek hizmetleri vb. başlıklarda örgütlenen hizmetlerin idari yapıda tek çatıda toplanması son derece önemli görülmektedir. Bu kapsamda Bilgi Sistemleri ve Veri Yönetimi ya da Bilgi ve Belge Yönetimi Daire Başkanlıklarının oluşturulmasında yarar vardır.

Çalışmada bilgi bilimin kavramsal çerçevesinde veri yönetimi ve veri analitiği uygulamaları ayrıntılı biçimde incelenmeye çalışılmıştır. Veri yönetimi ve veri analitiği başlı başına bir araştırma disiplini olmaktan çok farklı disiplinlerin kullanabileceği bir araştırma metodolojisi ortaya koymaktadır. Bilgi ve belge yönetimi alanında analitik yaklaşımlarda, hizmet geliştirme ve kalite aracı olarak çok daha yaygın kullanılabilmesi düşünülmektedir. Öte yandan veri yönetimi ve veri analitiği, bilgi merkezlerinde kullandığımız tüm diğer araçlardan ve yöntemlerden çok daha bizdendir. Çünkü veri kaynakları halihazırda çok büyük oranda bilgi merkezleridir. Biz veriyi sadece depolanabilir ve eriştirilmesi gereken bir kaynak olarak değil, işlememiz, analiz etmemiz, sentezlememiz ve dönüştürücü bilgi olarak sunmamız gereken bir hizmet olarak görmek durumundayız. Bunun için daha çok veriyi anlamaya, onu nasıl işleyebileceğimizi, analiz edebileceğimizi ve anlamlı bilgilere dönüştürebileceğimizi araştırmamıza gereksinimimiz vardır. İlgili konuda kuramsal içerik, ilişkili programlama dilleri, yaygın kullanılan paket programlar, örnek uygulamalar bilgi ve belge yönetimi alanının analitik ve uygulamalı çalışmalarının kapsamında daha çok yer almak durumundadır.

Kaynakça / References

- Abbott, D. (2008). What is digital curation. Erişim adresi: www.dcc.ac.uk/resources/briefing-papers/introduction-curation/what-digital-curation.
- AGIMO. (2013). AGIMO Big Data Strategy - Issues Paper. Erişim adresi: www.finance.gov.au/files/2013/03/Big-Data-Strategy-Issues-Paper1.pdf.
- Ajis, A.F., Baharin, S.H. (2019). Dark Data Management as frontier of Information. Erişim adresi: <https://ieeexplore.ieee.org/document/8743915>
- Ayankoya, K., Calitz, A.P., Greyling, J. (2014). Intrinsic relations between data science, big data, business analytics and datafication. *ACM International Conference Proceeding Series* 28 (2014), 192–198.
- Bates, M. (2015). The information professions: knowledge, memory, heritage. *Information Research*, 7 (1). Erişim adresi: www.informationr.net/ir/20-1/paper655.html#.WRx98jTVLM (accessed 9 February 2018).

- Boulton, G. (2014, December). Open data and the future of science. Septentrio conference series (no. 1). Erişim adresi: <https://septentrio.uit.no/index.php/SCS/article/view/3231>.
- Business Intelligence Improved by Data Mining Algorithms and Big Data Systems: An Overview of Different Tools Asslied in Industrial Research. *Computer Science and Information Technology* 7 (1): 1-21, 2019.
- CA. (2016). Canada Capitalizing on Big Data. Erişim adresi: http://www.sshrc-crsh.gc.ca/news_room-salle_de_presse/latest_news-nouvelles_recentes/big_data_consultation-donnees_massives_consultation-eng.aspx.
- Carole, L. P. (2008). Research practices and research libraries: Working toward high impact information services. Erişim adresi: <https://www.ideals.illinois.edu/bitstream/handle/2142/9742/researchpracticeslibraries.sst.pdf?sequence=3>.
- Chen, H., Chiang, R. H. L., & Storey, V. C. (2012). Business intelligence and analytics: From big data to big impact. *MIS Quarterly*, 36(4), 1165–1188.
- Clark, T, D, Jones, M, C ve Armstrong, C, P. (2007). The dynamic structure of management sussort systems: theory development, research focus, and direction. *MIS Quarterly*, 31 (3) (2007), ss. 579-615
- Chung, W. (2014). BizPro: Extracting and categorizing business intelligence factors from textual news articles. *International Journal of Information Management* 34: 272-284.
- CNSF. (2015). National Science Foundation China. Retrieved from <http://www.nsf.gov.cn/>.
- European Commission. 2014. Commission urges governments to embrace potential of big data. Erişim adresi: europa.eu/rapid/press-release_IP-14-769_en.htm.
- Cunningham, A. (2016). Describing archives in context: Peter J Scott and the Australian “series” system. In V. Lemieux (Ed.), *Building trust in information: Perspectives on the frontiers of provenance* (ss. 49–58). Berlin: Springer-Verlag, Research Triangle Park, NC.
- Çakın, İ. (2012). Bilgi profesyonellerinin eğitiminde 40 yıl: Hacettepe Üniver sitesinin lisans programındaki değişiklikler. *Türk Kütüphaneciliği*, 26(2), 262-290.
- Delen, D. Moscato, G., Toma, I.L. (2018). The impact of real-time business intelligence and advanced analytics on the behaviour of business decision makers. 2018 International Conference on Information Management and Processing (ICIMP). IEEE Xplore: 29 March 2018, DOI: 10.1109/ICIMP1.2018.8325840.
- Delen, D. (2014). *Real-World Data Mining: Asslied Business Analytics and Decision Making*. New Jersey: Pearson Education, Inc.
- Digital Curation Centre (2005). Digital curation and preservation: defining the research agenda for the next decad. Erişim adresi: www.dcc.ac.uk/sites/default/files/documents/data-forum/documents/events/Warwick_Introduction.pdf.
- Diebold, F. X. (2012). A personal perspective on the origin(s) and development of “big data”: The phenomenon, the term, and the discipline (Scholarly Paper No. ID 2202843). Social Science Research Network. Erişim adresi: http://papers.ssrn.com/sol3/papers.cfm?abstract_id=2202843.
- Dingwall, G., Marciano, R., Moore, R., & McLellan, E. (2007). From data to records: Preserving the geographic information system of the city of Vancouver. *Archivaria*, 64, 181–198.
- Duranti, L., & Michetti, G. (2015). The archival method: Rediscovering a research tradition. In A. Gilliland, A. Lau, & A. McKemmish (Eds.), *Research in the archival multiverse*. Clayton, Victoria, Australia: Monash University Publishing.
- Eckerson, W. (2010). *Performance Dashboards: Measuring Monitoring and Managing Your Business*. John Wiley & Sons, Inc.
- Elragal, A. (2014). ERP and big data: the inept couple, *Procedia Technology*, Vol. 16, 242-249.
- Eroshkin, S.Yu., Kameneva, N.A., Kovkov, D.V., Sukhoruko A.I. (2017). Conceptual system in the modern information management *Procedia Computer Science* 103 (2017) 609 – 612. II. International Symposium Intelligent Systems, INTELS 16, 5-7 October 2016, Moscow.

- European Commission. (2014). Commission urges governments to embrace potential of big data. Erişim adresi: europa.eu/rapid/press-release_IP-14-769_en.htm.
- Fairfielda, J. and Shteina, H. (2014). Big data, big problems: Emerging issues in the ethics of data science and journalism. *Journal of Mass Media Ethics* 29, 1 (2014), 38–51.
- Fan, J., Han, F., & Liu, H. (2014). Challenges of big data analysis. *National Science Review*, 1(2), 293–314.
- Feldman, R. (2013). Techniques and applications for sentiment analysis. *Communications of the ACM*, 56(4), 82–89.
- Ferran, N., Mor, E., & Minguillón, J. (2005). Towards personalization in digital libraries through ontologies. *Library Management*, 26(4/5), 206–217.
- Foley, E. ve Guillemette, M. G. (2010). What is Business Intelligence. *International Journal of Business Intelligence Research*, vol. 1, no. 4.
- Fresa, A., Justrell, B., Prandoni, C. (2015). Digital curation and quality standards for memory institutions: PREFORMA research Project. *Archival Science* (2015) 15: 191–216.
- Gandomi, A. Haider, M. (2015). Beyond the hype: Big data concepts, methods, and analytics. *International Journal of Information Management*, 35 (2): 137-144.
- Gopalakrishnan, S., Kessler, E., & Scillitoe, J. (2010). Navigating the innovation landscape: past research, present practice, and future trends. *Organ. Manag. J.*, 7: 262–277.
- Hahn, U., & Mani, I. (2000). The challenges of automatic summarization. *Computer*, 33(11), 29–36.
- Heard, J., & Marciano, R. (2011). A system for scalable visualization of geographic archival records. In *Proceedings of the 1st IEEE Symposium on Large-Scale Data Analysis and Visualization* (ss. 121–122). Erişim adresi: <http://www.slideshare.net/richardjmarciano/a-system-for-scalablevisualization-of-geographic-archival-records/>
- Hey, T. , Tansley, S. and Tolle, K. (2009). *The Fourth Paradigm: Data-Intensive Scientific Discovery*. Erişim adresi: <http://research.microsoft.com/en-us/collaboration/fourthparadigm/>: 2032–2033.
- Higgins, S. (2017). Time to become our own profession?. *Digital Preservation Coalition Blog*. Erişim adresi: www.dpconline.org/blog/time-to-become-our-own-profession
- Higgins, S. (2018). Digital curation: the development of a discipline within information science. *Journal of Documentation*, 74 (6): 1318-1338.
- Hirtle, P.B. (2008). The history and current state of digital preservation in the United States, in Westbrooks, E. and Jenkins, K. (Eds), *Metadata and Digital Collections: A Festschrift in*
- Huang, R. H., & Li, B. Y. (2016). Data literacy education: Expansion of information literacy instruction in the big data era. *Document, Information & Knowledge*, 1, 21–27.
- Jiang, J. (2012). Information extraction from text. In C. C. Aggarwal, & C. Zhai (Eds.), *Mining text data* (ss. 11–41). United States: Springer.
- Janssen, H. van der Voort, A. Wahyudi. (2012). Factors influencing big data decision-making quality”, *Journal of Business Research*, vol. 70, ss. 338-345.
- Johnson, A. P. (2002). *A Short guide to action research*. Boston: Allyn & Bacon.
- Katerina B., Dimovski, V. (2019). Business intelligence and analytics for value creation: The role of absorptive capacity. *International Journal of Information Management* (46):93-103
- Kelly, K. (2012). The quantified century. In *Quantified Self Conference*. Retrieved from <http://quantifiedself.com/conference/Palo-Alto-2012>.
- Khan, R. A. and S. M. K. Quadri. (2012). Business intelligence: an integrated approach, *Business Intelligence Journal*, Vol.5, No.1, 64-70.
- Kooiman, J. (2015E). *Fish for life, interactive governance for fisheries*. Amsterdam: Amsterdam University Press.

- Kourentzes, N., Petropoulos, F. and Trapero, J.R. (2014), "Improving forecasting by estimating time series structural components across multiple frequencies", *International Journal of Forecasting*, Vol. 30 No. 2, ss. 291-302.
- Kwon, O., Lee, N., & Shin, B. (2014). Data quality management, data usage experience and acquisition intention of big data analytics. *International Journal of Information Management*, 34(3), 387–394.
- Külcü, Ö. (2018). *Bilgi Kuramı ve Bilgi Yönetimi*. İstanbul. Hiper Yayın.
- Labrinidis, A., & Jagadish, H. V. (2012). Challenges and opportunities with big data. *Proceedings of the VLDB Endowment*, 5(12), 2032–2033.
- Laney, D. (2001). 3D data management: Controlling data volume, velocity and variety. 6(70). META Group Research Note.
- Lemieux, V. (2012). Using information visualization and visual analytics to achieve a more sustainable future for archives: A survey and critical analysis of some developments. *Comma*, 2012(2), 5570.
- Lemieux, V. (2014). Towards a 'third order' archival interface: Research notes on some theoretical and practical implications of visual explorations in the Canadian context of financial electronic records. *Archivaria*, 78, 53–93.
- Li, S., Jiao, F. Zhang, Y., Xu, X. (2019). Problems and Changes in Digital Libraries in the Age of Big Data From the Perspective of User Services. *The Journal of Academic Librarianship*. 45 (1): 22-30.
- Liu, B. (2012). Sentiment analysis and opinion mining. *Synthesis Lectures on Human Language Technologies*, 5(1), 1–167.
- Longbing, C. (2017). Data Science: A Comprehensive Overview. *ACM Computing Surveys*. 50 (3): 43.1-43.42
- Loukides, M. (2011). *The Evolution of Data Products*. O'Reilly, Cambridge.
- Marciano, R., Lemieux, V., Hedges, M., Esteva, M., Underwood, W., Kurtz, M. & Conrad, Muntean, M., Sabau, G., Bologa, A., Surcel, T. and Florea, A. (2013), "Performance dashboards for universities", 2nd International Conference on Manufacturing Engineering, Quality and Production Systems, ss. 206-211.
- Massaro, A., Vitti, V., Galiano, A., Morelli, A. (2019). Business Intelligence Improved by Data Mining Algorithms and Big Data Systems: An Overview of Different Tools Asslied in Industrial Research. *Computer Science and Information Technology* 7(1): 1-21
- Massaro, A., Vitti, V. Lisco, P., Galiano, A., Savino, N. (2019). A business intelligence platform implemented in a big data system embedding data mining: a case of study. *International Journal of Data Mining & Knowledge Management Process (IJDMP)* 9 (1): 1-20.
- Mills, G. E. (2003). *Action research: A guide for the teacher researcher*. NJ: Merrill/Prentice Hall.
- Ong, I., Siew, P. and Wong, S. (2011), "A five-layered business intelligence architecture", *Communications of the IBIMA*, ss. 1-11.
- Salo, D. (2010). Retooling libraries for the data challenge. *Ariadne*, 64.
- Scholtz, B., Calitz, A and Haupt, H. (2016). A business intelligence framework for sustainability information management in higher education. *International Journal of Sustainability in Higher Education* 19 (2): 266-290
- Shera, J. H. (1968). An epistemological foundation for library science. *FRONTIERS OF LIBRARIANSHIP- 193 SYRACUSE UNIVERSITY* (8), 7-&. Shepherd, S., Bunn, J., Flinn, A and Lomas, E. (2019). Open government data: critical information management perspectives. *Records Management Journal* 291(1/2): 152-167.
- Skeina, S. S. (2017). *The Data Science Design Manual*. New York: Springer.
- Sixin, X., Yayuan Y.,Jiang, Y. (2017). A New Governance Architecture for Government Information Resources based on Big Data Ecological Environment in China. 2017 IEEE International Symposium on Multimed. DOI 10.1109/ISM.2017.103: 524-530
- UK. (2016). UK Big Data. Erişim adresi: <http://www.rcuk.ac.uk/research/infrastructure/big-data/>.

- USD2D. (2016). US National Consortium for Data Science. Erişim adresi: data2discovery.org.
- UN. (2010). United Nation Global Pulse Projects. Erişim adresi: <http://www.unglobalpulse.org/>.
- USNSF. 2012. US Big Data Research Initiative. Erişim adresi: <http://www.nsf.gov/cise/news/bigdata.jsp>.
- Uthayasankar, S. Kamal, M., Irani, Z., Weerakkody, V. (2017). Critical analysis of Big Data challenges and analytical methods. *Journal of Business Research* (70), January 2017: 263-286
- Uzuner, Y. (2005). Özel eğitimden örneklerle eylem araştırmaları. Ankara Üniversitesi Eğitim Fakültesi Özel Eğitim Dergisi, 6(2), 1 – 12.
- TechAmerica Foundation's Federal Big Data Commission. (2012). Demystifying big data: A practical guide to transforming the business of Government. Erişim adresi: <http://www.techamerica.org/Docs/fileManager.cfm?f=techamerica-bigdatareport-final.pdf>
- Tonta, Y. (2012). Kütüphanecilik ve bilgi bilim eğitiminde gelişmeler ve program değişiklikleri. *Türk Kütüphaneciliği* 26 (2): 227-261)
- Uslu, M. (2016). Veri Biliminde Veri Analitiği Tipleri. Erişim adresi: <https://kod5.org/veri-biliminde-veri-analitiği-tipleri/>.
- Xu, S., Du, W., Wang, C., & Liu, D. (2017). The library big data research: Status and libraries in China. *Library Work and Study*, 7, 46–50.
- Wang, X. Y. (2015). Analysis of and reflection on the asslication of big data in university libraries. *Journal of Academic Library and Information Science*, 33(1), 55–60.
- Winling, L., Connolly, N., Nelson, R., & Marciano R. (2017). Integrating the Depression-era history of race and redlining. In I. Gregory & D. Lafreniere (Eds.), *Routledge handbook of spatial history*, (ss. 502–524). New York, NY: Routledge.
- Wu, K., Su, X. N., & Deng, S. H. (2013). Big data, cloud computing and user behavior analysis. *Digital Library Forum*. 6. *Digital Library Forum* (ss. 19–23).
- Wu, X. et al. (2008). Top 10 algorithms in data mining, *Knowl. Inf. Syst.*, v.14, 1–37.
- Wyk, J., Walt, I. (2018). A Research Data Management Readiness Training Toolkit. *International Data Week 2018, SciDataCon 2018 Conference*, 5-9 November 2018, Gaborone, Botswana.
- Yigitbasioğlu, O.M. and Velcu, O. (2012), “A review of dashboards in performance management: implications for design and research”, *International Journal of Accounting Information Systems*, Vol. 13 No. 1, ss. 41-59.
- Yiu, C. (2012). The Big Data Ossortunity. Erişim adresi: <http://www.policyexchange.org.uk/images/publications/thessortunity.pdf>.